# CLASSICAL *and* QUANTUM INFORMATION

DAN C. MARINESCU

GABRIELA M. MARINESCU

# Classical and Quantum Information*

Dan C. Marinescu and Gabriela M. Marinescu

November 18, 2010

*Copyright 2006, 2007, 2008, 2009, 2010 by Dan C. Marinescu and Gabriela M. Marinescu

*To Vera Rae*

# Contents

**1 Preliminaries** **14**
1.1 Elements of Linear Algebra . . . 17
1.2 Hilbert Spaces and Dirac Notations . . . 23
1.3 Hermitian and Unitary Operators; Projectors. . . . 27
1.4 Postulates of Quantum Mechanics . . . 34
1.5 Quantum State Postulate . . . 36
1.6 Dynamics Postulate . . . 42
1.7 Measurement Postulate . . . 47
1.8 Linear Algebra and Systems Dynamics . . . 50
1.9 Symmetry and Dynamic Evolution . . . 52
1.10 Uncertainty Principle; Minimum Uncertainty States . . . 54
1.11 Pure and Mixed Quantum States . . . 55
1.12 Entanglement; Bell States . . . 57
1.13 Quantum Information . . . 59
1.14 Physical Realization of Quantum Information Processing Systems . . . 65
1.15 Universal Computers; The Circuit Model of Computation . . . 68
1.16 Quantum Gates, Circuits, and Quantum Computers . . . 74
1.17 Universality of Quantum Gates; Solovay-Kitaev Theorem . . . 79
1.18 Quantum Computational Models and Quantum Algorithms . . . 82
1.19 Deutsch, Deutsch-Jozsa, Bernstein-Vazirani, and Simon Oracles . . . 89
1.20 Quantum Phase Estimation . . . 96
1.21 Walsh-Hadamard and Quantum Fourier Transforms . . . 102
1.22 Quantum Parallelism and Reversible Computing . . . 107
1.23 Grover Search Algorithm . . . 111
1.24 Amplitude Amplification and Fixed-Point Quantum Search . . . 123
1.25 Error Models and Quantum Algorithms . . . 130
1.26 History Notes . . . 132
1.27 Summary . . . 136
1.28 Exercises and Problems . . . 137

**2 Measurements and Quantum Information** **142**
2.1 Measurements and Physical Reality . . . 144
2.2 Copenhagen Interpretation of Quantum Mechanics . . . 147
2.3 Mixed States and the Density Operator . . . 149
2.4 Purification of Mixed States . . . 156
2.5 Born Rule . . . 158
2.6 Measurement Operators . . . 159
2.7 Projective Measurements . . . 161
2.8 Positive Operator Valued Measures (POVM) . . . 164
2.9 Neumark Theorem . . . 167
2.10 Gleason Theorem . . . 168
2.11 Mixed Ensembles and their Time Evolution . . . 172
2.12 Bipartite Systems; Schmidt Decomposition . . . 174

2.13 Measurements of Bipartite Systems . . . 176
2.14 Operator-Sum (Kraus) Representation . . . 182
2.15 Entanglement; Monogamy of Entanglement . . . 185
2.16 Einstein-Podolski-Rosen (EPR) Thought Experiment . . . 189
2.17 Hidden Variables . . . 193
2.18 Bell and CHSH Inequalities . . . 199
2.19 Violation of Bell Inequality . . . 202
2.20 Entanglement and Hidden Variables . . . 206
2.21 Quantum and Classical Correlations . . . 208
2.22 Measurements and Quantum Circuits . . . 210
2.23 Measurements and Ancilla Qubits . . . 214
2.24 Measurements and Distinguishability of Quantum States . . . 217
2.25 Measurements and an Axiomatic Quantum Theory . . . 221
2.26 History Notes . . . 223
2.27 Summary and Further Readings . . . 225
2.28 Exercises and Problems . . . 228

**3 Classical and Quantum Information Theory 230**
3.1 The Physical Support of Information . . . 233
3.2 Entropy; Thermodynamic Entropy . . . 236
3.3 Shannon Entropy . . . 241
3.4 Shannon Source Coding . . . 251
3.5 Mutual Information; Relative Entropy . . . 255
3.6 Fano Inequality; Data Processing Inequality . . . 259
3.7 Classical Information Transmission through Discrete Channels . . . 261
3.8 Trace Distance and Fidelity . . . 267
3.9 von Neumann Entropy . . . 269
3.10 Joint, Conditional, and Relative von Neumann Entropy . . . 274
3.11 Trace Distance and Fidelity of Mixed Quantum States . . . 275
3.12 Accessible Information in a Quantum Measurement; Holevo Bound . . . 282
3.13 No Broadcasting Theorem for General Mixed States . . . 292
3.14 Schumacher Compression . . . 295
3.15 Quantum Channels . . . 297
3.16 Quantum Erasure . . . 301
3.17 Classical Information Capacity of Noiseless Quantum Channels . . . 306
3.18 Entropy Exchange, Entanglement Fidelity, and Coherent Information. . . . 312
3.19 Quantum Fano and Data Processing Inequalities . . . 318
3.20 Reversible Extraction of Classical Information from Quantum Information . . 322
3.21 Noisy Quantum Channels . . . 324
3.22 Holevo-Schumacher-Westmoreland Noisy Quantum Channel Encoding Theorem 329
3.23 Capacity of Noisy Quantum Channels . . . 334
3.24 Entanglement-Assisted Capacity of Quantum Channels . . . 338
3.25 Additivity and Quantum Channel Capacity . . . 342
3.26 Applications of Information Theory . . . 345
3.27 History Notes . . . 347

3.28 Summary and Further Readings . . . 348
3.29 Exercises and Problems . . . 351

**4 Classical Error Correcting Codes 355**
4.1 Informal Introduction to Error Detection and Error Correction . . . 357
4.2 Block Codes. Decoding Policies . . . 359
4.3 Error Correcting and Detecting Capabilities of a Block Code . . . 363
4.4 Algebraic Structures and Coding Theory . . . 366
4.5 Linear Codes . . . 375
4.6 Syndrome and Standard Array Decoding of Linear Codes . . . 383
4.7 Hamming, Singleton, Gilbert-Varshamov, and Plotkin Bounds . . . 387
4.8 Hamming Codes . . . 392
4.9 Proper Ordering, and the Fast Walsh-Hadamard Transform . . . 394
4.10 Reed-Muller Codes . . . 400
4.11 Cyclic Codes . . . 405
4.12 Encoding and Decoding Cyclic Codes . . . 410
4.13 The Minimum Distance of a Cyclic Code; BCH Bound . . . 421
4.14 Burst Errors. Interleaving . . . 424
4.15 Reed-Solomon Codes . . . 427
4.16 Convolutional Codes . . . 438
4.17 Product Codes . . . 445
4.18 Serially Concatenated Codes and Decoding Complexity . . . 446
4.19 Parallel Concatenated Codes - Turbo Codes . . . 449
4.20 History Notes . . . 453
4.21 Summary and Further Readings . . . 454
4.22 Exercises and Problems . . . 457

**5 Quantum Error Correcting Codes 461**
5.1 Quantum Error Correction . . . 463
5.2 A Necessary Condition for the Existence of a Quantum Code . . . 468
5.3 Quantum Hamming Bound . . . 469
5.4 Scale-up and Slow-down . . . 470
5.5 A Repetitive Quantum Code for a Single Bit-flip Error . . . 471
5.6 A Repetitive Quantum Code for a Single Phase-flip Error . . . 478
5.7 The Nine Qubit Error Correcting Code of Shor . . . 483
5.8 The Seven Qubit Error Correcting Code of Steane . . . 485
5.9 An Inequality for Representations in Different Bases . . . 490
5.10 Calderbank-Shor-Steane (CSS) Codes . . . 494
5.11 The Pauli Group . . . 500
5.12 Stabilizer Codes . . . 503
5.13 Stabilizers for Perfect Quantum Codes . . . 512
5.14 Quantum Restoration Circuits . . . 515
5.15 Quantum Codes over $GF(p^k)$ . . . 518
5.16 Quantum Reed-Solomon Codes . . . 521
5.17 Concatenated Quantum Codes . . . 527

5.18 Quantum Convolutional and Quantum Tail-Biting Codes . . . . . 528
5.19 Correction of Time-Correlated Quantum Errors . . . . . 538
5.20 Quantum Error Correcting Codes as Subsystems . . . . . 541
5.21 Bacon-Shor Code . . . . . 544
5.22 Operator Quantum Error Correction . . . . . 549
5.23 Stabilizers for Operator Quantum Error Correction . . . . . 553
5.24 Correction of Systematic Errors Based on Fixed-Point Quantum Search . . . . 555
5.25 Reliable Quantum Gates and Quantum Error Correction . . . . . 557
5.26 History Notes . . . . . 560
5.27 Summary and Further Readings . . . . . 560
5.28 Exercises and Problems . . . . . 562

**6 Physical Realization of Quantum Information Processing Systems 565**
6.1 Requirements for Physical Implementations of Quantum Computers . . . . . 567
6.2 Cold Ion Traps . . . . . 573
6.3 First Experimental Demonstration of a Quantum Logic Gate . . . . . 583
6.4 Trapped Ions in Thermal Motion . . . . . 588
6.5 Entanglement of Qubits in Ion Traps . . . . . 590
6.6 Nuclear Magnetic Resonance - Ensemble Quantum Computing . . . . . 596
6.7 Liquid-State NMR Quantum Computer . . . . . 598
6.8 NMR Implementation of Single-Qubit Gates . . . . . 605
6.9 NMR Implementation of Two-Qubit Gates . . . . . 606
6.10 The First Generation NMR Computer . . . . . 612
6.11 Quantum Dots . . . . . 614
6.12 Fabrication of Quantum Dots . . . . . 621
6.13 Quantum Dot Electron Spins and Cavity QED . . . . . 624
6.14 Quantum Hall Effect . . . . . 628
6.15 Fractional Quantum Hall Effect . . . . . 631
6.16 Alternative Physical Realizations of Topological Quantum Computers . . . . . 641
6.17 Photonic Qubits . . . . . 643
6.18 Summary and Further Readings . . . . . 649

**7 Appendix. Observable Algebras and Channels 652**

**8 Glossary 688**

"I want to know God's thoughts... the rest are details. " Albert Einstein.

## Preface

A new discipline, Quantum Information Science, has emerged in the last two decades of the twentieth century at the intersection of Physics, Mathematics, and Computer Science. Quantum Information Processing (QIP) is an application of Quantum Information Science which covers the transformation, storage, and transmission of quantum information; it represents a revolutionary approach to information processing

We have witnessed the development of microprocessors, high-speed optical communication, high-density storage technologies, followed by the widespread use of sensors, and more recently multi- and many-core processors and spintronics technology. We are now able to collect humongous amounts of information, process the information at high speeds, transmit the information through high-bandwidth and low-latency channels, store it on digital media, and share it using numerous applications built around the World Wide Web. Thus, the full cycle at the heart of information revolution was closed, Figure 1 [285], and this revolution became a reality that profoundly affects our daily life.

Now, at the beginning of the twenty first century, information processing is facing new challenges: heat dissipation, leakage, and other physical phenomena limit our ability to build

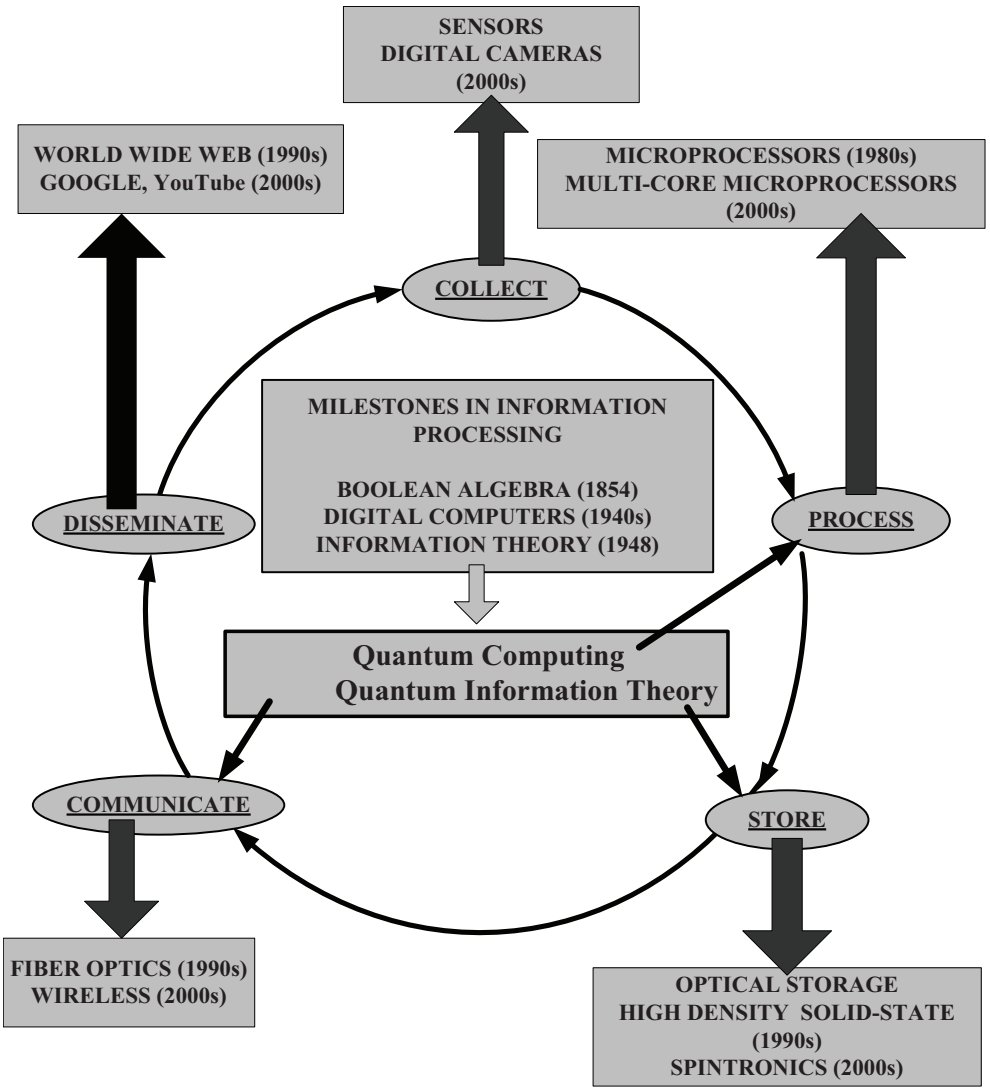



Figure 1: Our ability to collect, process, store, communicate, and disseminate information has increased considerably during the last two decades of the twentieth century. 1980s was the decade of microprocessors; advances in solid state technologies allowed the increase of the number of transistors on a chip by three order of magnitude and a substantial reduction of the cost of a microprocessor. In 1990s we have seen major breakthroughs in optical storage, high density solid-state storage technologies, fiber optics communication, and the widespread acceptance of the Word Wide Web. The first decade of the twenty first century is the decade of sensors, rapid information dissemination, and multi-core microprocessors.

increasingly faster and, implicitly, increasingly smaller solid-state devices; it is very difficult to ensure the security of our communication; we are overwhelmed by the volume of information we are bombarded with, and it is increasingly more difficult to extract useful information from the vast ocean of information surrounding us.

Information, either classical or quantum, is physical; this is the mantra repeated throughout the book. Therefore, we must understand the physical processes that affect the state

of the systems used to carry information. The physical processes for the storage, transformation, and transport of classical information are governed by the laws of classical Physics which limit our ability to process information increasingly faster using present day solid-state technologies. The speed of charge carriers in semiconductors is finite; to increase the speed of the device we have to pack the logic gates as tightly as possible.

The heat dissipated by a device increases with the clock rate to the power of 2 or 3, depending upon the solid-state technology. Heat removal is a hard problem for densely packed devices; the heat produced by a solid-state device is proportional to the number of gates thus, to the volume of the device. If we pack the gates into a sphere, the heat dissipated is proportional to the volume of the sphere and can be removed through the surface of the sphere; while the amount of the heat increases as the cube of the radius, our ability to remove it only increases as the square of the radius of the sphere. We are thus limited in our ability to increase the speed and density of classical circuits.

These facts provide a serious motivation to search for alternative physical realization of computing and communication systems. Scientists are now exploring revolutionary means to overcoming the limitations of computing and communication systems based upon the laws of classical Physics. Quantum and biological information processing provide a glimpse of hope in overcoming some of the limitations we mentioned and could revolutionize computing and communication in the third millennium. DNA computing together with quantum computing and quantum communication are the most promising avenues explored nowadays. While a significant progress has been made in understanding the properties of quantum information, fundamental questions regarding biological information are still waiting for answers. For example,how to explain the semantic aspect of biological information; how is information from a damaged region of the brain recovered?

Quantum information is information stored as a property of a quantum system e.g., the polarization of a photon, or the spin of an electron. Quantum information can be transmitted, stored, and processed following the laws of Quantum Mechanics. Several physical embodiments of quantum information are possible; for example, quantum communication involves a source that supplies quantum systems in a given state, a noisy channel that "transports" the quantum system, and the recipient that receives and decodes the quantum information. The source could be a laser producing monochromatic photons, the channel could be an optical fiber and the recipient a photocell; the source could also be an ion trap controlled by laser pulses, the channel a series of trapped ions, and the receiver a photo detector reading out the state of the ions via laser-induced fluorescence [276]. The diversity of the processes and technologies to process quantum information gives us hope that practical applications of quantum information will emerge sooner rather than later.

The physical processes for photonic, ion-traps, quantum dots, NMR, and other quantum systems are very different and could distract us from the goal of discovering the common properties of quantum information independent of its physical support. To study the properties of quantum information we use an abstract model which captures the critical aspects of quantum behavior; this model, Quantum Mechanics, describes the properties of physical systems as entities in a finite-dimensional Hilbert space. Therefore, quantum information theory requires a basic understanding of Quantum Mechanics and familiarity with the mathematical apparatus used by Quantum Mechanics and information theory.

Quantum information has special properties: the state of a quantum system cannot be

measured or copied without disturbing it; the quantum state of two systems can be *entangled*, the two-system ensemble has a definite state, though neither individual system has a well defined state of its own; we cannot reliably distinguish non-orthogonal states of a quantum system. Charles Bennett noted that "Speaking metaphorically, quantum information is like the information in a dream: attempting to describe your dream to someone else changes your memory of it, so you begin to forget the dream and remember only what you said about it." [49].

The properties of quantum information are remarkable and could be exploited for information processing: in quantum computing systems an exponential increase in parallelism requires only a linear increase in the amount of space needed thus, in principle, a quantum computer will be able to solve problems that cannot be solved with today's computers; reversible quantum computers avoid logically irreversible operations and can, in principle, dissipate arbitrarily little energy for each logic operation. Quantum information theory allows us to design algorithms for quantum key distribution and for quantum teleportation. Eavesdropping on a quantum communication channel can be detected with very high probability.

*Decoherence,* the randomization of the internal state of a quantum computer due to interactions with the environment, is a major problem in quantum information processing; quantum computers rely on undisturbed evolution of quantum coherence. Quantum error correction allows reliable communication over noisy quantum channels, provided that the channels are not too noisy. We should caution the reader that the complexity of the circuits involved in quantum error correction is far beyond today's technological possibilities; a fault-tolerant implementation of Shor's quantum factoring algorithm would most likely require thousands of physical qubits, at least two orders of magnitude more qubits than the systems reported in the literature have been able to harness. It may be possible though to resort to techniques which exploit the specific properties of individual physical realizations of quantum devices to manage the complexity of the quantum circuits for fault-tolerant systems. Fault-tolerant quantum computing still requires many more years of research.

Quantum information processing involves several areas including: quantum algorithms, quantum complexity theory, quantum information theory, quantum error correcting codes, quantum cryptography, and quantum reliability. This book covers basic concepts in quantum computing, quantum information theory, and quantum error correcting codes.

Classical information theory is a mathematical model for the transfer, storage and processing of information based on the laws of classical Physics. In the late 1940s Claude Shannon proved that it is possible to reliably transmit information over noisy classical communication channels; this discovery triggered the search for classical error correcting codes and the first codes were discovered by Richard Hamming in the early 1950s. Error correction is a critical component of modern technologies for reliable transfer, storage and processing of classical information. Quantum information theory (QIT) combines classical information theory with Quantum Mechanics to model information-related processes in quantum systems. The foundations of quantum information theory were established in the late 1980s by Charles Bennett and others and the interest in quantum information increased dramatically in mid 1990s after Peter Shor and Andrew Steane showed that quantum error correction is feasible and, together with Robert Calderbank, demonstrated that good quantum error correcting codes exist.

New discoveries add to the excitement of quantum information science: topological quantum computing proposed by Kitaev in 1997 and further developed by Friedman, Kitaev,

Larsen, and Wang has the potential to revolutionize fault-tolerance; in 2005 Grover discovered the fixed-point quantum search. In 2008 Smith and Yard showed that communication is possible over zero capacity quantum channels and in 2009 Hastings provided an answer to one of the most important open question in quantum information theory showing that the minimum entropy output of a quantum communication channel is not additive. In 1999 Knill, Laflamme, and Viola reformulated quantum error correction and proposed to view quantum error correcting codes as subsystems where the information resides in noiseless subspaces rather than considering a quantum code a subspace of a larger Hilbert space; in 2004 Kribs, Laflamme, and Poulin proposed a unified approach to quantum error correction and extended the concept of noiseless subsystems and their work led to the introduction of operator quantum error subsystems. These theoretical developments are mirrored by advances in quantum communication e.g., applications of Quantum Cryptography are close to commercialization.

The book organization is summarized in Figure 2; we first discuss classical concepts and then, gradually, we move to the corresponding concepts for quantum information. We adopted this philosophy for several reasons. First, the classical concepts are easier to grasp. Natural sciences develop increasingly more accurate and, at the same time, more complex models of physical reality; the level of abstraction makes it harder to develop the intuition behind the formalism and it is more difficult to master the mathematical apparatus the models are based on. The second reason why we discuss first the classical concepts is because the targeted audience for this book are not physicists familiar with Quantum Mechanics, but the larger population of scientists, engineers, students, or ordinary people puzzled by the "strange" properties of quantum information. Some of them are familiar with the classical information theory concepts and with classical error correcting codes; for them the significant leap is to transpose their intuition, and knowledge to a different frame of reference.

We follow the same philosophy in the presentation of quantum algorithms; we analyze first quantum oracles, the easier to understand algorithms for "toy" problems proposed by Deutsch, Jozsa, Bernstein and Vazirani, and Simon followed by an in depth analysis of phase estimation and of Grover search algorithm. The chapter covering information theory starts with the thermodynamic and Shannon entropy and classical channels and then we introduce von Neumann entropy and quantum channels. We discuss first linear codes and gradually move to more sophisticated cyclic, convolutional, and other families of classical codes; similarly, we analyze first the Shor, Steane, and CSS quantum error correcting codes before introducing stabilizer and subsystem codes. We hope that the numerous examples will facilitate the understanding of the more abstract concepts introduced throughout the book and will make the book accessible to a larger audience. Whenever possible we use the traditional notations in the literature or in the original papers which introduced the basic concepts. This required a careful selection of characters and fonts; for example, an $2^n$-dimensional Hilbert space is denoted as $\mathcal{H}_{2^n}$, Shannon entropy is $H$, the parity check matrix of a code is H the Hadamard transform is $\mathbf{H}$ and the transfer matrix of a Hadamard gate is $H$.

The authors are indebted to several colleagues who have read the manuscript and have made many constructive suggestions. Among them special thanks are due to Professors Dan Burghelea from the Mathematics Department at Ohio State University, Eduardo Mucciolo from the Physics Department, and Pawel Wocjan from the Computer Science Department at University of Central Florida. Of course, the authors are responsible for the errors that, in spite of our efforts, may still be found in the text.

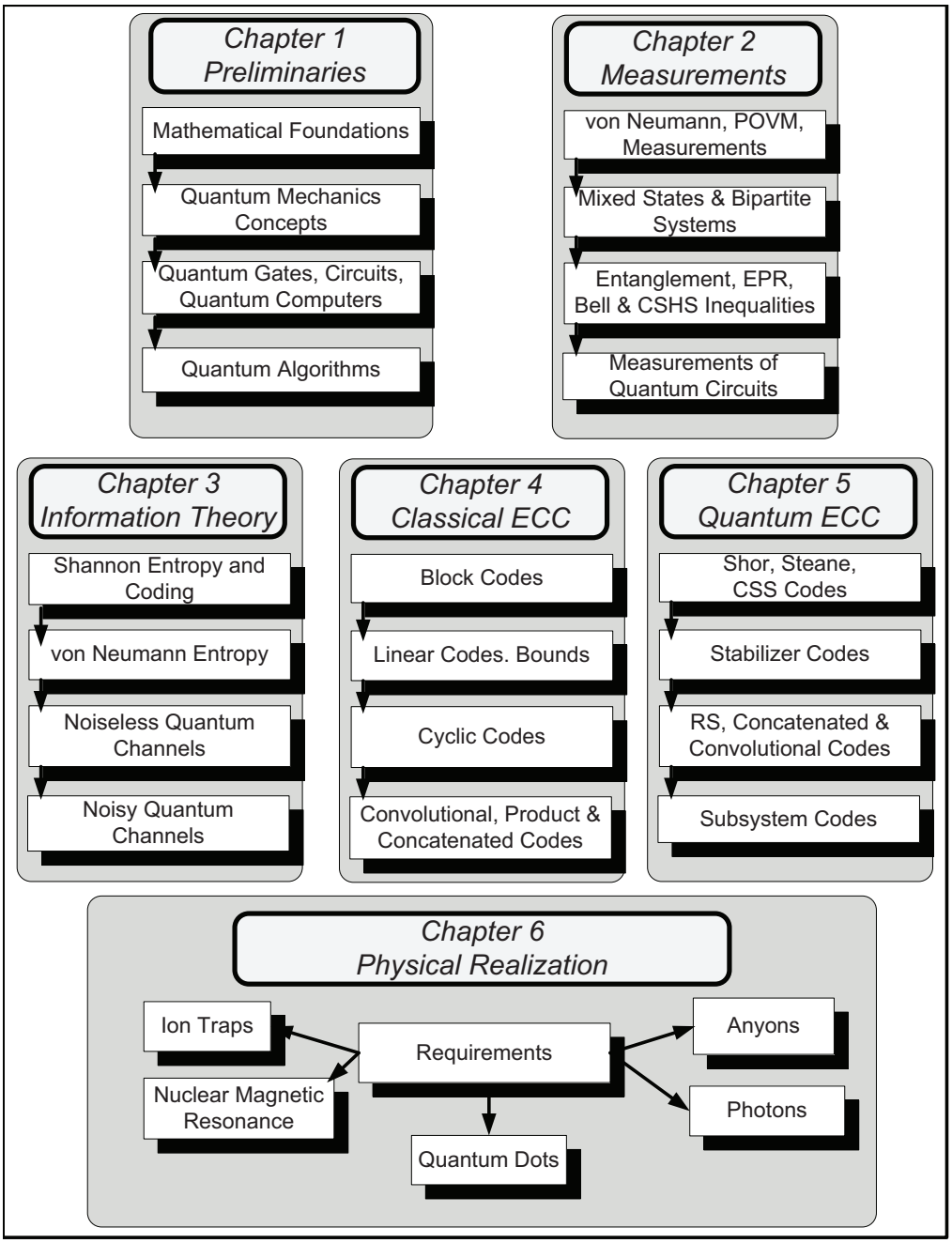



Figure 2: Book organization at a glance.

The artwork was created by George Dima, a gifted artist, concertmaster of the Bucharest Philharmonic and accomplished creator of computer-generated graphics (see http://picasaweb.google.com/degefe2008). We express our thanks to Patricia Osborne, Gavin Becker, and the editorial staff from Elsevier for their constructive suggestions.

"No amount of experiments can ever prove me right; a single experiment can prove me wrong." Albert Einstein.

# 1 Preliminaries

**What is information?** Carl Friederich von Weizsäcker's answer, *information is what is understood,* implies that information has a sender and a receiver who have a common understanding of the representation and the means to convey information using some properties of the physical systems [447]. He adds, "Information has no absolute meaning; it exists relatively between two semantic levels" [448].

Once asked the question *what is time,* Richard Feynman answered: "time is what happens when nothing else happens." Unfortunately, history did not record Feynman's answer to the question "what is information" and thus we do not have a crisp, witty, and insightful answer to a question central to the 21st century science. Indeed, the questions *what is information* and *what is its relationship with the physical world* become more important as we try to better understand physical phenomena at quantum scale and the behavior of biological systems.

It is easy to understand why there is no simple answer to the question we posed at the very beginning of this section; like matter and energy, information is a primitive concept thus,

it is rather difficult to rigorously define it. Informally, we can state that information abstracts properties of and allows us to distinguish among objects/entities/phenomena/thoughts; information is a common denominator for the very diverse contents of our material and spiritual world. There is a common expression of information as strings of bits, regardless of the objects/entities/processes/thoughts it describes. Moreover, these bits are independent of their physical embodiment. Information can be expressed using pebbles on the beach, mechanical relays, electronic circuits, and even atomic and subatomic particles.

*Classical information* is information encoded to some property of a physical system obeying the laws of classical Physics. Classical information is transformed using logic operations. Classical gates implement logic operations and allow for processing of classical information with classical computing devices.

*Quantum information* is information encoded to some property of quantum particles and obeys the laws of Quantum Mechanics. Quantum information is transformed using quantum gates, the building blocks for quantum circuits, which, in turn, can be assembled to build quantum computing and communication devices. The societal impact of information increases if the physical embodiments of bits and gates become smaller and we need less energy to process, store, and transmit information. This justifies our interest in quantum information.

**This book.** This book covers topics in quantum computing, quantum information theory, and quantum error correction, three important areas of quantum information processing. Quantum information theory and quantum error correction build on the scope, concepts, methodology, and techniques developed in the context of their close relatives, classical information theory and classical error correcting codes. It seems natural to follow the historical evolution of the concepts, and in this book, we first introduce the classical version of the concepts and techniques which are often simpler and easier to grasp, and then discuss in detail the significant leaps forward necessary to apply the concepts and techniques to quantum information

Information theory is a mathematical model for transmission and manipulation of classical information. Quantum information theory studies fundamental problems related to transmission of quantum information over classical and quantum communication channels such as: the entropy of quantum systems, the capacity of classical and quantum channels, the effect of the noise, fidelity, and optimal information encoding. Quantum information theory promises to lead to a deeper understanding of fundamental properties of nature and, at the same time, support new and exciting applications.

Error correcting codes allow us to detect and then correct errors during transmission of classical information over classical channels and to build fault-tolerant computing and communication systems which obey the laws of classical Physics. Quantum error correcting codes exploit the fundamental properties of quantum information investigated by quantum information theory and play an important role in the fault-tolerance of quantum computing and communication systems. Quantum error correcting codes are critical for the practical use of quantum computing and communication systems.

The first chapter of the book provides basic concepts from Mathematics, Quantum Mechanics, and Computer Science necessary for understanding the properties of quantum information. Then we discuss the building blocks of a quantum computer, the quantum circuits and quantum gates and survey some of the properties of quantum algorithms. Figure 3 provides a structured view of the topics covered in this chapter: (1) the mathematical apparatus

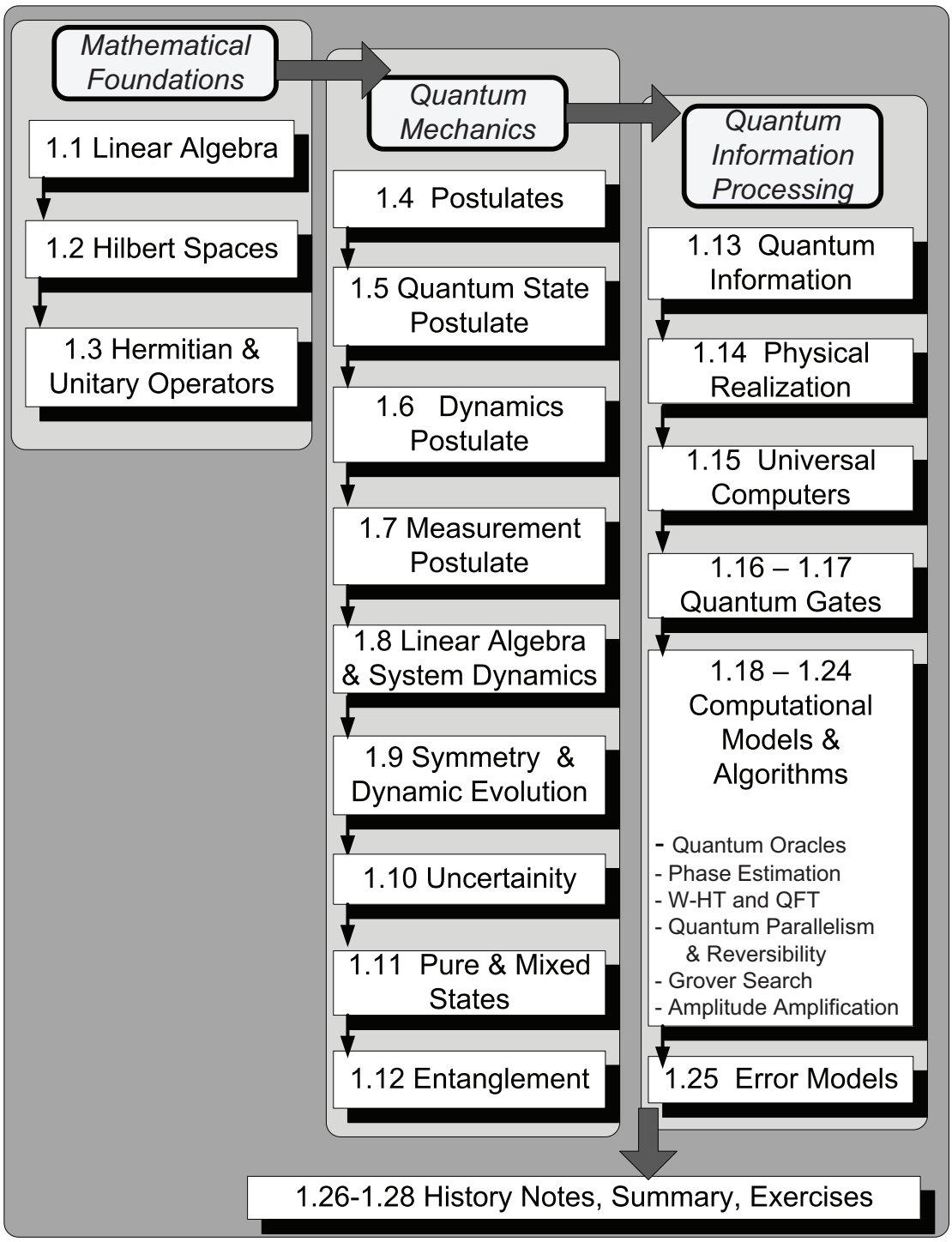



Figure 3: Chapter 1 at a glance.

used by Quantum Mechanics; (2) the fundamental ideas of Quantum Mechanics; (3) the circuits and algorithms for quantum computing devices.

## 1.1 Elements of Linear Algebra

Familiarity with complex numbers, algebraic structures such as groups, Abelian groups, and fields [58] and linear algebra [170] is required to understand the mathematical formalism of Quantum Mechanics. A review of algebraic structures used in coding theory is given in Section 4.4; in this section we review concepts such as vector spaces, inner product, norm, distance, orthogonality, basis, orthonormal basis, dimension of a vector space, linear transformation and matrices, eigenvectors and eigenvalues, and trace.

A vector space is an algebraic structure consisting of:

1. An Abelian group $(V, +)$ whose elements $\{\bar{v}_i\}$ are called "vectors" and whose binary operation "+" is called *addition*;

2. A field $\mathbb{F}$ of numbers whose elements are called "scalars"; we restrict $\mathbb{F}$ to be either $\mathbb{R}$ (the field of real numbers) or $\mathbb{C}$ (the field of complex numbers).

3. An operation called "multiplication with scalars" and denoted by "$\cdot$", which associates to any scalar $c \in \mathbb{F}$ and vector $\bar{v}_i \in V$ a new vector $\bar{v}_j = c \cdot \bar{v}_i \in V$. $\mathbb{F}$ acts linearly on $V$: if $a, b \in \mathbb{F}$ and $\bar{u}, \bar{v} \in V$ then $a \cdot (\bar{u} + \bar{v}) = a \cdot \bar{u} + a \cdot \bar{v}$ and $(a + b) \cdot \bar{u} = a \cdot \bar{u} + b \cdot \bar{u}$.

Assume that $\mathbb{F} \equiv \mathbb{C}$; it is easy to show that $\mathbb{C}^{m \times n}$, the set of all matrices $A = [a_{ij}]$ with entries $a_{ij} \in \mathbb{C}, 1 \leq i \leq n, 1 \leq j \leq m$ is a vector space where addition of two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is defined as $A + B = [a_{ij} + b_{ij}]$, the inverse with respect to addition of $A = [a_{ij}]$ is $-A = [-a_{ij}]$ and the identity element is $E = [0]$.

A set $\mathcal{B}$ of vectors is called a basis in $V$ if: (i) every vector $\bar{v} \in V$ can be expressed as a linear combination of vectors from $\mathcal{B}$; (ii) the vectors in $\mathcal{B}$ are linearly independent. The dimension of a vector space is the cardinality of $\mathcal{B}$. We consider only finite-dimensional vector spaces and in this case the cardinality is the number of elements of $\mathcal{B}$. An $n$-dimensional vector space will be denoted as $V_n$.

An inner product in the vector space $V_n$ over the field $\mathbb{F}$ is a mapping $g : V_n \times V_n \mapsto \mathbb{F}$ with several properties; $\forall \bar{v}_i, \bar{v}_j, \bar{v}_k \in V_n$ and $c \in \mathbb{F}$:

1. Obeys the addition rule in $V_n$:

$$g(\bar{v}_i + \bar{v}_j, \bar{v}_k) = g(\bar{v}_i, \bar{v}_k) + g(\bar{v}_j, \bar{v}_k) \quad \text{and} \quad g(\bar{v}_i, \bar{v}_j + \bar{v}_k) = g(\bar{v}_i, \bar{v}_j) + g(\bar{v}_i, \bar{v}_k).$$

2. Obeys the multiplication with a scalar rule in $V_n$:

$$g(c \cdot \bar{v}_i, \bar{v}_j) = c \times g(\bar{v}_i, \bar{v}_j) \quad \text{and} \quad g(\bar{v}_i, c \cdot \bar{v}_j) = c^* \times g(\bar{v}_i, \bar{v}_j)$$

   with $c^*$ the complex conjugate of $c$ when $\mathbb{F} \equiv \mathbb{C}$.

3. Satisfies the following relations

$$g(\bar{v}_i, \bar{v}_j) = g(\bar{v}_j, \bar{v}_i) \ \text{ if } \ \mathbb{F} = \mathbb{R} \quad \text{and} \quad g(\bar{v}_i, \bar{v}_j) = g^*(\bar{v}_j, \bar{v}_i) \ \text{ if } \ \mathbb{F} \equiv \mathbb{C}.$$

4. The inner product is non-degenerate, $g(\bar{v}_i, \bar{v}_i) \geq 0$ and $g(\bar{v}_i, \bar{v}_i) = 0$ if and only if $\bar{v}_i = 0$;

To simplify the notation the inner product $g(\bar{v}_i, \bar{v}_j)$ will be written as $\langle \bar{v}_i, \bar{v}_j \rangle$ and we shall use this notation from now on.

If an inner product in $V$ is provided, then the <u>norm</u> $|| \ \bar{v} \ ||$ of the vector $\bar{v} \in V_n$ is the square root of the inner product of the vector with itself:

$$|| \ \bar{v} \ ||= \sqrt{\langle \bar{v}, \bar{v} \rangle}.$$

The <u>distance</u> $d(\bar{v}_i, \bar{v}_j)$ of two vectors $\bar{v}_i, \bar{v}_j \in V_n$ is

$$d(\bar{v}_i, \bar{v}_j) =|| \ \bar{v}_i - \bar{v}_j \ ||= \sqrt{\langle (\bar{v}_i - \bar{v}_j), (\bar{v}_i - \bar{v}_j) \rangle}.$$

Two vectors, $\bar{v}_i, \bar{v}_j \in V_n$ are <u>orthogonal</u> if

$$\langle \bar{v}_i, \bar{v}_j \rangle = 0.$$

The vectors of $\{\bar{b}_1, \bar{b}_2, \ldots \bar{b}_n\} \in V_n$ form an <u>orthonormal basis</u> if the inner product of any two of them is zero, $\langle \bar{b}_i \bar{b}_j \rangle = 0, \ \forall (i,j) \in \{1, n\}, i \neq j$, and the norm of a vector is equal to unity, $|| \ \bar{b}_i \ ||= \langle \bar{b}_i, \bar{b}_i \rangle = 1, \ \forall i \in \{1, n\}$.

A <u>linear operator</u> $\mathbf{A}$ between two vector spaces $V$ and $W$ over the field $\mathbb{F}$ is any mapping from $V$ to $W$, $\mathbf{A} : V \mapsto W$, linear in its inputs

$$\mathbf{A}\left(\sum_i c_i \bar{v}_i\right) = \sum_i c_i \mathbf{A}(\bar{v}_i).$$

The identity operator $\mathbf{I}$ maps $\bar{v} \in V_n$ to itself, $\mathbf{I}(\bar{v}) = \bar{v}$. A <u>linear transformation</u> is a linear operator with $V = W$.

The <u>dual</u> of a vector space $V$, denoted as $V^*$, is the set of all scalar-valued linear maps $\varphi : V \mapsto \mathbb{F}$. If we define the addition and scalar multiplications in $V^*$ as

$$(\varphi + \phi)(\bar{v}) = \varphi(\bar{v}) + \phi(\bar{v}) \quad \text{and} \quad (c\varphi)(\bar{v}) = c\varphi(\bar{v})$$

with $\bar{v} \in V, c \in \mathbb{F}$ and $\varphi, \phi \in V^*$, then the dual is also a vector space over the field $\mathbb{F}$. If $V$ is an $n$-dimensional vector space so is $V^*$ and if the vectors $\{\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_j, \ldots, \bar{b}_n\}$ form a basis for $V_n$ then $V^*$ is also an $n$-dimensional vector space and the vectors $\{\bar{b}^1, \bar{b}^2, \ldots, \bar{b}^i, \ldots, \bar{b}^n\}$ defined by the property:

$$\bar{b}^i(\bar{b}^j) = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

form a basis for $V^*$. An inner product in $V$ provides an isomorphism, i.e., an invertible linear operator, $\mathbf{A} : V \mapsto V^*$, $\mathbf{A}(v)(w) = \langle v, w \rangle$.

If we choose a basis $\mathcal{B} = \{\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_n\}$ then the linear transformation $A$ is represented by an $n \times n$ <u>matrix</u>, $A = [a_{ij}], \ 1 \leq i, j \leq n$. Let the vectors $\bar{v}$ and $\bar{w}$ be

$$\bar{v} = \sum_i v_i \bar{b}_i \quad \text{and} \quad \bar{w} = \sum_i w_i \bar{b}_i$$

and $\bar{w} = A\bar{v}$. Then $w_i = \sum_{j=1}^{n} a_{ij} v_j$ which can be written as

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

with the right side of the equality representing the product of the $n \times n$ matrix whose elements are $a_{ij}$'s with the $n \times 1$ matrix whose elements are the $v_i$'s.

Addition and multiplication of matrices satisfy standard algebraic laws:

$$\begin{array}{lll} A + B = B + A & A + (B + C) = (A + B) + C & A(B + C) = AB + AC \\ A(BC) = (AB)C & (A + B)C = AC + BC & AI = IA = A \end{array}$$

with $I$ the identity matrix; $I_n$ is an $n \times n$ matrix with main diagonal elements equal to one and off-diagonal elements equal to zero. Non-zero matrices do not always have inverses and the product of two matrices is in general noncommutative, $AB \neq BA$.

The <u>determinant</u> of the $n \times n$ matrix $A$, $\det(A)$, is a number calculated from the elements of matrix $A$; it vanishes if and only if the matrix represents a linear transformation which is not one-to-one. The determinant can be written as a polynomial

$$\begin{vmatrix} a_1 & a_2 & a_3 & \cdots \\ b_1 & b_2 & b_3 & \cdots \\ c_1 & c_2 & c_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} = \sum_{ijk\cdots} \epsilon_{ijk\cdots}\, a_i b_j c_k \cdots$$

where $(ijk....)$ a permutation of indices $\{1, 2, 3, \ldots\}$ and

$$\epsilon_{ijk\cdots} = \left\{ \begin{array}{rl} 1 & \text{for even permutations,} \\ -1 & \text{for odd permutations.} \end{array} \right\}$$

The determinants of two $n \times n$ matrices $A$ and $B$ have the following property

$$\det(AB) = \det(A)\det(B).$$

An <u>eigenvector</u>, $\bar{v}$, of a linear transformation $\mathbf{A}$ is a non-zero vector such that

$$\mathbf{A}\bar{v} = \lambda\bar{v}.$$

The scalar $\lambda$ is called an <u>eigenvalue</u> corresponding to the eigenvector $\bar{v}$ of $\mathbf{A}$. The previous expression can be also written as

$$\mathbf{A}\bar{v} = \lambda\mathbf{I}\bar{v}.$$

Thus:

$$(\mathbf{A} - \lambda \mathbf{I})\bar{v} = 0.$$

This equation can be transformed to a matrix equation by choosing a basis $\{\bar{b}_i\}$, for $\bar{V}_n$. With respect to this basis $\bar{v}$ can be expressed as

$$\bar{v} = \sum_i c_i \bar{b}_i.$$

Then,

$$(\mathbf{A} - \lambda \mathbf{I}) \sum_i c_i \bar{b}_i = 0$$

where the coefficients must satisfy the equation

$$\sum_i (\mathbf{A} - \lambda \mathbf{I})_{i,j} c_i = 0 \quad \text{for any fixed} \;\; j.$$

A nontrivial solution exists only if the determinant

$$\det(\mathbf{A} - \lambda \, \mathbf{I}) = 0.$$

The scalar $\lambda$ for which that happens is an eigenvalue of $\mathbf{A}$. This condition becomes:

$$\begin{vmatrix} (a_{11} - \lambda) & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & (a_{22} - \lambda) & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & (a_{33} - \lambda) & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & (a_{nn} - \lambda) \end{vmatrix} = 0.$$

This is called the characteristic equation. The characteristic equation above is a polynomial of degree $n$ in $\lambda$, where $n$ is the dimension of the vector space. If $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$ then the polynomial has $n$ possibly complex numbers as roots and by the "fundamental theorem of algebra" can be expressed as a product of linear factors:

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) \cdots (\lambda_n - \lambda) = 0.$$

If $\mathbb{F} = \mathbb{C}$ then the $n$ roots, $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$ are the *eigenvalues* of the operator and are independent of the basis chosen to represent the operator as a matrix. If $\mathbb{F} = \mathbb{R}$ then the real roots are eigenvalues.

If the characteristic equation has less than $n$ distinct roots, there are *multiple roots*; the root $\lambda$ of multiplicity larger than one is said to be *multiple*. The *multiplicity $m$ of a root* $\lambda_i$ is the number of times the factor $(\lambda_i - \lambda)$ appears in the product above. For a multiple eigenvalue it is possible to have more than one eigenvector, all linearly independent and the corresponding linear space is of dimension $\leq m$.

If there are multiple eigenvalues of $\mathbf{A}$ and if for each multiple eigenvalue the multiplicity $m$ is equal to the dimension $d$ then it is possible to find a basis $\{\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_j, \ldots, \bar{b}_n\}$ for $\mathbf{A}$ such that each basis vector is an eigenvector of $\mathbf{A}$:

$$\mathbf{A}\bar{b}_1 = \lambda_1\bar{b}_1, \quad \mathbf{A}\bar{b}_2 = \lambda_2\bar{b}_2, \quad \mathbf{A}\bar{b}_3 = \lambda_3\bar{b}_3, \ \ldots \ , \mathbf{A}\bar{b}_n = \lambda_n\bar{b}_n.$$

Then $\mathbf{A}$ with respect to the basis $\{\bar{b}_1, \bar{b}_2, \ldots, \bar{b}_j, \ldots, \bar{b}_n\}$ can be expressed by the diagonal matrix

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

**Example.** *Let $\mathbb{F} = \mathbb{C}$ and consider the matrix:*

$$A = \begin{pmatrix} -1 & 2i \\ -2i & 2 \end{pmatrix}$$

representing a linear transformation $A : \mathbb{C} \mapsto \mathbb{C}^2$. $\lambda$ is an eigenvalue of $A$ if the determinant of the matrix:

$$A - \lambda I = \begin{pmatrix} -1-\lambda & 2i \\ -2i & 2-\lambda \end{pmatrix}$$

is zero. The resulting characteristic polynomial

$$\lambda^2 - \lambda - 6 = 0$$

has the roots $\lambda_1 = -2$ and $\lambda_2 = 3$. Hence, the eigenvectors of $A$ satisfy one or the other of the following systems of equations:

$$\begin{cases} -x_1 - 2iy_1 &= -2x_1 \\ 2ix_1 + 2y_1 &= -2y_1 \end{cases} \quad \text{or,} \quad \begin{cases} -x_2 - 2iy_2 &= 3x_2 \\ 2ix_2 + 2y_2 &= 3y_2 \end{cases}$$

where $x_1, x_2$ and $y_1, y_2$ represent the components corresponding to the basis vectors used to represent matrix $A$.

The equations can be rewritten as

$$\begin{cases} x_1 - 2iy_1 &= 0 \\ ix_1 + 2y_1 &= 0 \end{cases} \quad \text{or,} \quad \begin{cases} 2x_2 + iy_2 &= 0 \\ 2ix_2 - y_2 &= 0 \end{cases}$$

Solving these systems of equations we obtain the eigenvectors $(1, -\frac{1}{2}i)$ and $(1, 2i)$. These eigenvectors are not unique, $(\lambda, 2i\lambda)$ are also eigenvectors. These eigenvectors can be used as a new basis and the transformed matrix $A$ has a diagonal form:

$$\begin{pmatrix} -2 & 0 \\ 0 & 3-\lambda \end{pmatrix}$$

with respect to this basis. Now we review several important properties of square matrices with elements from $\mathbb{C}$ useful for the proof of the proposition presented at the end of this section.

The <u>trace</u> of a square matrix $A = [a_{ij}]\,,\ 1 \leq i, j \leq n,\ a_{ij} \in \mathbb{C}$, is the sum of the elements on the main diagonal of $A$, $\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$. From this definition it is easy to prove several properties of the trace.

(i) The trace is a linear map. Given a scalar $c \in \mathbb{R} \mid \mathbb{C}$ then $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(cA) = c\ \mathrm{tr}(A)$. As a consequence, if $A(t)$ is a matrix-valued function and $\frac{d}{dt}[A(t)]$ denotes the matrix whose entries are the derivatives of the entries of $A(t)$ then

$$\frac{d}{dt}\left[\mathrm{tr}(A(t))\right] = \mathrm{tr}\left(\frac{d}{dt}[A(t)]\right).$$

(ii) The trace is invariant to transposition, $\mathrm{tr}(A) = \mathrm{tr}(A^T)$. Indeed, the diagonal elements $a_{ii}$ of a square matrix $A = [a_{ij}]$ are invariant to transposition.

(iii) The trace is not affected by the order of the matrices in a product of two matrices

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

If $A = [a_{ij}]$ and $B = [b_{ij}]\,,\ 1 \leq i, j \leq n$, the diagonal elements of the products $AB$ and $BA$ are $\sum_{i=1}^{n} a_{ki}b_{ik}$ and $\sum_{i=1}^{n} b_{ki}a_{ik}$; thus, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$. Consequently if $U$ is a square $n \times n$ matrix and it is invertible then

$$\mathrm{tr}(U^{-1}AU) = \mathrm{tr}(A).$$

Given three $n \times n$ matrices $A, B$, and $C$ we have $\mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$. Another consequence, if $\lambda_1, \lambda_2, \ldots \lambda_n$ are the eigenvalues of a square matrix $A$ then

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i.$$

(iv) If $A$ is a square matrix and $A^*$ is its complex conjugate, i.e., the entry of index $ij$ is the complex conjugate of the entry $ji$, then

$$\mathrm{tr}(A^*A) \geq 0$$

Indeed,

$$\mathrm{tr}(A^*A) = \sum_{1 \leq i,j \leq n} a_{ij}a^*_{ji} = \sum_{1 \leq i,j \leq n} \mid a_{ij} \mid^2 .$$

It follows immediately that

$$\mathrm{tr}(A^*A) = 0 \quad \Leftrightarrow \quad A = 0.$$

From properties (i)-(iv) it follows that *the set of $n \times n$ square matrices with elements from $\mathbb{C}$ form a vector space with the inner product defined as*

$$\langle A, B \rangle = \mathrm{tr}(A^*B).$$

(v) The determinant and the trace are related

$$\det(I + \epsilon A) = 1 + \epsilon \times \mathrm{tr}(A) + o(\epsilon^2).$$

## 1.2 Hilbert Spaces and Dirac Notations

In this section we recast the linear algebra concepts using the formalism introduced by Paul Dirac for Quantum Mechanics. Hilbert spaces could be of infinite dimension but we shall only discuss finite-dimensional Hilbert spaces.

An n-dimensional Hilbert space $\mathcal{H}_n$ is an $n$-dimensional vector space over the field of complex numbers $\mathbb{C}$, equipped with an inner product. In this case, the isomorphism $A$ considered above, will be called Hermitian conjugation and denoted as $()^\dagger$. The state of a physical system will be represented either as a vector in $\mathcal{H}_n$ and referred to as a ket vector or, equivalently, by its Hermitian conjugate and referred to as a bra vector. In this vector space distances and angles between vectors can be measured.

An example of a Hilbert space of dimension $n$ is $\mathbb{C}_n$ when equipped with the inner product:

$$\langle \bar{a}, \bar{b} \rangle = \sum_{i=1}^{n} a_i b_i^* \quad \text{with} \quad \bar{a} = (a_1, a_2, \dots a_n) \;\text{ and }\; \bar{b} = (b_1, b_2, \dots b_n).$$

This Hilbert space has a canonical orthonormal base:

$$(1, 0, \dots, 0, 0), (0, 1, 0, \dots, 0, 0), \dots, (0, 0, \dots, 1, \dots, 0), \dots, (0, 0, \dots, 0, 1).$$

All $n$-dimensional Hilbert spaces are isomorphic with $\mathbb{C}_n$ by isomorphisms which identify the inner products (isometries). More precisely, if $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_i, \dots, \bar{v}_n$ is an orthonormal base, then the unique linear operator $\mathbf{A} : \mathcal{H}_n \mapsto \mathbb{C}_n$ defined by $\mathbf{A}(\bar{v}_i) = (0, 0, \dots, 1, \dots, 0)$ and extended linearly to all $\bar{v}_i \in \mathcal{H}_n$ is such an isometry. Conversely, for any isometry $\mathbf{A} : \mathcal{H}_n \mapsto \mathbb{C}_n$, $\bar{v}_i = \mathbf{A}^{-1}(0, 0, 0, \dots, 1, \dots, 0)$ provides an orthonormal basis in $\mathcal{H}_n$. In view of these remarks we can restrict ourselves to $\mathbb{C}_n$ as the "standard" representation of an $n$-dimensional Hilbert space, $\mathcal{H}_n$.

Given two Hilbert spaces $\mathcal{H}_m$ and $\mathcal{H}_n$ over the same field $\mathbb{F}$, their tensor product $\mathcal{H}_m \otimes \mathcal{H}_n$ is an $(m \times n)$-dimensional Hilbert space, $\mathcal{H}_{mn}$. If $\{e_1, e_2, \dots, e_m\}$ is an orthonormal basis in $\mathcal{H}_m$ and $\{f_1, f_2, \dots, f_n\}$ is an orthonormal basis in $\mathcal{H}_n$ then $\{e_i \otimes f_j\}, 1 \le i \le m, \;\; 1 \le j \le n$, is an orthonormal basis in $\mathcal{H}_{mn}$.

Quantum Mechanics uses a special notation for vectors; instead of the vector $\bar{v}$ we write $\mid v\rangle$ and call it the `ket` vector representation of $\bar{v}$; instead of the adjoint $v^\dagger$ we write $\langle v \mid$, and call it the `bra` vector representation of $\bar{v}$. With this convention the inner product $(v, w) = (w^\dagger, v)$ can also be written as $\langle w \mid v\rangle$.

The canonical orthonormal basis of $\mathbb{C}^n$ can be expressed as `ket` vectors as

$$\{\mid 0\rangle, \;\; \mid 1\rangle, \;\; \dots, \;\; \mid i\rangle, \;\; \dots, \;\; \mid n-1\rangle\},$$

or, as `bra` vectors as

$$\{\langle 0 \mid, \;\; \langle 1 \mid, \;\; \dots, \;\; \langle i \mid, \;\; \dots, \;\; \langle n-1 \mid\}.$$

We will also use the matrix representation of each `ket` vector $\mid i\rangle$ in this canonical representation as a column vector with a 1 in the $i^{th} + 1$ row and 0 in all the others. For example,

$$| 0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad | 1\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, | i\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, | n-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

and the matrix representation of the `bra` vector $\langle i \mid$ as a row vector with 1 in the $i^{th}+1$ position and 0 in all the others:

$$\begin{aligned} \langle 0 \mid &= \begin{pmatrix} 1 & 0 & \ldots & 0 & \ldots 0 \end{pmatrix}, \\ \langle 1 \mid &= \begin{pmatrix} 0 & 1 & \ldots & 0 & \ldots 0 \end{pmatrix}, \\ &\vdots \\ \langle i \mid &= \begin{pmatrix} 0 & 0 & \ldots & 1 & \ldots 0 \end{pmatrix}, \\ &\vdots \\ \langle n-1 \mid &= \begin{pmatrix} 0 & 0 & \ldots & 0 & \ldots 1 \end{pmatrix}. \end{aligned}$$

We have:

$$\langle i \mid j\rangle = \delta_{i,j} = \begin{cases} 0 & \text{if } \ i \neq j \\ 1 & \text{if } \ i = j \end{cases} \qquad 0 \leq (i,j) \leq n-1.$$

where $\delta_{i,j}$ is the *Kronecker delta* symbol.

An $n$-dimensional `ket` $\mid \psi\rangle$ can be expressed in this basis as a linear combination of the orthonormal basis `ket` vectors:

$$| \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle + \ldots + \alpha_i \mid i\rangle + \ldots + \alpha_{n-1} \mid n-1\rangle$$

where $\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots, \alpha_{n-1}$ are complex numbers.

The vector $\mid \psi\rangle$ can be expressed in different bases. For example, if $n = 2$ when the vector is expressed in the basis $\mid 0\rangle, \mid 1\rangle$ as

$$| \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$$

then in a new basis $\{\mid x\rangle, \mid y\rangle\}$ defined by

$$\{\mid x\rangle = \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle), \ \ \mid y\rangle = \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle)\}$$

the vector $\mid \psi\rangle$ is expressed as

$$| \psi\rangle = \frac{1}{\sqrt{2}}(\alpha_0 \ + \alpha_1) \mid x\rangle + \frac{1}{\sqrt{2}} \ (\alpha_0 - \alpha_1) \ \mid y\rangle.$$

For each `ket` vector $\mid \psi\rangle$ there is a *dual*, the `bra` vector denoted by $\langle \psi \mid$. As mentioned earlier the `bra` and `ket` vectors are related by Hermitian conjugation:

$$\mid \psi\rangle = (\langle\psi \mid)^{\dagger}, \qquad \langle\psi \mid = (\mid \psi\rangle)^{\dagger}.$$

The `bra` vector $\langle\psi \mid$ is expressed as a linear combination of the orthonormal `bra` vectors:

$$\langle\psi \mid = \alpha_0^*\langle 0 \mid + \alpha_1^*\langle 1 \mid + \ldots + \alpha_i^*\langle i \mid + \ldots + \alpha_{n-1}^*\langle n-1 \mid$$

where $\alpha_0^*, \alpha_1^*, \ldots, \alpha_i^*, \ldots, \alpha_{n-1}^*$ are the complex conjugates of $\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots, \alpha_{n-1}$.

In matrix representation, a `ket` vector is expressed as the column matrix:

$$\mid \psi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_{n-1} \end{pmatrix}$$

and a dual `bra` vector is expressed as the row matrix:

$$\langle\psi \mid = \begin{pmatrix} \alpha_0^* & \alpha_1^* & \ldots & \alpha_i^* & \ldots & \alpha_{n-1}^* \end{pmatrix}.$$

Recall that the inner product of two vectors $(\mid \psi\rangle, \mid \varphi\rangle)$ is a complex number. The inner product denoted as $\langle\psi \mid \varphi\rangle$ has the following properties:

1. The inner product of a vector with itself is a non-negative real number.
Indeed:

$$\langle\psi \mid \psi\rangle = \begin{pmatrix} \alpha_0^* & \alpha_1^* & \alpha_2^* \ldots \alpha_{n-1}^* \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \end{pmatrix}$$

$$= \mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 + \mid \alpha_2 \mid^2 + \ldots + \mid \alpha_{n-1} \mid^2$$

with $\mid \alpha_i \mid^2$, the square of the modulus of the complex number $\alpha_i$,

$$\mid \alpha_i \mid^2 = \{\Re e(\alpha_i)\}^2 + \{\Im m(\alpha_i)\}^2.$$

Thus, $\langle\psi \mid \psi\rangle \in \mathbb{R}$ and:

$$\langle\psi \mid \psi\rangle \begin{cases} = 0 & \text{if } \langle\psi \mid = (00\ldots0) \\ > 0 & \text{otherwise} \end{cases}$$

2. Linearity. If $\mid \psi\rangle, \mid \varphi\rangle, \mid \xi\rangle \in \mathcal{H}_n$ and $(a, b, c) \in \mathbb{C}$

$$\langle\psi \mid (c \mid \varphi\rangle) = c\langle\psi \mid \varphi\rangle;$$

$$(a\langle\psi \mid + b\langle\varphi \mid) \mid \xi\rangle = a\langle\psi \mid \xi\rangle + b\langle\varphi \mid \xi\rangle;$$

$$\langle \xi \mid (a \mid \psi\rangle + b \mid \varphi\rangle) = a\langle \xi \mid \psi\rangle + b\langle \xi \mid \varphi\rangle.$$

3. Hermitian symmetry:

$$\langle \psi \mid \varphi\rangle = \langle \varphi \mid \psi\rangle^*.$$

Let $\mathbf{A}$ and $\mathbf{B}$ be two liner operators represented as $m \times n$ and $p \times q$ matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ b_{31} & b_{32} & \dots & b_{3q} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pq} \end{pmatrix}$$

then, their <u>tensor product $\mathbf{A} \otimes \mathbf{B}$</u> is an $mp \times nq$ matrix defined as

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ a_{31}B & a_{32}B & \dots & a_{3n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}.$$

Here, $a_{ij}B, \quad 1 \le i \le m, \quad 1 \le j \le n$ is a sub-matrix whose entries are the products of $a_{ij}$ and all the elements of matrix $B$. Consistent with this definition the tensor product of two-dimensional vectors ($2 \times 1$ matrices), $(a, b)$ and $(c, d)$, is the four-dimensional vector (a $4 \times 1$ matrix)

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}.$$

**Example.** *The tensor product of vectors* $(\mid 0\rangle, \mid 1\rangle) \in \mathcal{H}_2$.

$$\mid 0\rangle \otimes \mid 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

An $m \times n$ matrix obtained as a product of an $m \times 1$ matrix (a `ket` vector $\mid \psi\rangle$) and a $1 \times n$ matrix (a `bra` vector $\langle \varphi \mid$) is sometimes called <u>outer product</u>. For example, let $\mid \psi\rangle, \mid \varphi\rangle \in \mathcal{H}_4$ be

$$\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle + \alpha_2 \mid 2\rangle + \alpha_3 \mid 3\rangle$$

$$\mid \varphi\rangle = \beta_0 \mid 0\rangle + \beta_1 \mid 1\rangle + \beta_2 \mid 2\rangle + \beta_3 \mid 3\rangle$$

Then the outer product $\mid \psi\rangle\langle\varphi \mid$ is the matrix

$$\mid \psi\rangle\langle\varphi \mid = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \begin{pmatrix} \beta_0^* & \beta_1^* & \beta_2^* & \beta_3^* \end{pmatrix} = \begin{pmatrix} \alpha_0\beta_0^* & \alpha_0\beta_1^* & \alpha_0\beta_2^* & \alpha_0\beta_3^* \\ \alpha_1\beta_0^* & \alpha_1\beta_1^* & \alpha_1\beta_2^* & \alpha_1\beta_3^* \\ \alpha_2\beta_0^* & \alpha_2\beta_1^* & \alpha_2\beta_2^* & \alpha_2\beta_3^* \\ \alpha_3\beta_0^* & \alpha_3\beta_1^* & \alpha_3\beta_2^* & \alpha_3\beta_3^* \end{pmatrix}.$$

We conclude with the observation that in the case of Hilbert spaces any root of the characteristic polynomial of a linear transformation $A$ is an eigenvalue. If an eigenvalue has multiplicity 1 the eigenvector is unique up to multiplication with a scalar but, if we insist that the eigenvector be unitary (of length 1), then the eigenvector is unique up to a phase factor. As we shall see in Sections 1.5, 1.6, and 1.7 vectors in Hilbert space represent the states of a quantum system; the transformation of states and the measurements are represented by linear operators.

## 1.3 Hermitian and Unitary Operators; Projectors.

A linear operator $\mathbf{A}$ maps vectors in a space $\mathcal{H}_n$ to vectors in the same space $\mathcal{H}_n$. Given the vectors $\mid \psi\rangle, \mid \varphi\rangle \in \mathcal{H}_n$ the transformation performed by the linear operator $\mathbf{A}$ can be described as

$$\mid \varphi\rangle = \mathbf{A} \mid \psi\rangle.$$

With respect to an orthonormal basis chosen once and for all, the linear operator $\mathbf{A}$ on $\mathcal{H}_n$ can be represented by the matrix

$$A = [a_{ij}], \quad \text{with} \quad a_{ij} \in \mathbb{C}, \quad 1 \leq i, j \leq n.$$

The adjoint of a linear operator $\mathbf{A}$ is denoted $\mathbf{A}^\dagger$. The matrix $A^\dagger$ describing the adjoint operator $\mathbf{A}^\dagger$ is the transpose conjugate matrix of $A$

$$A^\dagger = [a_{ji}^*], \quad 1 \leq i, j \leq n.$$

Throughout this book the term *operator* means *linear operator* and we shall use the terms operator $\mathbf{A}$ and matrix $A$ interchangeably. An operator acts on the left of a `ket` vector and on the right of a `bra` vector. Moreover,

$$[\mathbf{A} \mid \varphi\rangle]^\dagger = \mid \varphi\rangle^\dagger A^\dagger = \langle\varphi \mid \mathbf{A}^\dagger.$$

An operator $\mathbf{A}$ is normal if $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$. An operator $\mathbf{A}$ is Hermitian (self-adjoint) if $\mathbf{A} = \mathbf{A}^\dagger$. Clearly, a Hermitian operator is normal. The sum of two Hermitian operators $\mathbf{A} = \mathbf{A}^\dagger$ and $\mathbf{B} = \mathbf{B}^\dagger$ is Hermitian

$$(\mathbf{A} + \mathbf{B})^\dagger = \mathbf{A}^\dagger + \mathbf{B}^\dagger = \mathbf{A} + \mathbf{B}$$

while their product is not necessarily Hermitian. The necessary and sufficient condition for the product of two Hermitian operators to be Hermitian is $\mathbf{AB} = \mathbf{BA}$ or $\mathbf{AB} - \mathbf{BA} = 0$. The expression:

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$$

is called the <u>commutator</u> of $\mathbf{A}$ and $\mathbf{B}$. Thus, *the product of two Hermitian operators is a Hermitian operator if and only if their commutator is equal to zero.*

We now introduce a class of Hermitian operators of special interest in Quantum Mechanics; a <u>projector</u>, $\mathbf{P}$, is a Hermitian operator with the property: $\mathbf{P}^2 = \mathbf{P}$. The outer product of a state vector $\mid \varphi\rangle$ with itself is a projector: $\mathbf{P}_\varphi = \mid \varphi\rangle\langle\varphi \mid$. Two projectors $\mathbf{P}_i$, $\mathbf{P}_j$ are <u>orthogonal</u> if, for *every* state $\mid \psi\rangle \in \mathcal{H}_n$ the following equality holds

$$\mathbf{P}_i\mathbf{P}_j \mid \psi\rangle = 0$$

This condition is often written as

$$\mathbf{P}_i\mathbf{P}_j = 0.$$

A set of orthogonal projectors $\{\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2, \ldots\}$ is <u>complete/exhaustive</u> if

$$\sum_i \mathbf{P}_i = I.$$

It it easy to see that

$$(\mathbf{P}_\varphi)^\dagger = (\mid \varphi\rangle\langle\varphi \mid)^\dagger = \mid \varphi\rangle\langle\varphi \mid = \mathbf{P}_\varphi.$$

If $\mid \varphi\rangle$ is normalized, $\langle\varphi \mid \varphi\rangle = 1$, then $\mathbf{P}_\varphi^2 = \mathbf{P}_\varphi$.

$$(\mathbf{P}_\varphi)^2 = (\mathbf{P}_\varphi)^\dagger(\mathbf{P}_\varphi) = (\mid \varphi\rangle\langle\varphi \mid)(\mid \varphi\rangle\langle\varphi \mid) = \mid \varphi\rangle\langle\varphi \mid = \mathbf{P}_\varphi.$$

**Properties of Hermitian operators.** Hermitian operators enjoy a set of remarkable properties; some of them, *P1* - *P5* are discussed next.

*P1. The eigenvalues of a Hermitian operator in $\mathcal{H}_n$ are real.*

*Proof:* We calculate the inner product $\langle\psi \mid \mathbf{A} \mid \psi\rangle$ where $\lambda_\psi$ is the eigenvalue corresponding to the eigenvector $\mid \psi\rangle$ of $\mathbf{A}$:

$$\langle\psi \mid \mathbf{A} \mid \psi\rangle = \langle\psi \mid [\mathbf{A} \mid \psi\rangle] = \langle\psi \mid [\lambda_\psi \mid \psi\rangle] = \lambda_\psi\langle\psi \mid \psi\rangle$$

If we take the adjoint of $\mathbf{A} \mid \psi\rangle = \lambda_\psi \mid \psi\rangle$ and use the fact that $\mathbf{A}$ is Hermitian, $\mathbf{A} = \mathbf{A}^\dagger$, it follows that $\langle\psi \mid A = \lambda_\psi^*\langle\psi \mid$. We now rewrite the inner product:

$$\langle\psi \mid \mathbf{A} \mid \psi\rangle = [\langle\psi \mid \mathbf{A}] \mid \psi\rangle = \left[\lambda_\psi^*\langle\psi \mid\right] \mid \psi\rangle = \lambda_\psi^*\langle\psi \mid \psi\rangle.$$

It follows that

$$\lambda_\psi\langle\psi \mid \psi\rangle = \lambda_\psi^*\langle\psi \mid \psi\rangle \quad \Longrightarrow \quad (\lambda_\psi - \lambda_\psi^*)\langle\psi \mid \psi\rangle = 0, \quad \forall \mid \psi\rangle \in \mathcal{H}_n.$$

Since $\langle\psi \mid \psi\rangle \neq 0$, it follows that $\lambda_\psi = \lambda_\psi^*$ thus, the eigenvalue $\lambda_\psi$ is a real number.

*P2. Two eigenvectors $\mid \psi\rangle$ and $\mid \varphi\rangle$ of the Hermitian operator $\mathbf{A}$ with distinct eigenvalues, $\lambda_\psi \neq \lambda_\varphi$ are orthogonal, $\langle\psi \mid \varphi\rangle = 0$.*

*Proof:* We calculate the inner product $\langle\psi \mid \mathbf{A} \mid \varphi\rangle$ and use the fact that $\lambda_\varphi$ is an eigenvalue of $\mathbf{A}$ corresponding to the eigenvector $\mid \varphi\rangle$ thus, $\mathbf{A} \mid \varphi\rangle = \lambda_\varphi \mid \varphi\rangle$:

$$\langle\psi \mid \mathbf{A} \mid \varphi\rangle = \langle\psi \mid [\mathbf{A} \mid \varphi\rangle] = \langle\psi \mid [\lambda_\varphi \mid \varphi\rangle] = \lambda_\varphi\langle\psi \mid \varphi\rangle.$$

The eigenvalue $\lambda_\psi$ corresponds to the eigenvector $\mid \psi\rangle$ thus, $\mathbf{A} \mid \psi\rangle = \lambda_\psi \mid \psi\rangle$. If we take the adjoint of this expression and use the fact that $\mathbf{A}$ is Hermitian, $\mathbf{A} = \mathbf{A}^\dagger$, it follows that $\langle\psi \mid A = \lambda_\psi^*\langle\psi \mid$. We can now rewrite the inner product:

$$\langle\psi \mid \mathbf{A} \mid \varphi\rangle = [\langle\psi \mid \mathbf{A}] \mid \varphi\rangle = \left[\lambda_\psi^*\langle\psi \mid\right] \mid \varphi\rangle = \lambda_\psi^*\langle\psi \mid \varphi\rangle.$$

It follows that

$$\lambda_\varphi\langle\psi \mid \varphi\rangle = \lambda_\psi^*\langle\psi \mid \varphi\rangle \quad \forall\ \lambda_\psi \neq \lambda_\varphi.$$

This is possible if and only if: $\langle\psi \mid \varphi\rangle = 0$.

*P3. For every Hermitian operator* $\mathbf{A}$ *there exists a basis of orthonormal eigenvectors such that the matrix* $A$ *is diagonal in that basis and its diagonal elements are all the eigenvalues of* $\mathbf{A}$.

*Proof:* Let $\mid \psi_i\rangle$, $1 \leq i \leq n$, be an eigenvector of $\mathbf{A}$ and $\lambda_i$ the corresponding eigenvalue. If the eigenvalues are distinct, $\lambda_i \neq \lambda_j$ for $i \neq j$, $1 \leq i, j \leq n$, we construct the orthonormal basis as follows: we select an eigenvector $\mid \psi_j\rangle$ from a subspace that is orthogonal to the subspace spanned by $(\mid \psi_1\rangle, \mid \psi_2\rangle, \cdots, \mid \psi_i\rangle)$.

Then:

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & \\ 0 & \lambda_2 & 0 & \cdots & \\ 0 & 0 & \lambda_3 & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

If the eigenvalues $\lambda_j$ are degenerate (not distinct), then there are many bases of eigenvectors that diagonalize matrix $A$. If $\lambda$ is an eigenvalue of multiplicity $K > 1$, the set of corresponding eigenvectors generate a subspace of dimension $K$, the eigenspace corresponding to that $\lambda$; the eigenspaces corresponding to different eigenvalues are orthogonal. Assume that $\lambda$ is a degenerate eigenvalue and the corresponding eigenvectors are $\mid \psi_{\lambda_1}\rangle$ and $\mid \psi_{\lambda_2}\rangle$:

$$\mathbf{A} \mid \psi_{\lambda_1}\rangle = \lambda \mid \psi_{\lambda_1}\rangle \ \text{ and } \ \mathbf{A} \mid \psi_{\lambda_2}\rangle = \lambda \mid \psi_{\lambda_2}\rangle.$$

Then, $\forall a_1, a_2 \in \mathbb{C}$:

$$\mathbf{A}(a_1 \mid \psi_{\lambda_1}\rangle + a_2 \mid \psi_{\lambda_2}\rangle) = \lambda(a_1 \mid \psi_{\lambda_1}\rangle + a_2 \mid \psi_{\lambda_2}\rangle).$$

We can say that there is a whole subspace spanned by vectors $\mid \psi_{\lambda_1}\rangle$ and $\mid \psi_{\lambda_2}\rangle$, the elements of which are eigenvectors of $\mathbf{A}$ with eigenvalues $\lambda$.

*P4. The necessary and sufficient condition for two Hermitian operators,* $\mathbf{A}_1$ *and* $\mathbf{A}_2$*, to have the same set of eigenvectors is that they commute with each other:*

$$[\mathbf{A}_1, \mathbf{A}_2] = \mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1 = 0$$

*P5. The eigenvectors of a projector operator,* $\mathbf{P}_\psi = \mid \psi\rangle\langle\psi \mid$, *where* $\langle\psi \mid \psi\rangle = 1$, *are either perpendicular to or collinear to the vector* $\mid \psi\rangle$ *and their eigenvalues are* 0 *and* 1, *respectively.*

*Proof:* Assume $\mid \varphi\rangle$ is an eigenvector of $\mathbf{P}_\psi$ corresponding to eigenvalue $\lambda$:

$$\mathbf{P}_\psi \mid \varphi\rangle = \lambda \mid \varphi\rangle$$

or,

$$\mid \psi\rangle\langle\psi \mid \varphi\rangle = \lambda \mid \varphi\rangle.$$

The inner product $\langle\psi \mid \varphi\rangle$ is a number $\gamma$

$$\gamma \mid \psi\rangle = \lambda \mid \varphi\rangle$$

This implies that

$$\lambda = 0 \quad \text{if} \quad \gamma = 0, \quad \text{when} \quad \mid \psi\rangle \text{ and } \mid \varphi\rangle \text{ are perpendicular}$$

or,

$$\lambda = 1 \quad \text{if} \quad \gamma = 1, \quad \text{when} \quad \mid \psi\rangle \text{ and } \mid \varphi\rangle \text{ are parallel and normalized.}$$

An operator $\mathbf{A}$ is <u>unitary</u> if $\mathbf{A}_n\mathbf{A}_n^\dagger = \mathbf{A}_n^\dagger\mathbf{A}_n = \mathbf{I}_n$. Clearly, a unitary operator is normal. The product of unitary operators is unitary, but the sum is not; "product" in this case means composition, not to be confused with tensor product.

*A unitary operator* $\mathbf{A}$ *preserves the inner product,* thus, it preserves the distance in a Hilbert space. If $\mid \psi_1\rangle, \mid \psi_2\rangle, \mid \varphi_1\rangle, \mid \varphi_2\rangle \in \mathcal{H}_n$ and $\mid \varphi_1\rangle = \mathbf{A} \mid \psi_1\rangle$ and $\mid \varphi_2\rangle = \mathbf{A} \mid \psi_2\rangle$ then:

$$\langle\varphi_1 \mid \varphi_2\rangle = \langle\psi_1 \mid \psi_2\rangle$$

Indeed the inner product $\langle\varphi_1 \mid \varphi_2\rangle$ can be written as

$$\langle\varphi_1 \mid \varphi_2\rangle = [\mathbf{A} \mid \psi_1\rangle]^\dagger \mathbf{A} \mid \psi_2\rangle = \langle\psi_1 \mid \mathbf{A}^\dagger\mathbf{A} \mid \psi_2\rangle = \langle\psi_1 \mid \psi_2\rangle.$$

It follows immediately that a unitary operator $\mathbf{A}$ preserves the norm of a vector:

$$[\langle\varphi \mid A^\dagger]\ [A \mid \varphi\rangle] = \langle\varphi \mid [A^\dagger A] \mid \varphi\rangle = \langle\varphi \mid \varphi\rangle$$

If $\mathbf{A}$ is a unitary operator then we can undo its action, i.e., a unitary operator is invertible and $\mathbf{A}^{-1} = \mathbf{A}^\dagger$.

Every normal operator has a complete set of orthonormal eigenvectors. If $\mid e_i\rangle$ is an eigenvector (eigenstate) of the operator $\mathbf{A}$ and $\lambda_i$ is the associated eigenvalue, then we can write:

$$\mathbf{A} \mid e_i\rangle = \lambda_i \mid e_i\rangle.$$

The <u>spectral decomposition</u> of a normal operator $\mathbf{A}$ is

$$\mathbf{A} = \sum_i \lambda_i \, \mathbf{P}_i$$

with $\mathbf{P}_i = \mid e_i \rangle \langle e_i \mid$ the projector corresponding to the eigenvector $\mid e_i \rangle$.

*Proof:* Note that $(\mathbf{A} + \mathbf{A}^\dagger)/2$ and $(\mathbf{A} - \mathbf{A}^\dagger)/2i$ are Hermitian and commute. By *P3* and *P4* there exists an orthonormal base $\{e_1, e_2, \ldots, e_n\}$ of eigenvectors simultaneously for $(\mathbf{A}+\mathbf{A}^\dagger)/2$ and $(\mathbf{A} - \mathbf{A}^\dagger)/2i$. Hence they are also eigenvectors for

$$\mathbf{A} = \left( \frac{\mathbf{A} + \mathbf{A}^\dagger}{2} + i \frac{\mathbf{A} - \mathbf{A}^\dagger}{2i} \right).$$

Let $\lambda_1, \lambda_2, \ldots \lambda_n$ be the eigenvalues corresponding to the eigenvectors $\{e_1, e_2, \ldots, e_n\}$.

$$\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i \quad \text{with} \quad \mathbf{P}_i = \mid e_i \rangle \langle e_i \mid .$$

Indeed, both $\mathbf{A}$ and $\sum_i \lambda_i \mathbf{P}_i$ have the same matrix representation with respect to the base $\{e_1, e_2, \ldots, e_n\}$.

**Example.** *Choose an orthonormal basis in $\mathcal{H}_2$. Let $\mathbf{A}$ be a $2 \times 2$ matrix in this basis is*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

*Calculate its eigenvalues $\lambda_1$ and $\lambda_2$, possibly equal. Find linearly independent solutions of the two linear equations:*

$$A \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \lambda_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad \text{and} \quad A \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \lambda_2 \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}.$$

Define

$$v_1 = \begin{pmatrix} \frac{x_1}{\sqrt{|x_1|^2+|y_1|^2}} \\ \frac{y_1}{\sqrt{|x_1|^2+|y_1|^2}} \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} \frac{x_2}{\sqrt{|x_2|^2+|y_2|^2}} \\ \frac{y_2}{\sqrt{|x_2|^2+|y_2|^2}} \end{pmatrix}$$

The vectors $\mid v_1 \rangle$ and $\mid v_2 \rangle$ form another basis for $\mathbf{A}$ such that we can express for $\mathbf{A}$ as

$$\mathbf{A} = \lambda_1 \mid v_1 \rangle \langle v_1 \mid + \lambda_2 \mid v_2 \rangle \langle v_2 \mid .$$

Now we introduce the concept of positive operators and define the square root and the modulus of a positive operator; then we describe the canonical decomposition of an invertible operator [355].

**Positive operators.** An operator $\mathbf{V} : \mathcal{H}_n \mapsto \mathcal{H}_n$ is a <u>positive operator</u>, $\mathbf{V} > 0$, iff:

$$\langle \mathbf{V}\varphi \mid \varphi \rangle > 0, \qquad \forall \mid \varphi \rangle \in \mathcal{H}_n, \quad \mid \varphi \rangle \neq 0.$$

The eigenvalues of a positive definite operator are real and positive.

An operator is positive semi-definite if $\forall \mid \varphi\rangle \neq 0 \quad \langle \mathbf{V}\varphi \mid \varphi\rangle = \langle \varphi \mid \mathbf{V} \mid \varphi\rangle \geq 0$; the eigenvalues of $\mathbf{V}$ are real and non-negative thus, $\mathrm{tr}V \geq 0$. A positive semi-definite operator is self-adjoint. If two positive operators commute then their product is a positive operator.

If $\mathbf{V}$ is a positive semi-definite operator there exists a unique positive semi-definite operator $\mathbf{Q}$ such that $\mathbf{Q}^2 = \mathbf{V}$ denoted also by $\mathbf{Q} = \sqrt{\mathbf{V}}$. Indeed, there exits an orthonormal base in $\mathcal{H}_n$ consisting of eigenvectors of $\mathbf{V}$; with respect to this basis $\mathbf{V}$ can be represented as a diagonal matrix, $\mathrm{diag}\,(p_1, p_2, \ldots p_n)$. Since $\mathbf{V}$ is positive then $p_i \geq 0$, hence they have unambiguously defined positive square roots. Define:

$$\mathbf{Q} = \sqrt{\mathbf{V}} \quad \text{to be represented by} \quad \mathrm{diag}\,(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n})\,.$$

The modulus of any operator $\mathbf{V}$ is defined as

$$\mid \mathbf{V} \mid = \sqrt{\mathbf{V}^\dagger \mathbf{V}} \geq 0,$$

it is always positive. One can also consider $\sqrt{\mathbf{V}^\dagger \mathbf{V}} \geq 0$ but, unless $\mathbf{V}$ is normal, that is not the same as $\sqrt{(\mathbf{V}\mathbf{V}^\dagger)} \geq 0$.

The operator $\mathbf{Q}$ is invertible if:

$$\mid \mathbf{Q} \mid\mid \mathbf{Q} \mid^{-1} = \mathbf{I}$$

*If an operator* $\mathbf{Q}$ *is invertible then it can be written as a product of a positive-definite operator,* $\mathbf{V}$*, and a unitary operator,* $\mathbf{U}$*:*

$$\mathbf{Q} = \mathbf{V}\mathbf{U}, \qquad \text{with} \qquad \mathbf{V} = \mid \mathbf{Q} \mid = \sqrt{\mathbf{Q}\mathbf{Q}^\dagger} \quad \text{and} \quad \mathbf{U} = \mid \mathbf{Q} \mid^{-1}\mathbf{Q}.$$

From definition it follows immediately that $\mathbf{V}$ is positive as the modulus of $\mathbf{Q}$. To check that $U$ is unitary, observe that $\mid \mathbf{Q} \mid$ and $\mid \mathbf{Q} \mid^{-1}$ are self-adjoint hence $\mid \mathbf{Q} \mid^\dagger = \mid \mathbf{Q} \mid$ .
Then we have

$$\mathbf{U}\mathbf{U}^\dagger = \mid \mathbf{Q} \mid^{-1} \mathbf{Q}\mathbf{Q}^\dagger (\mid \mathbf{Q} \mid^{-1})^\dagger = \mid \mathbf{Q} \mid^{-1} \mathbf{Q}\mathbf{Q}^\dagger \mid \mathbf{Q} \mid^{-1} = \mid \mathbf{Q} \mid^{-1} \mid \mathbf{Q} \mid\mid \mathbf{Q} \mid\mid \mathbf{Q} \mid^{-1} = \mathbf{I}.$$

To prove that this decomposition is unique we assume that

$$\mathbf{Q} = \mathbf{V}_1\mathbf{U}_1 \quad \text{and} \quad \mathbf{Q} = \mathbf{V}_2\mathbf{U}_2.$$

Then, we use the fact that $\mathbf{U}_1$ and $\mathbf{U}_2$ are unitary, $\mathbf{U}_1\mathbf{U}_1^\dagger = \mathbf{U}_2\mathbf{U}_2^\dagger = \mathbf{I}$, to get:

$$\mathbf{Q}\mathbf{Q}^\dagger = [\mathbf{V}_1\mathbf{U}_1]\,[\mathbf{V}_1\mathbf{U}_1]^\dagger = \mathbf{V}_1\mathbf{U}_1\mathbf{U}_1^\dagger\mathbf{V}_1^\dagger = \mathbf{V}_1\mathbf{V}_1^\dagger = \mathbf{V}_\mathbf{1}^\mathbf{2}.$$

and

$$\mathbf{Q}\mathbf{Q}^\dagger = [\mathbf{V}_2\mathbf{U}_2]\,[\mathbf{V}_2\mathbf{U}_2]^\dagger = \mathbf{V}_2\mathbf{U}_2\mathbf{U}_2^\dagger\mathbf{V}_2^\dagger = \mathbf{V}_2\mathbf{V}_2^\dagger = \mathbf{V}_\mathbf{2}^\mathbf{2}$$

It follows that $\mathbf{V}_1 = \mathbf{V}_2$ and immediately that $\mathbf{U}_1 = \mathbf{U}_2$.

Another useful property of a positive operator $\mathbf{V}$ regards the trace of the product of $\mathbf{V}$ and a unitary operator $\mathbf{U}$. The maximum with respect to $\mathbf{U}$ of the trace of the product $\mathbf{V}\mathbf{U}$ is

$$\mu = \max_{\mathbf{U}} [\text{tr } (\mathbf{VU})] = \text{tr } (\mathbf{V}), \quad \text{if} \quad (\mathbf{VU}) \geq 0.$$

We show this only for the case when $\mathbf{V}$ is invertible and $\mathbf{V} =\mid \mathbf{V} \mid \mathbf{U}$. Let $p_1, p_2, \ldots, p_i, \ldots, p_n \geq 0$ be the eigenvalues of $\mid \mathbf{V} \mid$:

$$\mid \mathbf{V} \mid = \text{diag}\,[p_1, p_2, \ldots p_n]\,.$$

The elements of $U = [u_{ij}]$, the matrix associated with the unitary operator $\mathbf{U}$, cannot be larger than unity, $u_{ij} \leq 1$. It follows that

$$\text{tr } (\mathbf{VU}) =\mid \sum_i p_i u_{ii} \mid .$$

Thus the maximum is obtained when $u_{ii} = 1$ and $\mathbf{U} = \mathbf{I}$; then $\mu = \text{tr } (\mathbf{V})$, as stated.

A useful relation is the Schwarz inequality for the inner product of two operators (see Problem 1 at the end of this Chapter):

$$\text{tr}\left(\mathbf{VV}^\dagger\right) \text{tr}\left(\mathbf{QQ}^\dagger\right) \geq\mid \text{tr}\left(\mathbf{VQ}^\dagger\right) \mid^2 .$$

Last, we consider a Hilbert space $\mathcal{H}_n$ with an orthonormal basis $\{\mid e_1\rangle, \mid e_2\rangle, \ldots, \mid e_n\rangle\}$. Consider now a vector $\mid \varphi\rangle$ in the extended Hilbert space $\mathcal{H}_n \otimes \mathcal{H}_n$:

$$\mid \varphi\rangle = \sum_{i=1}^{n} \mid e_i\rangle \otimes \mid e_i\rangle.$$

Let $\mathbf{V}$ be any operator in $\mathcal{H}_n$ and let $\mathbf{V}^T$ be its transpose in $\mathcal{H}_n$. If we want to apply the operator $\mathbf{V}$ defined on $\mathcal{H}_n$ to the state vector $\mid \varphi\rangle$ defined on $\mathcal{H}_n \otimes \mathcal{H}_n$, it is easy to see that we have to write:

$$(\mathbf{V} \otimes \mathbf{I}) \mid \varphi\rangle = \left(\mathbf{I} \otimes \mathbf{V}^T\right) \mid \varphi\rangle$$

or

$$\sum_{i=1}^{n} [\mathbf{V} \mid e_i\rangle] \otimes \mid e_i\rangle = \sum_{i=1}^{n} \mid e_i\rangle \otimes \left[\mathbf{V}^T \mid e_i\rangle\right].$$

We have now concluded the survey of the mathematical apparatus required to describe information embodied by a quantum system. As we shall see in Chapter 6 the physical processes for the storage, transformation, and transport of information for photonic, ion-traps, quantum dots, NMR, and other quantum systems are very different and could distract us from the goal of discovering the common properties of quantum information independent of its physical support. To study the properties of quantum information we have to resort to an abstract model which captures the essence of quantum behavior; this model, Quantum Mechanics, describes the properties of physical systems as objects in a finite-dimensional Hilbert space.

## 1.4 Postulates of Quantum Mechanics

A model of a physical system is an abstraction based on *correspondence rules* which relate the entities manipulated by the model to the physical objects or systems in the real world. Once such rules are established, we can operate only with the abstractions according to a set of *transformation rules*. To ensure the usefulness of the model and its ability to describe physical reality we have to *validate* the model and compare its prediction with the physical reality. To ensure expressiveness, the ability of the model to describe the physical system, the correspondence and the transformation rules must be kept as simple as possible, but, at the same time, complete, in other words, capable of capturing the relevant properties of the physical system and of its dynamics, its evolution in time.

*Distinguishability* and system dynamics require the model to abstract the concepts of *observable* and of *state* of the physical object. An observable is a property of the system state that can be revealed as a result of some physical transformation. The state at time $t$ is a synthetic characterization of the object that could be revealed by the measurement of relevant observables at time $t$.

The model must also abstract the concept of *measurement;* it should describe the relation between the state of the object before and after the measurement and specify how to interpret the results of a measurement, how to map the range of possible results to abstractions. In the physical world we often have to deal with a collection of physical objects. If $A, B, C, \ldots$ are the abstractions of the objects $a, b, c, \ldots$, respectively, we need another transformation rule to specify how to construct $\{A, B, C, \ldots\}$ the abstraction corresponding to the collection $\{a, b, c, \ldots\}$. Last, but not least we need transformation rules to describe the system dynamics, the evolution of the system in time.

Quantum Mechanics is a model of the physical world at all scales; it describes more accurately than classical Physics systems at the atomic and sub-atomic scale. This model allows us to abstract our knowledge of a quantum system, to describe the state of single and composite quantum systems, the effect of a measurement on the system's state, and the dynamics of quantum systems. A quantum state summarizes our knowledge about a quantum system at a given moment in time, it allows us to describe what we know, as well as, what we do not know about the system. An impressive number of experiments have produced results consistent with the prediction of Quantum Mechanics and so far there is no experimental evidence to disprove it thus, we shall used this model to study the properties of quantum information.

The correspondence and transformation rules are captured by the *postulates of Quantum Mechanics,* Figure 4. We find it useful to expand the traditional three postulates of Quantum Mechanics, the state postulate, the dynamics postulate, and the measurement postulate, to emphasize some aspects important for quantum information processing:

1. A quantum system $\mathcal{Q}$ is described in an $n$-dimensional Hilbert space $\mathcal{H}_n$, where $n$ is finite. The Hilbert space $\mathcal{H}_n$ is a linear vector space over the field of complex numbers with an inner product. The dimension, $n$, of the Hilbert space is equal to the maximum number of reliably distinguishable states the system $\mathcal{Q}$ can be in.

2. A state $\mid \psi\rangle$ of the quantum system $\mathcal{Q}$ corresponds to a direction (or ray) in $\mathcal{H}_n$. In Section 1.11 we shall see that the most general representation of a quantum state is

1. A PHYSICAL SYSTEM IS REPRESENTED BY A HILBERT SPACE.

2. A STATE OF THE SYSTEM IS A RAY IN THIS SPACE.

3. THE SPONTANEOUS EVOLUTION OF THE SYSTEM IN ISOLATION IS DESCRIBED BY A CERTAIN UNITARY TRANSFORMATION IN THIS HILBERT SPACE.

4. TWO OR MORE SYSTEMS ARE REPRESENTED BY THE TENSOR PRODUCT OF THE HILBERT SPACES REPRESENTING EACH COMPONENT SYSTEM.

5. A MEASUREMENT OF A QUANTUM SYSTEM CORRESPONDS TO A PROJECTION OF ITS STATE INTO ORTHOGONAL SUBSPACES. THE SUM OF THESE PROJECTIONS IS ONE.

Figure 4: The postulates of Quantum Mechanics.

any density operator over an $n$-dimensional Hilbert space with $n$ finite. The density operator is Hermitian, has non-negative eigenvalues, and its trace is equal to unity.

3. When the internal conditions and the environment of a quantum system are completely specified and no measurements are performed on the system, then the system's evolution is described by a unitary transformation in $\mathcal{H}_n$ defined by the Hamiltonian operator. A unitary transformation $\mathbf{U}$ is linear and preserves the inner product. The spontaneous evolution of an unobserved quantum system with the density matrix $\rho$ is

$$\rho \mapsto \mathbf{U}\rho\mathbf{U}^\dagger$$

with $\mathbf{U}^\dagger$ the adjoint of $\mathbf{U}$.

4. Given two *independently* prepared quantum systems, $\mathcal{Q}$ described in $\mathcal{H}_n$ and $\mathcal{S}$ described in $\mathcal{H}_m$, the bi-partite system consisting of both $\mathcal{Q}$ and $\mathcal{S}$ is described in a Hilbert space $\mathcal{H}_n \otimes \mathcal{H}_m$, the tensor product of the two Hilbert spaces.

5. A measurement of the quantum system $\mathcal{Q}$ in state $\mid \psi\rangle$ described in $\mathcal{H}_n$ corresponds to a resolution of $\mathcal{H}_n$ to orthogonal subspaces $\{\mathcal{H}^j\}$ and a projection of the system's state to these subspaces, $\{\mathbf{P}_j\}$, such that the sum of the projections is $\sum \mathbf{P}_j = 1$. The measurement produces the result $j$ with probability

$$\text{Prob}(j) = \mid \mathbf{P}_j \mid \psi\rangle \mid^2 .$$

The state after the measurement is

$$\mid \varphi\rangle = \frac{\mathbf{P}_j \mid \psi\rangle}{\mid \mathbf{P}_j \mid \psi\rangle \mid} = \frac{\mathbf{P}_j \mid \psi\rangle}{\sqrt{\text{Prob}(j)}}.$$

Manipulation of coherent quantum states is at the heart of quantum computing and quantum communication. A quantum computation involves a single entity and consists of unitary transformations of the quantum state. Quantum communication involves multiple entities and involves the transmission of quantum states over noisy communication channels.

## 1.5 Quantum State Postulate

| |---|
| A state is a complete description of a physical system. In Quantum Mechanics, a state $\mid \psi\rangle$ of a system is a vector, in fact, a direction (ray) in the Hilbert space $\mathcal{H}_n$ |

Consider a canonical base $\{\mid 0\rangle, \mid 1\rangle, \ldots, \mid k\rangle, \ldots, \mid (n-1)\rangle\} \in \mathcal{H}_n$; the state $\mid \psi\rangle \in \mathcal{H}_n$ can be written as a linear combination of basis states:

$$\mid \psi\rangle = \sum_{k=0}^{n-1} \alpha_k \mid k\rangle.$$

The coefficients

$$\alpha_k = \langle k \mid \psi\rangle$$

are complex numbers and represent the <u>probability amplitudes</u>; the probability of observing the state $\mid k\rangle$ is $p_k = \mid \alpha_k \mid^2$.

By convention, state vectors are assumed to be normal(ized), i.e., $\langle \psi \mid \psi\rangle = 1$. Therefore:

$$\sum_{k=0}^{n-1} \mid \alpha_k \mid^2 = 1$$

The *length* of a `bra` vector $\langle \psi \mid$ or of the corresponding `ket` vector $\mid \psi\rangle$ is defined as the square root of the positive number $\langle \psi \mid \psi\rangle$. For a given state, the `bra` or `ket` vector representing it is defined only as direction and its length is undetermined up to a factor; the factor is chosen so that *the vector length is usually set equal to unity.* Even then the vector could be undetermined because it can be multiplied by a quantity of modulus 1. Such a quantity, the complex number $e^{i\gamma}$, where $\gamma$ is real, is called *phase factor.*

The inner product of two state vectors, $\mid \psi\rangle$ and $\mid \varphi\rangle$, represents the generalized "angle" between these states and gives an estimate of their *overlap*.

The inner product $\langle\psi \mid \varphi\rangle = 0$ defines orthogonal states; the implication of $\langle\psi \mid \varphi\rangle = 1$ is that $\mid \psi\rangle$ and $\mid \varphi\rangle$ are parallel, in fact, one and the same state.

The inner product of two state vectors is a complex number, but the square of the inner product $\mid \langle\psi \mid \varphi\rangle \mid^2$, a real number, can be interpreted as a quantitative measure of the "relative orthogonality" between these states.

**Superpositions of quantum states.** We assume that the state of a dynamical quantum system at a particular time corresponds to a vector $\mid \psi\rangle$; if this state is the result of a superposition of other states, $\mid \varphi\rangle$ and $\mid \xi\rangle$, it can be represented by the linear expression

$$\mid \psi\rangle = \alpha \mid \varphi\rangle + \beta \mid \xi\rangle.$$

The superposition principle: every vector (*ray*) in the Hilbert space corresponds to a possible state; given two states $\mid \varphi\rangle$ and $\mid \xi\rangle$, we can form another state as a superposition of these two states, $\alpha \mid \varphi\rangle + \beta \mid \xi\rangle$. This is a characteristic of the Hilbert space, which contains all possible superpositions of its vectors, and is suited for the description of *interference effects*.

The state superposition has several properties derived from the properties of linear transformations:

(1) Symmetry - the order of the state vectors in the superposition is not important

$$\mid \psi\rangle = \alpha \mid \varphi\rangle + \beta \mid \xi\rangle = \beta \mid \xi\rangle + \alpha \mid \varphi\rangle$$

(2) Each state in the superposition can be expressed as a superposition of the other states

$$\mid \varphi\rangle = \frac{1}{\alpha}(\mid \psi\rangle - \beta \mid \xi\rangle).$$

(3) The superposition of a state with itself results in the original state:

$$\alpha_1 \mid \varphi\rangle + \alpha_2 \mid \varphi\rangle = (\alpha_1 + \alpha_2) \mid \varphi\rangle.$$

If $\alpha_1 + \alpha_2 = 0$, there is no superposition and the two components cancel each other by an interference effect. If $a_1 + a_2 = a_3$ we assume that the result of the superposition is the original state itself and we can conclude that if the `ket` (`bra`) vector corresponding to a state is multiplied by any non-zero complex number, the resulting `ket` (`bra`) vector will correspond to the same state.

There is a *fundamental difference* between a quantum and a classical superposition. For example, a superposition of a membrane vibration state with itself results in a different state with a different magnitude of the oscillation; the magnitude of such a classical oscillation has no correspondence in any physical characteristic of a quantum state. A classical state with amplitude of oscillation zero everywhere is a membrane at rest. No corresponding state exists for a quantum system since a zero `ket` vector corresponds to no state at all.

Quantum states have several other properties:

- A state is specified by the direction of a `ket` vector; the length of the vector is irrelevant.
- The states of a dynamical system are in one-to-one correspondence with all the possible orientations of a `ket` vector.
- The directions of the `ket` vectors $\mid \psi\rangle$ and $(-\mid \psi\rangle)$ are not distinct.
- When a state $\mid \psi\rangle$ is the result of a superposition of two other states, the ratio of the complex coefficients $\alpha$ and $\beta$ effectively determines the state $\mid \psi\rangle$.

A quantum state is a ray in a Hilbert space; a ray is an *equivalence class* of vectors that differ by multiplication by a nonzero complex scalar. We can choose an element of this class (for any non-vanishing vector) to have unit norm $\langle \psi \mid \psi\rangle = 1$. For such a normalized vector we can say that $\mid \psi\rangle$ or $e^{i\gamma} \mid \psi\rangle$, where $\mid e^{i\gamma} \mid \; = 1$ describe the same physical state. The phase factor $e^{i\gamma}$ becomes physically significant when it appears in a superposition state $(\alpha \mid \varphi\rangle + e^{i\gamma}\beta \mid \xi\rangle)$ as a *relative phase*.

**Transformation of quantum states.** If we apply a transformation to a quantum system in state $\mid \psi\rangle$ e. g., rotate it, or let it evolve for some time $\Delta t$, the system evolves to a different state, $\mid \varphi\rangle$. The physical interaction of a quantum system, e.g., the interaction of an atom with a magnetic field, is represented in our formalism by an operator.

Consider a canonical base $\{\mid 0\rangle, \ldots, \mid j\rangle, \ldots \mid k\rangle \ldots \mid (n-1)\rangle\} \in \mathcal{H}_n$. A Hermitian operator $\mathbf{A} = \mathbf{A}^\dagger$ applied to state $\mid \psi\rangle = \sum_k \alpha_k \mid k\rangle$, with $\alpha_k = \langle k \mid \psi\rangle$, produces a new state $\mid \varphi\rangle$:

$$\mid \varphi\rangle = \mathbf{A} \mid \psi\rangle \quad \text{or} \quad \langle \varphi \mid = \langle \psi \mid \mathbf{A}^\dagger = \langle \psi \mid \mathbf{A}.$$

It follows that

$$\langle j \mid \varphi\rangle = \langle j \mid \mathbf{A} \mid \psi\rangle = \langle j \mid \mathbf{A} \sum_k \alpha_k \mid k\rangle = \sum_k \langle j \mid \mathbf{A} \mid k\rangle \alpha_k = \sum_k \langle j \mid \mathbf{A} \mid k\rangle \langle k \mid \psi\rangle.$$

The expressions $\langle j \mid \varphi\rangle$ give the amount that each basis state $\mid j\rangle$ contributes to the state $\mid \varphi\rangle$; this amount is given in terms of a linear superposition of $\langle k \mid \psi\rangle$, the probability amplitudes in each basis state of the original state, $\mid \psi\rangle$. The numbers $\langle j \mid \mathbf{A} \mid k\rangle$ tell how much of each amplitude $\langle k \mid \psi\rangle$ goes to the sum for each $\mid j\rangle$; these coefficients are the components $A_{jk}$ of the matrix $A = [A_{jk}]$ associated with the linear operator $\mathbf{A}$

$$A_{jk} = \langle j \mid \mathbf{A} \mid k\rangle.$$

Once we have determined the matrix $A$ associated with the operator $\mathbf{A}$ for one basis, we can calculate the corresponding matrix for another basis; the matrix can be transformed from one *representation* to another.

**The qubit.** Consider a quantum system in a two-dimensional Hilbert space, a qubit. The state of a qubit, $\mid \psi\rangle \in \mathcal{H}_2$, can be expressed using the canonical base $\mid 0\rangle$ and $\mid 1\rangle$ as

$$\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle, \quad \alpha_0, \alpha_1 \in \mathbb{C}, \quad \mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 = 1.$$

Recall that using polar coordinates we can express an arbitrary complex number $z = x+iy$ as $z = r\cos\theta + ir\sin\theta$ (where $x = r\cos\theta,\ y = r\sin\theta$ and $i = \sqrt{-1}$). Then using Euler's identity $e^{i\theta} = \cos\theta + i\sin\theta$ we can express $z$ as $z = re^{i\theta}$.

It follows immediately that the state of a qubit can be defined using four real parameters $r_0, \delta_0, r_1, \delta_1$ and substituting $\alpha_0 = r_0 e^{i\delta_0}$ and $\alpha_1 = r_1 e^{i\delta_1}$

$$\mid \psi\rangle = r_0 e^{i\delta_0} \mid 0\rangle + r_1 e^{i\delta_1} \mid 1\rangle.$$

The only measurable quantities for the state $\mid \psi\rangle$ are the probability amplitudes $\mid \alpha_0 \mid^2$ and $\mid \alpha_1 \mid^2$ thus, multiplying the state by an arbitrary global phase $e^{i\gamma}$ has no observable consequences; indeed,

$$\mid e^{i\gamma}\alpha_0 \mid^2 = (e^{i\gamma}\alpha_0)^*(e^{i\gamma}\alpha_0) = (e^{-i\gamma}\alpha_0^*)(e^{i\gamma}\alpha_0) = \alpha_0^*\alpha_0 = \mid \alpha_0 \mid^2 .$$

A similar derivation can be carried out for $\mid \alpha_1 \mid^2$; if we multiply the expression of the state by $e^{-i\delta_0}$ we get a new expression for the state of the qubit

$$\mid \psi'\rangle = r_0 e^{i\delta_0} e^{-i\delta_0} \mid 0\rangle + r_1 e^{i(\delta_1-\delta_0)} \mid 1\rangle = r_0 \mid 0\rangle + r_1 e^{i\varphi} \mid 1\rangle.$$

We notice that now the state depends only on three real parameters, $r_0, r_1$ and $\varphi = \delta_1 - \delta_0$. We express the state $\mid \psi'\rangle$ in Cartesian coordinates as

$$\mid \psi'\rangle = r_0 \mid 0\rangle + (x+iy) \mid 1\rangle$$

and observe that the normalization constraint $\langle \psi' \mid \psi'\rangle = 1$ requires that

$$\mid r_0 \mid^2 + \mid x+iy \mid^2 = 1.$$

This condition can be rewritten as

$$r_0^2 + x^2 + y^2 = 1.$$

This is the equation of unit sphere in a real `3D` space with the Cartesian coordinates $x, y, r_0$. Thus, the state $\mid \psi'\rangle$ is represented by a vector connecting the center to a point on the unit sphere. If we use polar coordinates, Figure 5(a), then the `3D` Cartesian coordinates can be expressed as

$$x = r\sin\theta'\cos\varphi, \quad y = r\sin\theta'\sin\varphi, \quad z = r\cos\theta'.$$

But $r = 1$ thus

$$\mid \psi'\rangle = z \mid 0\rangle + (x+iy) \mid 1\rangle = \cos\theta' \mid 0\rangle + \sin\theta'(\cos\varphi + i\sin\varphi) \mid 1\rangle = \cos\theta' \mid 0\rangle + e^{i\varphi}\sin\theta' \mid 1\rangle.$$

The state requires only two parameters $\theta'$ and $\varphi$. We also notice that when $\theta' = 0$ then $\mid \psi'\rangle = \mid 0\rangle$ and when $\theta' = \pi/2$ then $\mid \psi'\rangle = e^{i\varphi}\sin\theta' \mid 1\rangle$; this suggests that $0 \le \theta' \le \pi/2$ and $0 \le \varphi \le 2\pi$ may generate all the points on the unit sphere.

To show that we can only consider the upper hemisphere of the so called Bloch sphere we consider two opposite points $\mid \psi\rangle$ with polar coordinates $(1, \theta, \varphi)$ and $\mid \psi'\rangle$ with polar coordinates $(1, \theta' = \pi - \theta, \varphi' = \varphi + \pi)$. It is easy to see that the two states differ only by

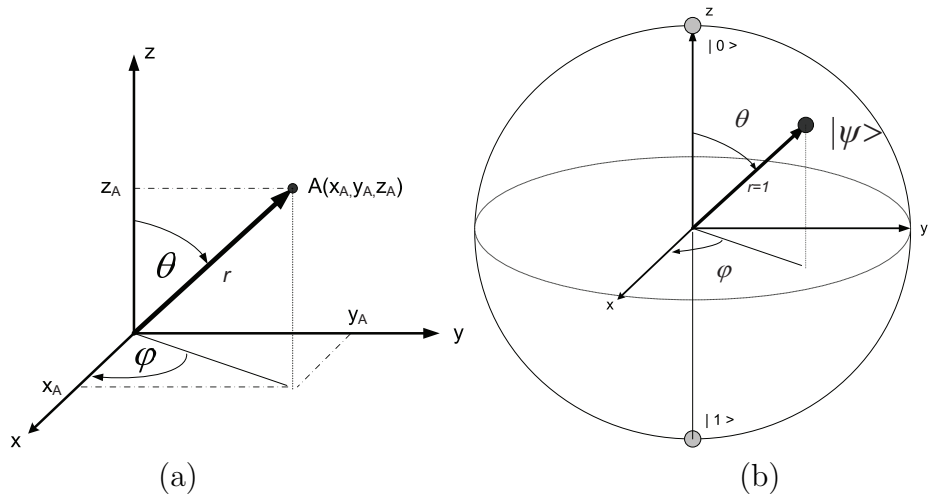



Figure 5: (a) The Cartesian coordinates of the point $A(x_A, y_A, z_A)$ with polar coordinates $(r, \theta, \varphi)$ are: $x_A = r\sin\theta\cos\varphi$, $y_A = r\sin\theta\sin\varphi$, $z_A = r\cos\theta$. (b) The Bloch sphere representation of a qubit in state $\mid \psi\rangle$.

a phase factor of $-1 = e^{-i\pi}$ and we recall that multiplication by $e^{i\gamma}, \forall\gamma$, has no observable consequence:

$$\mid \psi'\rangle = \cos(\pi-\theta) \mid 0\rangle + e^{i(\varphi+\pi)}\sin(\pi-\theta) \mid 1\rangle = -\cos\theta \mid 0\rangle - e^{i\varphi}\sin\theta \mid 1\rangle = - \mid \psi\rangle.$$

Thus, when we use $\theta' = \theta/2$ we can represent the state of a qubit as a point on the Bloch sphere, Figure 5(b), as

$$\mid \psi\rangle = \cos\frac{\theta}{2} \mid 0\rangle + e^{i\varphi}\sin\frac{\theta}{2}, \quad 0 \le \theta \le \pi, \quad 0 \le \varphi \le 2\pi.$$

**Examples of quantum mechanical operators.** We present now a few basic quantum operators:

1. The Pauli spin (rotation) operators for spin-$\frac{1}{2}$ particles: $\sigma_I, \sigma_x, \sigma_y, \sigma_z$ also denoted as $I, X, Y$ and are:

$$\sigma_I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

2. The Hadamard operator $H$ rotates the basis with an angle of 45 degrees

$$H \mid \psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

3. The rotations around the $x, y, z$ axes with an angle $\theta$ are carried out by the following unitary operators:

$$\mathcal{R}_x(\theta) = \cos\frac{\theta}{2}\sigma_I - i\sin\frac{\theta}{2}\sigma_x = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix},$$

$$\mathcal{R}_y(\theta) = \cos\frac{\theta}{2}\sigma_I - i\sin\frac{\theta}{2}\sigma_y = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix},$$

and

$$\mathcal{R}_z(\theta) = \cos\frac{\theta}{2}\sigma_I - i\sin\frac{\theta}{2}\sigma_z = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix}.$$

3. The rotation around an axis connecting the origin with the point $A(x_A, y_A, z_A)$ on the Bloch sphere with an angle $\theta$

$$\mathcal{R}_A(\theta) = \cos\frac{\theta}{2}\sigma_I - i\sin\frac{\theta}{2}(x_A\sigma_x + y_A\sigma_y + z_A\sigma_z)$$

4. The inversion (parity) operator $\mathbf{\Pi}$ creates a new state by reversing the sign of all coordinates.

5. The flip operator in a two-dimensional Hilbert space, $\mathcal{H}_2$, written as a linear combination of orthonormal basis vectors

$$\mathbf{A} = \mid 0\rangle\langle 1 \mid + \mid 1\rangle\langle 0 \mid .$$

When applied to a `ket` state vector, $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$, or the `bra` $\langle\psi \mid$, we obtain, respectively,

$$\mathbf{A} \mid \psi\rangle = (\mid 0\rangle\langle 1 \mid + \mid 1\rangle\langle 0 \mid)\ (\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) = \alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle,$$

$$\langle\psi \mid \mathbf{A} = (\alpha_0^*\langle 0 \mid + \alpha_1^*\langle 1 \mid)\ (\mid 0\rangle\langle 1 \mid + \mid 1\rangle\langle 0 \mid) = \alpha_0^*\langle 1 \mid + \alpha_1^*\langle 0 \mid .$$

6. The $z$-component of the angular momentum , $\mathbf{J}_z$. This operator is defined in terms of the rotation operator for an infinitesimally small angle $\Delta\theta$ around the $z$-axis

$$\mathbf{R}_z(\Delta\theta) = 1 + \frac{i}{\hbar}\mathbf{J}_z\Delta\theta$$

with $\hbar = h/2\pi$, the reduced Planck constant. Among the new states resulting from application of this rotation operator, some are the same as the initial state, except for a phase factor. The phase is proportional to the angle $\Delta\theta$, thus:

$$\mathbf{R}_z(\Delta\theta)\ \mid \psi_0\rangle = \left(1 + \frac{i}{\hbar}\ \mathbf{J}_z\ \Delta\theta\right)\ \mid \psi_0\rangle = e^{im\Delta\theta}\ \mid \psi_0\rangle = (1 + im\Delta\theta)\ \mid \psi_0\rangle$$

When we compare with the definition of $\mathbf{J}_z$ above, we see that

$$\mathbf{J}_z\ \mid \psi_0\rangle = m\hbar\ \mid \psi_0\rangle$$

where $m\hbar$ is the amount of the $z$-component of the angular momentum. This expression can be interpreted in the following way: if we operate with $\mathbf{J}_z$ on a state with a definite angular

momentum about the $z$-axis, we get the same state multiplied by a factor $m\hbar$. The state $\mid \psi_0\rangle$ is an eigenvector of the operator $J_z$ with eigenvalue $m\hbar$.

7. The displacement operator $\mathbf{D}_x(L)$ by distance $L$ along $x$-axis; a small displacement $\Delta x$ along $x$, transforms a state $\mid \psi\rangle$ to $\mid \psi'\rangle$, where

$$\mid \psi'\rangle = \mathbf{D}_x(\Delta x) \ \mid \psi\rangle = \big(1 + \frac{i}{\hbar}\ \mathbf{p}_x\ \Delta x\big) \ \mid \psi\rangle.$$

When $\Delta x \mapsto 0$, the, $\mid \psi'\rangle$ should become the initial state $\mid \psi\rangle$, that is $\mathbf{D}_x(0) = 1$. For infinitesimally small $\Delta x$, the change of $\mathbf{D}_x$ from its value 1 should be proportional to $\Delta x$; the proportionality quantity $\mathbf{p}_x$, is the $x$-component of the momentum operator.

## 1.6 Dynamics Postulate

| |
|---|
| The dynamics of a quantum system is specified by a Hermitian operator $\mathbf{H}$, called the Hamiltonian; the time-evolution of the system is described by the Schrödinger equation: $i\hbar\ \frac{d}{dt} \mid \psi(t)\rangle = \mathbf{H} \mid \psi(t)\rangle$, where $\hbar$ is the reduced Plank's constant. |

More precisely, $\psi(t) = e^{-it/\hbar}\mathbf{H} \mid \psi(0)\rangle$ with $\mid \psi(0)\rangle$ the initial state of the system at $t = 0$. The Schrödinger equation of a finite dimensional quantum system is in fact a *coupled system*[1] *of linear differential equations.*

The Schrödinger equation represents the relation between the states of a quantum system observed at different instants of time. When we make an observation (measurement) on a dynamical quantum system, the state of the quantum system is changed - it is projected with a certain probability onto one of the basis vectors, it "collapses". However, between observations the evolution of the quantum system is expected to be governed by equations of motion and that makes the state at one time determine the state at a later time (causality is assumed to apply in a similar way as in classical Physics). The equations of motion will apply as long as the quantum system is left undisturbed by any observation.

Schrödinger equation, as written above, gives the general law for the variation with time of the vector corresponding to the state at any time. The operator $\mathbf{H}(t)$, the Hamiltonian, is a real, linear operator, characteristic of the dynamical system. The Hamiltonian is a special operator: it describes the complete dynamics of a quantum system under time evolution and it also determines the energy eigenstates (equilibrium states) of the system.

The Hamiltonian matrix $H = [H_{ij}]$ contains all the physics behind the actions which, when applied to the quantum system, cause it to change; thus, the matrix may depend on time. If we know this matrix, we have a complete description of the behavior of the system with respect to time. The Hamiltonian $\mathbf{H}$, as the total energy, takes different forms for different physical systems. For example, the energy of a body of mass $m$ in a *simple harmonic motion* is given by

$$\frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}$$

[1] A set of partial differential equations is a *coupled system* if the same functions $f(x)$ and/or their derivatives appear in several of the equations, so that solving one equation depends on the solutions of another.

where $x, p$, and $\nu$ are, respectively, the position, the momentum, and the oscillation frequency of the body at a given moment in time and $\omega = 2\pi\nu$. A body executes a *simple harmonic motion* when the force that tries to restore the body to its equilibrium (rest) position is proportional to the displacement of the body[2].

To describe the state of a quantum system we need to select a set of basis states and to express through the matrix $H_{ij}$ the physical laws which apply to that particular system. The rule is to find the Hamiltonian corresponding to a particular physical situation, such as the interaction with a magnetic field, with an electric field, and so on. For non-relativistic phenomena and for some other special cases there are very good approximations. For example, the form of the Hamiltonian containing the kinetic energy and the Coulomb interaction between nuclei and electrons in atoms is a very good starting point to describe chemical phenomena.

The *Hamiltonian of a quantum system is a Hermitian operator*

$$\mathbf{H} = \mathbf{H}^\dagger \quad \Longrightarrow \quad H_{ij} = H^*_{ji}.$$

This property is required by the condition that the total probability that an isolated quantum system is in *some* state (any state) does not change in time and by the fact that $H$, the energy observable, must have real eigenvalues. If our system is an isolated particle, as the time goes on we'll find it in the very same state, $\mid \psi\rangle = \sum_i \alpha_i \mid i\rangle$ with $\{\mid i\rangle\}$ an orthonormal basis in $\mathcal{H}_n$; the probability to find it *somewhere* at time $t$ is

$$\sum_i \mid \alpha_i(t) \mid^2 = 1$$

This probability must not vary with time, though the individual probability amplitudes vary with time.

Let us consider some simple quantum systems and describe them using the Hamiltonian:

**A system with one basis state.** A hydrogen atom at rest can be described to a good approximation with only one basis state, which is an energy eigenvector of the hydrogen atom. Since the atom is at rest, we assume that the external conditions do not change in time thus, the Hamiltonian $\mathbf{H}$ is independent of time and

$$i\hbar \, \frac{d\alpha_1}{dt} = H_{11} \, \alpha_1.$$

The system is described with one differential equation for the probability amplitude $\alpha_1$ with $H_{11}$ is constant (in time) and its solution is

$$\alpha_1 = (\text{const}) \, e^{-\frac{i}{\hbar} \, H_{11} \, t}.$$

The solution expresses the time dependence of a state characterized by a definite total energy $E = H_{11}$ of the system.

**A system with two basis states.** Let us assume the basis states are $\mid 0\rangle$ and $\mid 1\rangle$ and the system is in the state $\mid \psi\rangle$. If the probability amplitude of being in state $\mid 0\rangle$ is

[2] A typical example is the motion of an object on a spring when it is subject to an elastic restoring force $F = -kx$ with $k$ the spring constant and $x$ the displacement of the body; the motion is sinusoidal in time.

$\alpha_0 = \langle 0 \mid \psi \rangle$ and the probability amplitude of being in state $\mid 1\rangle$ is $\alpha_1 = \langle 1 \mid \psi \rangle$, then the state vector $\mid \psi\rangle$ can be written as

$$\mid \psi\rangle = \mid 0\rangle\langle 0 \mid \psi\rangle + \mid 1\rangle\langle 1 \mid \psi\rangle$$

or

$$\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle.$$

If we assume that the system changes its state at any given moment in time, the coefficients $\alpha_0$ and $\alpha_1$ will change in time according to the equations

$$i\hbar \frac{d\alpha_0}{dt} = H_{11}\alpha_0 + H_{12}\alpha_1$$

and

$$i\hbar \frac{d\alpha_1}{dt} = H_{21}\alpha_0 + H_{22}\alpha_1$$

*Solution:* We have to make some assumptions about the $H$ matrix in order to solve these equations:

(i) *Stationary states hypothesis:* We assume that once the system is in one of the states, $\mid 0\rangle$ or $\mid 1\rangle$, there is no chance that it could transition to the other state. In this case, the matrix elements expressing transitions from one state to another are $H_{12} = 0$ and $H_{21} = 0$ and the equations become

$$i\hbar \frac{d\alpha_0}{dt} = H_{11}\alpha_0,$$
$$i\hbar \frac{d\alpha_1}{dt} = H_{22}\alpha_1.$$

The solutions of these differential equations for the probability amplitudes are

$$\alpha_0 = (\text{const})\ e^{-\frac{i}{\hbar}\ H_{11}\ t}$$
$$\alpha_1 = (\text{const})\ e^{-\frac{i}{\hbar}\ H_{22}\ t}$$

These are the amplitudes for *stationary* states with energies $E_0 = H_{11}$ and $E_1 = H_{22}$, respectively. These two state are separated by an energy barrier. If the two states $\mid 0\rangle$ and $\mid 1\rangle$ are symmetrical the two energies are equal

$$E_0 = E_1 = E$$

(ii) *Non-stationary states hypothesis:* Let us assume that there is a small probability (amplitude) that the system could transition from one state to the other thus, "tunnel" through the energy barrier separating the two (symmetrical) states. That means

$$H_{12} = H_{21}^* = H_{21} = -\mathcal{E} \quad \text{up to a phase factor.}$$

The initial two equations for the probability amplitudes become

$$i\hbar \frac{d\alpha_0}{dt} = E\alpha_0 - \mathcal{E}\alpha_1$$

$$i\hbar \frac{d\alpha_1}{dt} = E\alpha_1 - \mathcal{E}\alpha_0$$

We solve this system of two differential equations: first we take the sum of the two equations

$$i\hbar \frac{d}{dt}(\alpha_0 + \alpha_1) = (E - \mathcal{E})(\alpha_0 + \alpha_1),$$

then we take the difference of the two equations

$$i\hbar \frac{d}{dt}(\alpha_0 - \alpha_1) = (E + \mathcal{E})(\alpha_0 - \alpha_1)$$

The solutions of these new differential equations are

$$\alpha_0 + \alpha_1 = a\ e^{-\frac{i}{\hbar}\ (E-\mathcal{E})t}$$

and, respectively,

$$\alpha_0 - \alpha_1 = b\ e^{-\frac{i}{\hbar}\ (E+\mathcal{E})t}$$

The integration constants $a$ and $b$ are chosen to give the appropriate initial conditions for a particular system. By adding and subtracting the last two equations, we get the probability amplitudes

$$\alpha_0(t) = \frac{a}{2}\ e^{-\frac{i}{\hbar}\ (E-\mathcal{E})t} + \frac{b}{2}\ e^{-\frac{i}{\hbar}\ (E+\mathcal{E})t}$$

$$\alpha_1(t) = \frac{a}{2}\ e^{-\frac{i}{\hbar}\ (E-\mathcal{E})t} - \frac{b}{2}\ e^{-\frac{i}{\hbar}\ (E+\mathcal{E})t}.$$

These solutions have a physical interpretation:

If $b = 0$:

$$\alpha_0(t) = \alpha_1(t) = \frac{a}{2}\ e^{-\frac{i}{\hbar}\ (E-\mathcal{E})t}$$

the two probability amplitudes are equal and have the same frequency $\nu = (E - \mathcal{E})/\hbar$ in the exponent. We can say that the system is in a state of definite energy $(E-\mathcal{E})$ at this frequency, in a *stationary* state, when the amplitudes $\alpha_0$ and $\alpha_1$ for the system to be in state $\mid 0\rangle$ and, respectively $\mid 1\rangle$ are equal.

If $a = 0$:

$$\alpha_0(t) = \frac{b}{2}\ e^{-\frac{i}{\hbar}\ (E+\mathcal{E})t},$$

$$\alpha_1(t) = -\frac{b}{2}\ e^{-\frac{i}{\hbar}\ (E+\mathcal{E})t},$$

and

$$\alpha_0(t) = -\alpha_1(t).$$

This is another possible stationary state. This time the two amplitudes have the frequency $(E+\mathcal{E})/\hbar$. We say that the system is in a state of definite energy $(E+\mathcal{E})$ if the two amplitudes are equal, but of opposite sign, i.e., $\alpha_0 = -\alpha_1$.

Now, let us assume that at $t = 0$ the system is in state $\mid 0\rangle$. Then:

$$\alpha_0(0) = \frac{a+b}{2} = 1$$

$$\alpha_1(0) = \frac{a-b}{2} = 0$$

with the result

$$a = b = 1$$

The amplitudes become

$$\alpha_0(t) = e^{-\frac{i}{\hbar}\ Et}\left(\frac{e^{\frac{i}{\hbar}\ \mathcal{E}t} + e^{-\frac{i}{\hbar}\ \mathcal{E}t}}{2}\right)$$

$$\alpha_1(t) = e^{-\frac{i}{\hbar}\ Et}\left(\frac{e^{\frac{i}{\hbar}\ \mathcal{E}t} - e^{-\frac{i}{\hbar}\ \mathcal{E}t}}{2}\right)$$

and we can rewrite them as

$$\alpha_0(t) = e^{-\frac{i}{\hbar}\ Et}\cos\frac{\mathcal{E}t}{\hbar}$$

$$\alpha_1(t) = e^{-\frac{i}{\hbar}\ Et}\sin\frac{\mathcal{E}t}{\hbar}.$$

The probability that the system is found in state $\mid 0\rangle$ at time $t$ is

$$\mid \alpha_0(t) \mid^2 = \cos^2\frac{\mathcal{E}t}{\hbar}.$$

The probability that the system is in state $\mid 1\rangle$ at time $t$ is

$$\mid \alpha_1(t) \mid^2 = \sin^2\frac{\mathcal{E}t}{\hbar}.$$

At the initial moment $t = 0$ the probability that the system is in state $\mid 1\rangle$ is zero, increases to one and continues to oscillate between zero and one in time. The probability that the system is in state $\mid 0\rangle$ is one at the initial moment, $t = 0$, decreases to zero and then oscillates between one and zero in time. We say that the magnitude of the two probability amplitudes varies *harmonically* [3] with time. The probability to find the system in one of the two states varies back and forth between the magnitudes of the two individual probabilities.

[3] A harmonic variation is expressed in complex exponential form as $Ae^{i(\omega t+\epsilon)}$, where $\omega = 2\pi\nu$ is the angular velocity.

## 1.7 Measurement Postulate

| |
|---|
| The numerical outcome of a measurement of observable $A$ of a quantum system in state $\mid \varphi\rangle \in \mathcal{H}_n$ is an eigenvalue $\lambda_i$ of the operator $\mathbf{A}$ used to measure observable A; immediately after the measurement, the quantum state of the system is an eigenvector $\mid a_i\rangle$ of $\mathbf{A}$ corresponding to the eigenvalue $\lambda_i$. |

This postulate is sometimes called the *postulate of collapse* because as we shall see later a measurement projects, or collapses, the state of the quantum system being measured.

Another formulation of this postulate is that *mutually exclusive measurement outcomes correspond to orthogonal projection operators (projectors)* $\{\mathbf{P}_0, \mathbf{P}_1, \ldots\}$. From the definition of a complete set of orthogonal projectors it follows that the number of projectors in a complete orthogonal set must be *less* than, or *equal* to the dimension of the Hilbert space. The measurements postulate can be reformulated in terms of completeness of a set of projectors: *a complete set of orthogonal projectors specifies an exhaustive measurement.*

The resulting state after applying the transformation given by $\mathbf{P}_i$ to a quantum system in state $\mid \varphi\rangle$ is

$$\mid \psi\rangle = \mathbf{P}_i \mid \varphi\rangle.$$

The probability of obtaining a measurement outcome $\lambda_i$ is the norm of the resulting state

$$p(\lambda_i) = \mid \mid \psi\rangle \mid^2 = \mid \mathbf{P}_i \mid \varphi\rangle \mid^2 = (\mathbf{P}_i \mid \varphi\rangle)^\dagger \, \mathbf{P}_i \mid \varphi\rangle = \langle \varphi \mid \mathbf{P}_i^\dagger \mathbf{P}_i \mid \varphi\rangle = \langle \varphi \mid (\mathbf{P}_i)^2 \mid \varphi\rangle.$$

But $(\mathbf{P}_i)^2 = \mathbf{P}_i$ thus,

$$p(\lambda_i) = \langle \varphi \mid \mathbf{P}_i \mid \varphi\rangle.$$

The completeness of the set of projectors implies that the total probability for all possible measurement outcomes is $\sum_i p(\lambda_i) = 1$. It follows that

$$\sum_i \langle \varphi \mid \mathbf{P}_i \mid \varphi\rangle = 1.$$

After the measurement, the state $\mid \psi\rangle$ becomes a normalized pure quantum state defined as

$$\mid \psi'\rangle = \frac{\mid \psi\rangle}{\sqrt{\langle \varphi \mid \mathbf{P}_i \mid \varphi\rangle}} = \frac{\mathbf{P}_i \mid \varphi\rangle}{\sqrt{\langle \varphi \mid \mathbf{P}_i \mid \varphi\rangle}}.$$

**Example.** *Consider a two-dimensional Hilbert space, $\mathcal{H}_2$ where we have chosen the orthogonal basis vectors $\mid x\rangle$ and $\mid y\rangle$.* We define two possible state vectors of a system in this space

$$\mid \varphi_A\rangle = \alpha_x \mid x\rangle + \alpha_y \mid y\rangle \quad \text{and} \quad \mid \varphi_B\rangle = \beta_x \mid x\rangle + \beta_y \mid y\rangle$$

The initial state of the system could be $\mid \psi_A\rangle$ with probability $p$ or $\mid \psi_B\rangle$ with probability $1-p$; we say that the system is in a *mixed* ensemble of quantum states.

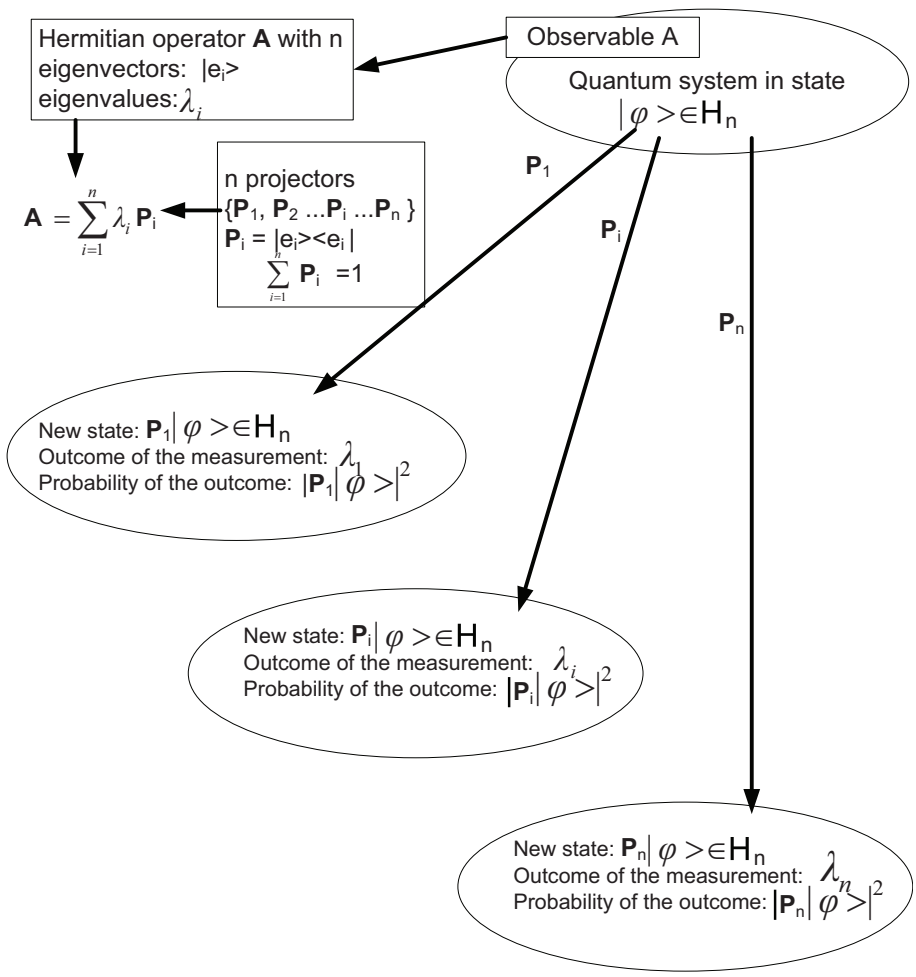



Figure 6: The measurement of observable $A$ of a quantum system in state $\mid \varphi\rangle \in \mathcal{H}_n$. The operator $\mathbf{A}$ corresponding to observable $A$ is Hermitian and has $n$ eigenvalues, $\lambda_i \in \mathbb{R}$, $1 \leq i \leq n$. The corresponding eigenvectors $\mid e_i\rangle$, $1 \leq i \leq n$ form an eigenbasis. The set of $n$ projectors $\mathbf{P}_i$ is complete if $\sum_{i=1}^{n} \mathbf{P}_i = \mathbf{I}$; the state of the system after the transformation due to $\mathbf{P}_i$ is $\mid \psi_i\rangle = \mathbf{P}_i \mid \varphi\rangle$, the outcome of the measurement is $\lambda_i$ and the $\text{Prob}(\lambda_i) =\mid \mathbf{P}_i \mid \varphi\rangle \mid^2$.

We perform a large number, $N$, of measurements corresponding to the projectors

$$\mathbf{P}_x =\mid x\rangle\langle x \mid \quad \text{and} \quad \mathbf{P}_y =\mid y\rangle\langle y \mid .$$

We wish to predict $n_x$, the number of times out of the total number $N$ of measurements, when we expect to obtain the measurement outcome corresponding to basis vector $\mid x\rangle$. We use the conditional probabilities $\text{Prob}(x\mid \mid \varphi_A\rangle)$ and $\text{Prob}(x\mid \mid \varphi_B\rangle)$ to express $n_x$ as

$$\begin{aligned} n_x &= N\left[\text{Prob}(\mid \varphi_A\rangle) \cdot \text{Prob}(x\mid \mid \varphi_A\rangle) + \text{Prob}(\mid \varphi_B\rangle) \cdot \text{Prob}(x\mid \mid \varphi_B\rangle)\right] \\ &= N[p\ \langle\varphi_A \mid \mathbf{P}_x \mid \varphi_A\rangle + (1-p)\langle\varphi_B \mid \mathbf{P}_x \mid \varphi_B\rangle] \\ &= N[p \mid \alpha_x \mid^2 +(1-p) \mid \beta_x \mid^2] \end{aligned}$$

The ratio $n_x/N$ is bounded from below by the smaller of $\mid \alpha_x \mid^2$ and $\mid \beta_x \mid^2$ as $0 \leq p \leq 1$.

Let us assume that the initial state of the system is a *coherent superposition* of the states $\mid \varphi_A\rangle$ and $\mid \varphi_B\rangle$ corresponding to the *pure* state $\mid \varphi(\gamma_A, \gamma_B)\rangle$, with

$$\begin{aligned} \mid \varphi(\gamma_A, \gamma_B)\rangle &= \gamma_A \mid \varphi_A\rangle + \gamma_B \mid \varphi_B\rangle \\ &= (\gamma_A\alpha_x + \gamma_B\beta_x) \mid x\rangle + (\gamma_A\alpha_y + \gamma_B\beta_y) \mid y\rangle \end{aligned},$$

were $\gamma_A$ and $\gamma_B$ are chosen such that the state $\mid \varphi(\gamma_A, \gamma_B)\rangle$ is normalized, $\mid \gamma_A \mid^2 + \mid \gamma_B \mid^2 = 1$.

In this case, the probability of a measurement outcome corresponding to basis vector $\mid x\rangle$ is $n_x/N$:

$$\begin{aligned} \tfrac{n_x}{N} &= \langle \varphi(\gamma_A, \gamma_B) \mid \mathbf{P}_x \mid \varphi(\gamma_A, \gamma_B)\rangle \\ &= \mid \gamma_A\alpha_x + \gamma_B\beta_x \mid^2 \end{aligned}$$

We notice that in certain cases it is possible to choose $\gamma_A$ and $\gamma_B$ such that $\gamma_A\alpha_x = -\gamma_B\beta_x \neq 0$ and then, $n_x = 0$ though $\mid \alpha_x \mid^2, \ \mid \beta_x \mid^2 \ > \ 0$. This phenomenon is known as *destructive interference*.

There is an important distinction between *coherent superpositions* (of the type that produce a single pure state) and the *incoherent admixtures* (of the type that produce a mixed ensemble of quantum states).

Now we summarize several important properties of the quantum observables and quantum operators related to the concepts discussed in this section:

1. An *observable* in Quantum Mechanics is a Hermitian operator with a complete set of eigenvectors. This set of eigenvectors can be chosen to be mutually orthogonal; they form a basis.

2. The eigenvalues of a Hermitian operator are real numbers.

3. Eigenvectors of a Hermitian operator corresponding to different eigenvalues are mutually orthogonal.

4. If two Hermitian operators commute, they have a common basis of orthonormal eigenvectors (an eigenbasis). If they do not commute, then no common eigenbasis exists.

5. A complete set of commuting observables is the minimal set of Hermitian operators with a unique common eigenbasis.

6. The eigenvalues of a unitary operator are complex numbers of modulus 1.

7. Eigenvectors of a unitary operator corresponding to different eigenvalues are mutually orthogonal.

8. In general, an operator must be normal for the property of orthogonality of eigenvectors to hold.

9. If $\mathbf{A}$ is a Hermitian operator, and $A$ is an observable of a system, then the measurement of observable $A$ of the system in state $\mid \varphi\rangle$ leaves the system in a state which is an eigenvector $\mid a\rangle$ of $\mathbf{A}$ and the probability of this outcome is $\mid \langle a \mid \varphi\rangle \mid^2$.

## 1.8 Linear Algebra and Systems Dynamics

Linear algebra allows us to describe the evolution in time of classical, as well as quantum systems. Consider first a discrete-time, finite-dimensional, non-deterministic classical system; the state transitions occur at distinct time instances $t_1, t_2, \ldots, t_k, \ldots$; the state space is of dimension $n$. The state of the system at time $t_k$ is a stochastic vector, $\sigma = (p_1^k, p_2^k, \ldots, p_n^k)$ with $\sum_i p_i^k = 1$ and $p_i^k$ the probability of the system being in state $i, 1 \leq i \leq n$, at time $t_k$. The dynamics of the system is captured by a directed graph $G$ with vertices corresponding to the states; the arcs correspond to transitions and are labelled with the transition probabilities. This graph is characterized by the adjacency matrix $A = [a_{ij}]$ with $a_{ij}$ the probability of the transition from state $i$ to state $j$. The adjacency matrix $A$ is a *row stochastic matrix*, the sum of all elements in a row is equal to 1.

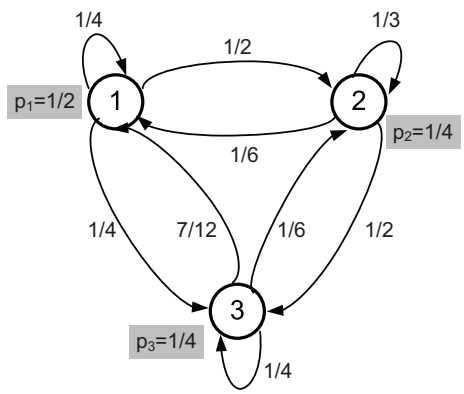



Figure 7: Systems dynamics. The directed graph $G$ with adjacency matrix $A$.

For example, in the graph in Figure 7 the state of the system is given by the vector $\sigma^T(t_k) = (1/2\ 1/4\ 1/4)$ and the adjacency matrix is

$$A = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 7/12 & 1/6 & 1/4 \end{pmatrix}$$

If the system is in state $\sigma_{t_k}$ at time $t_k$, then the state at time $t_{k+1}$ will be $\sigma_{t_{k+1}} = A\sigma_{t_k}$:

$$\sigma_{t_{k+1}} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/6 & 1/3 & 1/2 \\ 7/12 & 1/6 & 1/4 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} 5/16 \\ 7/24 \\ 19/48 \end{pmatrix}.$$

The next state of a classical Markovian discrete-time stochastic system depends only on the current state. The system could reach a state $\sigma_i$ from two distinct initial states, $\sigma_a$ and $\sigma_b$; the memoryless property does not allow to distinguish which path was taken to reach $\sigma_i$ and makes it impossible to return to the initial state, the system is *non-reversible*. Only if $A$ is a permutation matrix, a matrix with one non-zero element in each row and each column, the discrete-time system is reversible. The description of a classical stochastic system with a continuous-time and a possibly infinite state space is considerably more complex and will not be discussed.

We have been concerned with the dynamics of classical systems where probabilities are real numbers $0 \leq p_i, p_j \leq 1$ and $\sum_i p_i = 1$; then probabilities can only increase when added, $p_i + p_j \geq p_i$ and $p_i + p_j \geq p_j$. What if the probabilities are functions of complex numbers, e.g., they are the square of the modulus of complex numbers, $\mid q_i \mid^2 = \mid \alpha_i + i\beta_i \mid^2$, and we replace the condition $\sum_i p_i = 1$ with $\sum_i \mid q_i \mid^2 = \sum_i (\alpha_i^2 + \beta_i^2) = 1$.

With these rules in place we notice that $\mid q_1 + q_2 \mid^2$ can be smaller than $\mid q_1 \mid^2$ or $\mid q_2 \mid^2$. For example, if $q_1 = 3 + 5i$ and $q_2 = -4 - 4i$, then $\mid q_1 \mid^2 = 9 + 25 = 34$, $\mid q_2 \mid^2 = 16 + 16 = 32$ and $\mid q_1 + q_2 \mid^2 = \mid (3-4) + i(5-4) \mid^2 = 1 + 1 = 2$.

The arcs in the directed graphs are now labelled with complex numbers and we now require that the adjacency matrix be unitary, rather than row stochastic. If $q_{ij}$ is the label of the arc from state $i$ to state $j$ then the adjacency matrix is

$$Q = Q[q_{ij}] \quad \text{and} \quad QQ^\dagger = I$$

Then:

$$\mid \sigma_{t+1} \rangle = Q \mid \sigma_t \rangle \quad \Longrightarrow \quad \langle \sigma_t \mid = \langle \sigma_{t+1} \mid Q.$$

Indeed, $\mid \sigma_{t+1} \rangle = Q \mid \sigma_t \rangle$ implies that $\langle \sigma_{t+1} \mid = \langle \sigma_t \mid Q^\dagger$ and $\langle \sigma_{t+1} \mid Q = \langle \sigma_t \mid Q^\dagger Q = \langle \sigma_t \mid$.

A logical question we address next is why we choose the transition probabilities for a discrete-time quantum system to be functions of complex numbers. The answer is that the new probability rules allow us to capture the wave-like behavior of atomic and subatomic particles while the real-valued probabilities describe classical, particle-like behavior.

**The wave-particle duality and the new probability rule.** One of the greatest discoveries leading to the formulation of wave equation of Schrödinger in 1926 was the wave-particle duality principle formulated by Louis de Broglie in 1923; this principle states that atomic and subatomic particles (matter) and photons (energy) exhibit both wave-like and particle-like properties. The wavelength $\lambda$ and the momentum $p$ of any form of matter are related: $\lambda = h/p$ with $h$ Planck's constant.

To grasp the physical significance of the probability rules we consider the double-slit experiments discussed [149]. When the experiment is performed with bullets and when both slits are open the number of bullets $n(x)$ collected at a spot $x$ between the two slits is the sum $n(x) = n_{up}(x) + n_{low}(x)$, with $n_{up}(x)$ the number of bullets when only the upper slit is open and $n_{low}(x)$ is the number of bullets when only the lower slit is open. If $N$ bullets are shot and $p(x) = n(x)/N$ is the probability of a bullet hitting spot $x$ with both slits open, $p_{up}(x) = n_{up}(x)/N$ and $p_{low}(x) = n_{low}(x)/N$ correspond to the two probabilities with only one slit open, then the usual probability rule applies: $p(x) = p_{up}(x) + p_{low}(x)$.

When the double slit experiment is performed with waves, then in some spots the amplitude $a(x)$ of the waves is reinforced $a(x) > \max(a_{up}, a_{low})$, while in other spots it is diminished, $a(y) < \min(a_{up}(y), a_{low}(y))$. This is due to the phenomenon of *interference* characteristic to wave-like behavior.

In conclusion, *state transformation in a Hilbert space corresponds to application of a unitary operator to the current state and the probability of a state is the modulus of a complex number.* For example in $\mathcal{H}_2$ a system in state:

$$\mid \varphi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$$

has the probability $p_0 =\mid \alpha_0 \mid^2$ to be in state $\mid 0\rangle$ and the probability $p_1 =\mid \alpha_1 \mid^2$ to be in state $\mid 1\rangle$ with $\mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2= 1$. This system will be the embodiment of a unit of quantum information called a qubit.

## 1.9 Symmetry and Dynamic Evolution

Symmetry plays an important role in classical as well as quantum physics. The mathematical concept which describes the symmetry of a physical system or of a physical object is the concept of a group. An abstract group is a set $G$ equipped with a map $g : G \times G \mapsto G$ and $(g)^{-1} : G \mapsto G$ and an element $e \in G$ so that the familiar axioms discussed in Section 4.4 hold. A subgroup $G'$ of the group $G$ is a subset such that $g_1, g_2 \in G'$ implies $g_1 \cdot g_2^{-1} \in G'$.

The set of symmetries of a physical system form a group; one can compose two symmetries, consider the inverse of each symmetry, and there always exists the "identity" as an obvious symmetry. Most of these groups belong to a remarkable class of groups, the Lie groups. An abstract Lie group satisfies the axioms of a group and, in addition, is a "smooth manifold"[4] such that both the composition and the inverse are given by differentiable functions. The most familiar group is $GL(n, F)$, the group of symmetries of an $n$-dimensional vector space over the field $F$ of real or complex numbers.

Almost all Lie groups can be described as subgroups of $GL(n, \mathbb{R})$ or $GL(n, \mathbb{C})$. For example, $SL(n, \mathbb{R})$, the group of symmetries of $\mathbb{R}^n$ which preserves the $n$-dimensional volume, is the subgroup of $GL(n, \mathbb{R})$ consisting of matrices with a determinant equal to 1. As another example, $O(n, \mathbb{R})$, the group of symmetries of the vector space $\mathbb{R}^n$ which preserves the angles and the distances, is a subgroup of matrices $\mathbf{A}$ with real elements and $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. Also, $U(n)$, the group of symmetries of the Hilbert space $\mathbb{C}^n$ is the subgroup of matrices $\mathbf{A}$ in $GL(n, \mathbb{C})$ with $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}$.

Lie algebra is a mathematical object which is mathematically, rather than conceptually, simpler than a Lie group. To each Lie group one can associate a Lie algebra; if the Lie group is a group of symmetries, then the Lie algebra is usually called infinitesimal symmetries. Each finite-dimensional Lie algebra determines an unique (connected and simply connected) Lie group.

A Lie algebra is a vector space $V$ over a field $F$ together with a binary operation $[,]$ called a Lie bracket with the following properties:

1. Bilinearity:

$$[\alpha u + \beta v, w] = \alpha[u, w] + \beta[v, w], \quad [w, \alpha u + \beta v] = \alpha[w, u] + \beta[w, v], \quad \forall \alpha, \beta \in F, \quad \forall u, v, w \in V.$$

2. The Lie bracket $[u, u]$:

$$[u, u] = 0, \quad \forall u \in V.$$

3. Jacobi identity:

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0, \quad \forall u, v, w \in V.$$

[4] A smooth manifold is a geometrical object which is locally the same as the Euclidian space $\mathbb{R}^n$ [269]

The operation defined by the Lie bracket is not in general associative $[[x, y], z]$ need not equal $[x, [y, z]]$.

The Lie algebra of $GL(n, F)$ is denoted by $gl(n, F)$ and is the vector space of all $n \times n$ matrices with elements from the field $F$ and bracket $[A, B] = A \cdot B - B \cdot A$. The Lie algebra of $SL(n, \mathbb{R}$ and $O(n, \mathbb{R})$, are subspaces of $n \times n$ matrices $A$ with real elements and $\mathrm{tr}(A) = 0$, respectively $A + A^T = 0$ (skew symmetric matrices). The Lie algebra of $U(n)$, denoted by $u(n)$ consists of $n \times n$ matrices with complex elements which satisfy $A + A^\dagger = 0$, equivalently with matrices $A$ such that $(iA)$ is a Hermitian matrix.

An element in the Lie algebra defines a smooth one-parameter family of elements in the Lie group parameterized by $t \in \mathbb{R}$ which for $t = 0$ is exactly $e$. Any smooth one-parameter family of elements in the group which, for $t = 0$ is $e$, an element in the Lie algebra of the group. If an $n \times n$-matrix $A$ is an element of $gl(n, F)$ then the one-parameter family is

$$e^{tA} = I + \frac{tA}{1!} + \frac{(tA)^2}{2!} + \dots.$$

This series is convergent for any $t$ to an invertible matrix. If $\mathbf{A}(t)$ is an one-parameter family of elements in $gl(n, F)$ the

$$\frac{d\mathbf{A}(t)}{dt} \mid_{t=0}$$

is a matrix, not necessary invertible, hence an element in $gl(n, F)$.

The non-commutativity of the Lie group is reflected in the bracket operation; more precisely if

$$A = \frac{d\mathbf{A}(t)}{dt} \mid_{t=0} \quad \text{and} \quad B = \frac{d\mathbf{B}(t)}{dt} \mid_{t=0}$$

then

$$\lim_{t \to 0} \frac{\mathbf{A}(t)\mathbf{B}(t)\mathbf{A}(t)^{-1}\mathbf{B}(t)^{-1} - \mathbf{I}}{t^2} = [A, B].$$

In particular

$$[A, B] = \lim_{t \to 0} \frac{e^{tA} \cdot e^{tB} \cdot e^{-tA} \cdot e^{-tB} - I}{t^2}$$

The equality can be verified by applying l'Hospital rule from calculus.

Usually the elements of a Lie algebra of a Lie group of symmetries, i.e. infinitesimal symmetries, define conservation laws (first observed by Nöther [14]).

If an observable $\mathbf{A}$ (a self-adjoint operator) is invariant by the dynamics of a quantum system defined by the Hamiltonian $\mathbf{H}$ (i.e., all eigenstates and eigenvalues are preserved by the dynamics) which means $\mathbf{A}e^{-\frac{i}{\hbar}t\mathbf{H}} = e^{-\frac{i}{\hbar}t\mathbf{H}}\mathbf{A}$ (Section 1.6); one obtains (by derivation at $t = 0$) $[\mathbf{H}, \mathbf{A}] = 0$. This holds even in the more general case when the dynamics is defined by the time-dependent Hamiltonian $\mathbf{H}(t)$; precisely, $[\mathbf{H}(t), \mathbf{A}] = 0$.

## 1.10 Uncertainty Principle; Minimum Uncertainty States

In classical Physics the nondeterminism is due to *uncontrolled causes that are recognized to exist and that, if better known, would make the predictions better.* The quantum state postulate reveals the critical role of the non-determinism in quantum theory; the uncertainty principle shows that it has broader implications on our ability to observe the properties of a quantum system and makes us wonder if its nature is fundamentally different from the role it plays in classical Physics.

In quantum physics nondeterminism means that a *precise knowledge at the quantum level is impossible.* The uncertainty principle introduced by Werner Heisenberg in 1927, is at the heart of the special nature of the nondeterminism of Quantum Mechanics. Uncertainty is an intrinsic property of quantum systems. The accuracy of measured values of physical properties, such as position and momentum along the axis used to measure the position, is limited; the precise knowledge of both the position and the momentum is forbidden in a quantum system. This limitation cannot be avoided, as shown by many experiments performed over the years. Einstein, who doubted that "God is playing dice," questioned the truth of such an indeterminacy.

The uncertainty principle: consider canonically conjugated observables, $x$ the position of a particle and $p_x$ the momentum of the same particle at position $x$; $x$ and $p_x$ are three-dimensional vectors and we consider their projections along the same direction $d$. Then $\Delta x$, the uncertainty in determining the projection of the position, and $\Delta p_x$, the uncertainty in determining the projection of the momentum at position $x$ along the same direction, are constrained by the inequality:

$$\Delta x \; \Delta p_x \geq \frac{\hbar}{2}$$

where $\hbar = h/2\pi$ is the reduced Planck's constant.

Let us assume that we are interested in two observables $A$ and $B$ of a quantum particle, associated with the Hermitian operators $\mathbf{A}$ and $\mathbf{B}$, respectively. We prepare two disjoint sets $\mathcal{S}_1$ and $\mathcal{S}_2$ each set consisting of a large number of quantum systems in the identical state $\mid \varphi\rangle$. For the systems in $\mathcal{S}_1$ we measure first the observable $A$ and then the observable $B$ on all particles. We obtain the "same" value for the observable $A$, while we notice a large standard deviation of the observable $B$. For the systems in $\mathcal{S}_2$ we measure first the observable $B$ and then the observable $A$. Now we obtain the same value for the observable $B$, while we notice a large standard deviation of the observable $A$.

Call $\Delta\mathbf{A}$ and $\Delta\mathbf{B}$ the standard deviation of the measurements of observables $A$ and $B$, respectively. The uncertainty principle can be expressed as

$$\Delta\mathbf{A} \; \Delta\mathbf{B} \; \geq \; \frac{1}{2} \mid \langle\varphi \mid [\mathbf{A}, \mathbf{B}] \mid \varphi\rangle \mid .$$

with $[\mathbf{A}, \mathbf{B}]$ the commutator of the two operators, $\mathbf{A}$ and $\mathbf{B}$. If $A$ and $B$ are non-commutative quantum observables then $[\mathbf{A}, \mathbf{B}] > 0$; it follows that $\Delta\mathbf{A}\Delta\mathbf{B} \; > \; 0$ thus, there is a minimum level of uncertainty that cannot possibly be removed.

On the other hand, *when two operators commute, then they can be diagonalized simultaneously and we can measure the eigenvalues of one of them without disturbing the eigenvectors*

*of the other.* This property is important for a class of quantum error correcting codes, the stabilizer codes, discussed in Section 5.12.

The quantum mechanical uncertainty relation $\Delta x\ \Delta p_x \geq \frac{\hbar}{2}$ has a classical counterpart based on the wave phenomena; an acoustic signal with intensity $s(t)$ cannot have precise timing and precise pitch, the two must satisfy the inequality:

$$\Delta t\ \Delta\omega \geq \frac{1}{2}$$

where:

$$(\Delta t)^2 \equiv \int t^2 s(t)dt - \left(\int ts(t)dt\right)^2 \quad \text{and} \quad (\Delta\omega)^2 \equiv \int \omega^2 S(w)d\omega - \left(\int \omega S(\omega)d\omega\right)^2.$$

In this expression $S(\omega) = \mathcal{F}(s(t))$ is the Fourier transform of the function $s(t)$ and $\omega = (2\pi)/f$, with $f = 1/T$ the frequency and $T$ the period of the signal.

Peres observes [325] that in the case of acoustic signals we can have *approximate* values for the time and the frequency; musical scores indicate both time and frequency and allow musicians to produce sounds that capture the information generated by the composer. It seems quite reasonable to consider also approximate values for the observables of a quantum system. For example, we can consider a Gaussian distribution of the position and wavelength, $\lambda = h/p$, of a wave packet with mean: $\bar{x}$ and variance $\sigma$, and, respectively, $\bar{p}$. Then the wave function:

$$\phi(x) = (\pi\sigma^2)^{-1/4} \exp\left[\frac{x-\bar{x}}{2\sigma^2}\right] + \frac{1}{\hbar} i\bar{p}x$$

is a minimum uncertainty wave packet with:

$$\Delta x = \frac{1}{\sqrt{2}}\sigma, \qquad \Delta p = \frac{\hbar}{\sqrt{2}}\sigma \quad \text{and} \quad \Delta x\ \Delta p = \frac{\hbar}{2}\sigma^2.$$

Minimum uncertainty wave functions can be extended to non-commuting operators other than position and momentum [325]. Coherent superposition states, introduced in Section 1.7 and discussed in more depth in Section 3.16, are minimum uncertainty states.

## 1.11 Pure and Mixed Quantum States

So far, our discussion was focused on a class of states of a quantum system, described by Dirac's `ket` and `bra` vectors, or, by a *wave function* in a finite-dimensional Hilbert space. These states are called pure quantum states; a pure state can be expressed as a superposition of the basis vectors of an orthonormal basis $\mathcal{B} = \{\mid b_1\rangle, \mid b_2\rangle, \ldots, \mid b_n\rangle\} \in \mathcal{H}_n$ as

$$\mid \psi\rangle = \sum_{i=1}^{n} \alpha_i \mid b_i\rangle \qquad \text{with} \qquad \sum_{i=1}^{n} \mid \alpha_i \mid^2 = 1.$$

A pure state that is a linear combination of two or more component states is sometimes called a coherent superposition. According to the postulates of Quantum Mechanics, discussed in

Section 1.4, the evolution of a closed quantum system can be completely described as a unitary transformation of pure states in a Hilbert space.

For a qubit in a coherent superposition state there is always a basis in which any measurement of the qubit will produces the same result. For example, the state

$$\mid \psi\rangle = \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$$

represents a qubit which has a 50% probability to produce either a "0" or a "1" as a result of a measurement. If we rotate the basis by 45° by applying a Hadamard transformation, $H$, the state of the qubit becomes

$$H \mid \psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and all measurements of the resulting state $\mid 0\rangle$, produce the same result, "0."

We introduce a new characterization of the state of a quantum system by means of the density operator. We can associate every vector $\mid \psi\rangle \in \mathcal{H}_n$ with a matrix $A \in \mathbb{C}^{n\times n}$; indeed, $\mathbb{C}^{n\times n}$, the set of $n \times n$ matrices with complex elements, is an inner-product vector space thus, there is an isomorphism from a Hilbert space to $\mathbb{C}^{n\times n}$. The <u>density matrix of a pure state</u> is

$$\mid \psi\rangle = \sum_{i=1}^{n} \alpha_i \mid b_i\rangle \qquad \text{with} \qquad \sum_{i=1}^{n} \mid \alpha_i \mid^2 = 1.$$

The density matrix of a pure state $\mid \psi\rangle$ is given by the outer product

$$\rho = \mid \psi\rangle\langle\psi \mid = \begin{pmatrix} \alpha_1\alpha_1^* & \alpha_1\alpha_2^* & \dots & \alpha_1\alpha_n^* \\ \alpha_2\alpha_1^* & \alpha_2\alpha_2^* & \dots & \alpha_2\alpha_n^* \\ \vdots & \vdots & & \vdots \\ \alpha_n\alpha_1^* & \alpha_n\alpha_2^* & \dots & \alpha_n\alpha_n^* \end{pmatrix}.$$

The density matrix of a pure state is a Hermitian operator

$$\rho^\dagger = (\mid \psi\rangle\langle\psi \mid)^\dagger = \mid \psi\rangle\langle\psi \mid = \rho.$$

The trace of the density matrix of a pure state is equal to one

$$\mathrm{tr}(\rho) = \sum_{i=1}^{n} \alpha_i\alpha_i^* = \sum_{i=q}^{n} \mid \alpha_i \mid^2 = 1.$$

A <u>mixed state</u> is a statistical ensemble $\{(\mid \psi_1\rangle, p_1), (\mid \psi_2\rangle, p_2), \dots, (\mid \psi_n\rangle, p_n), \}$ with $p_i$ a probability in the classical sense, $\sum_{i=1}^{n} p_i = 1$, rather than a probability amplitude. The <u>density matrix of a mixed state</u> $\sum_{i=1}^{n} p_i \mid \psi_i\rangle \in \mathcal{H}_n$ is defined as

$$\rho = \sum_{i=1}^{n} p_i \mid \psi_i\rangle\langle\psi_i \mid$$

where more than one factor $p_i$ is greater than zero. The density matrix of a mixed state is a Hermitian operator and its trace is one. Different mixtures of pure states could have the same density matrix; for example, the mix of the pair of states

$$| +\rangle = \frac{1}{\sqrt{2}}(| 0\rangle + | 1\rangle) \quad \text{and} \quad | -\rangle = \frac{1}{\sqrt{2}}(| 0\rangle - | 1\rangle)$$

with probability $p = 1/2$ and the mix of pair of states $| 0\rangle$ and $| 1\rangle$ with probability $p = 1/2$ have the same density matrix. See also Section 2.3 where we show that the density matrix allows us to distinguish pure states from mixed (impure) states: tr $(\rho^2) = 1$ for a pure state and tr $(\rho^2) < 1$ for a mixed state. Pure states are represented as points on the surface of the *Bloch sphere* while mixed states are points inside the *Bloch ball*. An incoherent mixture remains a mixture whatever basis we choose to describe it.

The density matrix plays an important role in quantum information theory, it provides an answer to the question "how much information can we acquire about a quantum state?" *Pure states* are characterized by maximal knowledge or minimal ignorance; in principle there is nothing more to be learned about a quantum system in a pure state [44]. Whenever we can only attribute probabilities to possible states, or when we are allowed to observe only a subsystem of a composed system, we cannot acquire maximum information about the entire quantum system and we say that the system is in a *mixed state.* The density operator $\rho$ allows us to distinguish pure from mixed quantum states; $\rho$ is a Hermitian operator and $\text{tr}(\rho) = 1$.

## 1.12 Entanglement; Bell States

We discussed a number of intriguing properties of quantum states and we add to this list the entanglement, a phenomena without a classical counterpart. Quantum mechanical systems have a unique property: a bipartite system, a system composed of two subsystems, can be in a pure state for which it is not possible to assign a definite state to each of its component subsystems. This strong correlation of quantum states is called entanglement; entangled states are pure states of bipartite quantum systems.

Erwin Schrödinger discovered the phenomenon of entanglement[5] and in 1935 he made a crucial observation: "total knowledge of a composite system does not necessarily include maximal knowledge of all its parts, not even when these are fully separated from each other and at the moment are not influencing each other at all."

Charles Bennett and Peter Shor comment on the effect of entanglement for quantum information processing [44]: "classical information can be copied freely, but can only be transmitted forward in time to a receiver in the sender's forward light cone. Entanglement, by contrast cannot be copied, but can connect any two points in space-time. Conventional data-processing operations destroy entanglement, but quantum operations can create it, preserve it and use it for various purposes, notably speeding up certain computations and assisting in the transmission of classical data or intact quantum states (teleportation) from a sender to a receiver."

[5]Entanglement is the English translation of the German noun Verschränkung, the name used by Schrödinger to describe this phenomena.

According to the postulates of Quantum Mechanics the joint state of two or more quantum systems is a vector in a Hilbert defined as the tensor product of the Hilbert spaces $\mathcal{H}_{n_1}, \mathcal{H}_{n_2} \ldots \mathcal{H}_{n_k}$ used to represent the individual states of the component systems

$$\mathcal{H}_{n_1 \times n_2 \cdots \times n_k} = \mathcal{H}_{n_1} \otimes \mathcal{H}_{n_2} \ldots \otimes \mathcal{H}_{n_k}.$$

For example, the state of a quantum system consisting of two qubits is a vector in $\mathcal{H}_{2^2} = \mathcal{H}_2 \otimes \mathcal{H}_2$ with the orthonormal basis $\{\mid 00\rangle, \mid 01\rangle, \mid 10\rangle, \mid 11\rangle\}$ expressed as

$$\mid \psi\rangle = \alpha_{00} \mid 00\rangle + \alpha_{01} \mid 01\rangle + \alpha_{10} \mid 10\rangle + \alpha_{11} \mid 11\rangle$$

with $\mid \alpha_{00} \mid^2 + \mid \alpha_{01} \mid^2 + \mid \alpha_{10} \mid^2 + \mid \alpha_{11} \mid^2 = 1$.

Sometimes the state of a two-qubit system can be factored as the tensor product of the individual states of two qubits. For example, when $\alpha_{00} = \alpha_{10} = 1/2$ and $\alpha_{01} = \alpha_{11} = -i/2$ the state $\mid \psi\rangle$ is the tensor product of the states of the two qubits, $\mid \psi_1\rangle$ and $\mid \psi_2\rangle$

$$\begin{aligned}
\mid \psi\rangle &= \tfrac{1}{2}[\mid 00\rangle - i \mid 01\rangle + \mid 10\rangle - i \mid 11\rangle] \\
&= \tfrac{1}{2}[\mid 0\rangle \otimes (\mid 0\rangle - i \mid 1\rangle) + \mid 1\rangle \otimes (\mid 0\rangle - i \mid 1\rangle)] \\
&= \tfrac{1}{2}(\mid 0\rangle + \mid 1\rangle) \otimes (\mid 0\rangle - i \mid 1\rangle) \\
&= \mid \psi_1\rangle \otimes \mid \psi_2\rangle.
\end{aligned}$$

The individual states of the two qubits are well-defined

$$\mid \psi_1\rangle = \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \quad \text{and} \quad \mid \psi_2\rangle = \frac{1}{\sqrt{2}}(\mid 0\rangle - i \mid 1\rangle).$$

This factorization is not always feasible. For example, consider a special state of a two-qubit system when

$$\alpha_{00} = \alpha_{11} = 1/\sqrt{2} \quad \text{and} \quad \alpha_{01} = \alpha_{10} = 0.$$

The resulting state $\mid \beta_{00}\rangle$ is called a <u>Bell state</u> and the pair of qubits is called an <u>EPR pair</u>

$$\mid \beta_{00}\rangle = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}}.$$

There are three other *Bell states*

$$\mid \beta_{01}\rangle = \frac{\mid 01\rangle + \mid 10\rangle}{\sqrt{2}}, \quad \mid \beta_{10}\rangle = \frac{\mid 00\rangle - \mid 11\rangle}{\sqrt{2}}, \quad \text{and} \quad \mid \beta_{11}\rangle = \frac{\mid 01\rangle - \mid 10\rangle}{\sqrt{2}}.$$

These states form an orthonormal basis and can be distinguished from one another. Bell states are *entangled* states; the four states are called *maximally entangled* states and $\mid \beta_{11}\rangle$ is an *anti-correlated* state.

Entangled states are never in an ideal form; the source producing the entangled state is affected by noise and the communication channel used to transfer entangled states or the

quantum circuits used to manipulate the states can add noise. Thus we have to *purify*, or *distill* partial entanglement. Assume we have $N$ pairs of qubits, each pair in the same entangled but mixed state with density matrix $\rho$; the ensemble of pairs is in the state given by the $N$-fold tensor product $\rho^{\otimes N}$. Several protocols exist to transform the set of the $N$ pairs of partially entangled pairs to $M < N$ pairs of maximally entangled particles using only local operations and classical communication primitives [39, 40, 119].

Entangled states are affected by decoherence; *the environment conspires to disentangle the quantum systems we have prepared in an entangled state.* and we have to protect the information embodied by entangled quantum states. The fragility of quantum information requires that we encode quantum states and then manipulate them in an encoded form.

After this introduction of basic concepts of quantum physics we review several aspects of quantum information processing. We start with a discussion of quantum information and then introduces quantum computational devices and quantum algorithms.

## 1.13 Quantum Information

The unit of classical information is a *bit* with two possible values, $B = \{0, 1\}$; $n$ bits form a *register* $R = \{0, 1\}^n$. The state of a system can characterized by the $m = 2^n$ possible configurations of the bits in register $R$. The vector $(p_1 p_2 \ldots p_m)$ with $\sum_i p_i = 1$ and the $(m \times m)$ adjacency matrix $P$ describe the evolution of the system. A reversible transformation of the system state is described by a permutation matrix $P$, a matrix with only one non-zero element in each row and column; the new state is $P(p_1 p_2 \ldots p_m)^T$.

When the physical support of information is a quantum system we talk about *quantum information*. The unit of quantum information is a *qubit* abstracted as $\mathcal{C}^2 = \{| \ 0\rangle, | \ 1\rangle\}$; a register of $n$ qubits is $(\mathcal{C}^2)^{\otimes n}$ and the reversible transformation of the state of a quantum system are described by a *unitary* transformation. The density matrix $\rho$ allows the description of pure as well as mixed states with the probability of the individual components of the mix described by the vector $q_1 q_2 \ldots q_n$ and $\sum_i q_i = 1$; $\rho$ is a diagonal matrix with $q_i$ as its elements; a reversible transformation transforms the density matrix $\rho$ in $\rho'$ with $\rho' = P \rho P^\dagger$.

As expected, properties of quantum quantum systems such as entanglement and superposition, the inability to clone the state of a quantum system, the fact that a measurement is an irreversible process, and the instability of quantum states due to interactions with the environment have a profound effect on the ways we process quantum information. The question we pose now is if this new embodiment of information leads to special attributes that can be exploited in the process of manipulating information.

**Superposition.** The wave-particle duality characterizes the behavior of quantum systems; interference occurs as a manifestation of the wave behavior and is a consequence of the superposition of quantum states. Due to superposition we can compute the $2^n$ values of a function $f(x)$ with $x$ a binary $n$-tuple in one time step using a single copy of a quantum circuit implementing the transformation $f(x)$ of its input $x$. For example, consider a lengthy computation of a function $f(x)$ where the argument $x$ is one bit and the result is also one bit, $f(0)$ or $f(1)$. We are not interested in the actual value of the function, but only if the function is constant, $f(0) = f(1)$, or balanced, $f(0) \neq f(1)$. A sequential classical computer evaluates first $f(0)$, then $f(1)$, and finally compares the two to provide the answer; a parallel

classical computer evaluates $f(0)$ and $f(1)$ concurrently, but needs two processors for this task. A quantum computer with the input a superposition of $\mid 0\rangle$ and $\mid 1\rangle$ allows us to extract *global information* about the function $f$ and thus to solve Deutsch's problem in a single time step (the time needed to evaluate the function for one value of the argument), and with a single copy of the hardware, see Section 1.19.

We are now in a better position to understand Feynman's argument that the exact simulation of a quantum system is only possible with a quantum computer; assume that a state of the quantum system used for the exact quantum simulation can be expressed using $n$ bits with $n$ a large number, e.g., $n = 10^3$. Such a state is described by $2^n = 2^{1000} \approx 10^{333}$ complex numbers. No classical computer is able to store and process this colossal amount of data; on the other hand the exact simulation of the system can be carried out with a quantum computer, provided that we expect the answer either to be that the resulting state is identical with a reference state or not, or to provide a measure of their distance.

We are capable of simulating the behavior of quantum circuits using classical computers, a problem critical for the study of quantum fault-tolerance [1, 108, 439]. The simulation of quantum circuits is possible when the behavior of the quantum system is confined to a small region of the vast Hilbert space.

**Entanglement.** The *entanglement* of quantum systems has a profound impact on quantum information processing. Consider a composite quantum system consisting of two subsystems $A$ and $B$ with $n$ and $m$ qubits, respectively. The question we pose is if it is possible to reconstruct the state of the joint systems from measurements performed separately on one of the subsystems alone. The answer to this question was given by John Bell [25, 26] who established that the information about the quantum state of a composite system is contained in non-local correlations between the two subsystems; these non-local correlations cannot be revealed by any measurement performed on one of the subsystems alone.

We should point out that non-local correlations that cannot be revealed by measurements performed on only one of the component systems are not restricted to quantum phenomena. Think for example of two correlated random variables $X$ and $Y$ related to the behavior of two systems $A$ and $B$, respectively; there are no local measurements which allows either $A$ or $B$ to determine the joint probability density function, $p_{XY}(x, y)$. In addition to non-local correlations a critical property of the quantum entanglement is that a measurement of one of the two subsystems forces the other one to change its state, as we shall see when we discuss the EPR experiment in Section 2.16.

Entanglement has important implications for quantum error correction, a critical aspect of quantum information processing. Entanglement allows us to detect quantum errors without altering the state of the qubits in a quantum register, and then to correct the errors. First, we entangle the qubits in the register with ancilla qubits prepared in a well-defined state; as a result of the entanglement the ancilla qubits contain information about the qubits in error. Then we perform a non-demolition measurement of the ancilla qubits to determine the *error syndrome*. The error syndrome tells us if an error has occurred and if so which qubits were affected and what type of error was present. An undesirable consequence of entanglement is the occurrence of spatial and time-correlated quantum errors. Once a qubit is affected by an error, the error can propagate to other qubits correlated to the one in error; a time-correlated error re-occurs at a later time, on the same qubit, after we have corrected its initial occurrence.

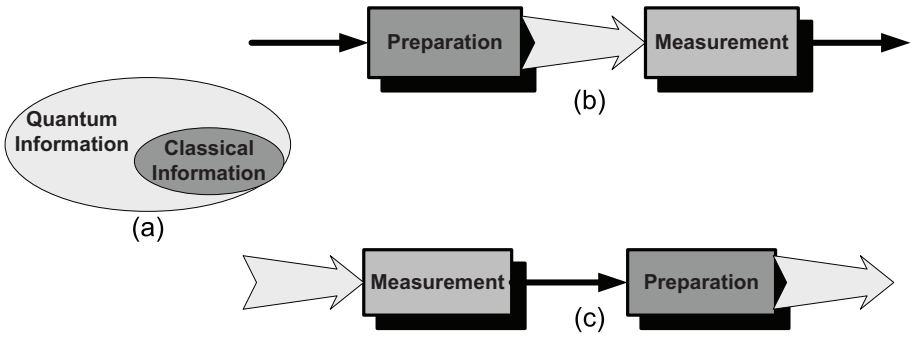



Figure 8: Classical and quantum information and conversion from one to another. Classical information is represented by thin arrows and quantum information as thick arrows. (a) Classical information can be regarded as a particular form of quantum information. (b) Classical information can be recovered from the quantum information when the preparation phase is followed by a measurement. The conversion path is: **classical** $\mapsto$ **quantum** $\mapsto$ **classical**. (c) Quantum information cannot be recovered when the preparation follows the measurement; the measurement is an irreversible process and alters the state of quantum systems. The conversion path in this case is **quantum** $\mapsto$ **classical** $\mapsto$ **quantum**.

**Measurements, preparation, and extraction of classical information from quantum information.** Classical information is independent of the medium used to transport it. Yet, classical information is often carried by the same types of particles as quantum information, e.g., electrons, or photons; why should we expect quantum information to be different from classical information?

To answer this question we should establish if it is possible to freely convert one type of information to another and then recover the original information [231]. We can convert classical to quantum information and then convert back the quantum information to the original classical information. Formally, this process consists of two stages: *preparation*, when the quantum information is generated from the classical one and *measurements*, when classical information is obtained from the quantum information, Figure 8(a).

The remaining question is if we can convert quantum to classical information and then convert the classical information to quantum information indistinguishable from the original one; this means that we should first perform a measurement to extract classical information and then use it to prepare quantum information, Figure 8(b). The only possibility to compare quantum mechanical systems is in terms of statistical experiments and this is not possible, since a measurement is an irreversible process, it alters the state of a quantum system; Chapter 2, devoted to quantum measurements, covers the arguments supporting this statement.

We conclude that indeed, quantum information is qualitatively different from the classical one. Even though classical and quantum information can be carried out by the same types of particles, the physical processes are different; in the former case the physical processes are subject to classical Physics and in the later to quantum Physics. As we shall see in Chapter 2 some states of a quantum system, the *pure states*, have a classical counterpart and a measurement allows us to distinguish orthogonal pure states. Therefore, classical information can be regarded as a particular form of quantum information, Figure 8 (c).

**Manipulation of quantum states.** Now we turn our attention to the the question if the states of a quantum system are stable for sufficiently long periods of time to allow the transformations prescribed by a quantum algorithm to progress without hinderance. Unfortunately, only rarely the state of a quantum system can be considered as being stable, more often they are unstable. For example, the famous state of Schrödinger's cat is a "superposition" of being at the same time "dead" and "alive" formally described as $|\text{ cat}\rangle = \frac{1}{\sqrt{2}}(|\text{ dead}\rangle + |\text{ alive}\rangle)$.

This state is possible in Quantum Mechanics, but never observed in practice; all the cats we have ever seen, or will ever see were and will be either dead or alive. The instability of quantum states is due to the interactions of the system with the environment; the correlation between a quantum system and the environment are very strong and lead to the phenomenon of *decoherence* which hinders quantum information processing.

The quantum information we wish to process is encoded as correlations among the components of the quantum system; the interactions of the environment with the quantum system erode in time these useful correlations which are transformed to correlations between the quantum system and the environment. The decoherence is only one of the problems we have to address when we think about building quantum computing and communication devices.

Another problem is the *accumulation of errors*, a problem we are familiar with from the study of classical analogue circuits; if each analog circuit performing the transformation $T_i$ introduces an error $\epsilon_i$ then, after $N$ steps, instead of the desired transformation, $\prod_{i=1}^{N} T_i(I)$ the transformation of the input $I$ will be $\prod_{i=1}^{N} T_i(1+\epsilon_i)(I)$. Similarly, if each quantum gate introduces a small error $\epsilon$ then, after $1/\epsilon$ gates, the cumulated error will be significant enough to affect the result.

In addition to the bit-flip errors we are familiar with from the study of fault-tolerance of classical systems, we have to deal with phase-flip errors and with combinations of bit- and phase-flip errors of a quantum system. This brief discussion motivates the attention paid to quantum error correction, one of the topics of this book.

**Physical embodiment of quantum information.** The quantum analog of a classical bit is called a *qubit*; following our mantra that "information is physical" we link the abstraction called qubit with the physical reality and provide a glimpse at some of the properties of quantum particles we have to manipulate in order to process quantum information.

Quantum particles have some properties with no classical counterpart. For example, the spin is an intrinsic angular momentum[6] of a quantum particle, related to its intrinsic rotation about an arbitrary direction. The spin of a quantum particle can be observed as the result of the interaction of the intrinsic angular momentum of the particle with an external magnetic field $\mathbf{B}$. Classical Physics operates with the concept of angular momentum arising from a rotation around a well-defined axis of a body.

There are two classes of quantum particles, those with spin multiple of one-half, called fermions, and those with spin multiple of one, called bosons. The spin quantum number of fermions can be $s = +1/2$, $s = -1/2$, or an odd multiple of $s = \pm 1/2$. Electrons, protons, and neutrons are fermions; the spin quantum number of bosons can be $s = +1$, $s = -1$, $s = 0$, or a multiple of $\pm 1$.

A quantum particle such as the electron is not a "body" in the classical sense and does not

[6]The intrinsic angular momentum of a quantum particle should be distinguished from its orbital angular momentum.

have a defined axis of rotation. The electron is characterized by a charge with a non-stationary spatial distribution. The variation in time of this charge distribution can be associated with an intrinsic rotation of the electron about directions randomly oriented in space.

One possible embodiment of a qubit is the spin of the electron, the quantum number characterizing the intrinsic angular momentum of the electron; the electron spin has either the value $+1/2$ or $-1/2$ along the measurement axis, regardless of what that axis is. The introduction of this two-valued quantum number for the electron led Pauli to postulate his *exclusion principle*. According to Pauli's exclusion principle, no more than two electrons can occupy the same "orbit" and those two electrons must have anti-parallel spins.

Indistinguishability is a principle of Quantum Mechanics saying that all quantum particles of the same type are alike; for example, we cannot distinguish an electron from another. Therefore the operation of swapping the position of two electrons in a systems with many electrons leaves the system's state unchanged or, in other words, the operation is symmetric and it is represented by a unitary transformation acting on the wave function as discussed in Section 1.7. In three dimensions an exchange of two bosons is represented by an identity operator; the wave function is invariant and we say that the particles obey Bose statistics. The exchange of two fermions in three dimensions changes the sign of the wave function; the particles are said to obey Fermi statistics.

A photon[7], a particle of light, is another important two-state quantum system used to embody a qubit; the quantum information can be encoded as the polarization of the photon. Photons differ from the 1/2-spin electrons in two ways: (1) they are massless and (2) they have spin one. A photon is characterized by its vector momentum (the vector momentum determines the frequency) and its polarization.

Light has a dual nature, wave-like and corpuscular-like; as an *electromagnetic radiation*, light consists of an electric and a magnetic field perpendicular to each other and, at the same time, perpendicular to the direction the energy is transported by the electromagnetic wave; the electric field oscillates in a plane perpendicular to the direction of light and the way the electric field vector travels in this plane defines the polarization of light.

When the end of the electric field vector oscillates along a straight line, we say that the light is *linearly polarized*. When the end of the electric field vector moves along an ellipse, the light is *elliptically polarized* and when it moves around a circle, the light is *circularly polarized*. If the light comes toward us and the end of the electric field vector moves around in a counterclockwise direction, we say that the light has *right-hand polarization*; if it moves in a clockwise direction, we say that the light has *left-hand polarization*.

The last embodiment of quantum information we discuss are the anyons, quantum particles of interest for the topological quantum computing. Anyons are indistinguishable particles defined in two dimensions; they are different from either fermions or bosons. Consider a gas of electrons squeezed between two slabs of semiconductor materials such that, the movement of electrons is restricted to two dimensions only. At very low temperature and in a strong magnetic field the two-dimensional electron gas has a strongly entangled ground (lowest energy) state separated from all other states by an energy gap. This lowest-energy state carry an electric charge that is not an integer multiple of the electron charge and does not have the quantum numbers associated with electrons [345]. The properties of the anyons manifest themselves as the Fractional Quantum Hall Effect (FQHE) discovered by Daniel Tsui, Horst

[7]The name photon comes from the Greek word "photos" meaning light.

Störmer and Arthur Gossard [425]; the FQHE is discussed in Section 6.15.

We have only mentioned three possible embodiments of quantum information, others are discussed in detail in Chapter 6 that covers physical realizations of quantum computing and communication devices. The diversity of potential implementations of quantum information processing devices is a source of optimism and, at the same time, anxiety. On one hand, it gives us great hope that at least some of the theoretical ideas will ultimately lead to feasible quantum information processing devices; on the other hand, it gives as some feeling about the vastness of the field and the depth of knowledge required to master a discipline involving multiple areas of Mathematics, Physics, and Computer Science.

**Quantum information processing systems.** The last subject of our survey of quantum information covers quantum information processing systems. A quantum computer is a physical device capable of transforming quantum information as required by a quantum algorithm; a quantum communication system transmits either classical, or quantum information from one place to another through a quantum channel.

The operation of a quantum computer takes advantage of fundamental principles of quantum Physics: superposition, quantum interference, quantum entanglement, and the high dimensionality of the state space of a quantum system to solve some computationally "hard" problems efficiently. A quantum algorithm aims to increase the probability of obtaining the correct answer by arranging that all computational path leading to the right answer interfere constructively, while the computational paths to wrong answers interfere destructively. This strategy has been applied successfully for solving a subset of "hard" computational problems such as factoring large numbers and searching large unstructured databases; the quantum algorithms of Shor [380] and Grover [183] are milestones in quantum computing.

To process information with a quantum system we first prepare the system in an initial quantum state, then transform this state through a set of operations prescribed by a quantum algorithm, and, finally, measure the resulting state of the system. This sequence of steps resembles the ones followed by a classical device up to the point when we examine the result. The qualitative difference between a classical system outputting the result of a transformation and the measurement of the final state of the quantum system is due to the *randomness of the quantum measurement process.* If the result of the quantum computation is one qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ then a measurement of the qubit will produce the result 0 with probability $p_0 =\mid \alpha_0 \mid^2$ or 1 with probability $p_1 =\mid \alpha_2 \mid^2$.

A quantum algorithm generates a probability distribution of the results while a randomizing algorithm for a classical computer uses randomness as part of its logic. There are qualitative differences between randomization algorithms and quantum algorithms. The former are designed with the hope of achieving good performance in the "average case" over all possible random choices; the later exploit the entanglement and superposition of quantum information, the randomness is not part of their logic, but of the physical transformations leading to the result.

In [265] Landauer observes that classical computers function based on the intuitive and completely abstract computability theory due to Church, Turing, Post, and Gödel, while "the properties of quantum computers are not postulated in abstracto, but deduced from the laws of physics" [118]. A quantum computer coupled with a quantum communication system performs functions related to quantum cryptography that are not feasible with a classical system; however, a quantum computer can only perform computations that can be carried

out by a classical computer, albeit much more efficiently.

A quantum communication channel transmits information using quantum effects when, for example, the information is encoded as the spin of an electron or in the polarization of a photon. Eavesdropping on a quantum communication channel can be detected with a very high probability because the effect of intrusion is an irreversible transformation, a measurement of the quantum state.

We conclude that the special properties of quantum information have important consequences for computation and for communication and we should investigate the the physical realization of quantum information processing devices.

## 1.14 Physical Realization of Quantum Information Processing Systems

A quantum computer is a physical device designed to transform quantum information embodied by the state of a quantum system. The physical processes required to transform the quantum state in a controlled manner are different for ion traps, solid-state, optical, Nuclear Magnetic Resonance (NMR), or other possible physical implementations. Thus, we need first to define the basic requirements to build a quantum computer regardless of the physical phenomena involved. DiVincenzo [131] provides a crisp and clear formulation of five plus two (additional) requirements (discussed in depth in Section 6.1) for a quantum information processing system:

1. A scalable physical system with well-characterized qubits.
2. The ability to initialize the state of the qubits.
3. Long relevant decoherence times, much longer than the gate operation times.
4. A "universal" set of quantum gates.
5. A qubit-specific measurement capability.

The two additional requirements are related to the ability to communicate:

1. The ability to inter-convert "stationary" (memory) and "flying" (communication) qubits.
2. The ability to faithfully transmit "flying" qubits between specified locations.

These requirements come naturally to mind, they have an immediate correspondent for classical computers. Indeed, we cannot conceive a state of the art computer built with circuits whose state cannot be controlled or be initialed to a desired state. It would be impractical to build a computer unless we have a finite set of building blocks and it seems obvious that we should have access to the results of a computation.

A <u>well-characterized</u> qubit means that the relevant parameters of the physical process, including the Hamiltonian $\mathbf{H}(t)$ of the system and the coupling between a qubit and the

environment and between this qubit and other qubits, must be known. For an implementation of a quantum computer we should only consider those quantum systems which satisfy this requirement. For example, the Super Selection Rules (SSR)[8] prohibit entangled states involving different particle numbers thus, a two-qubit system consisting of two quantum dots and an electron in a superposition state as being on one or the other quantum dot would not satisfy the first requirement [131].

The scalability requirement is not only related to the ability to carry out complex transformations, but also to the necessity, discussed later, of having reliable circuits; as we shall see shortly, every "useful qubit" requires a large number of additional qubits, 10 or more, to ensure fault-tolerance.

It is self-evident that we should prepare a qubit in a well-defined initial state in order to carry out a quantum computation with predictable results. An $n$-qubit quantum computer operates on a $2^n$-dimensional Hilbert space $\mathcal{H}_{2^n}$ with computational basis states $\mid x_1 x_2 \ldots x_n\rangle$ with $x_i = 0$ or $x_i = 1$. This requirement translates to the fact that *any computational basis state* $\mid x_1 x_2 \ldots x_n\rangle$ *should be prepared in at most* $n$ *steps.* The practical questions are how to initialize a qubit in a state with maximal entropy and how fast the preparation of qubits can be done. A possible solution is to have an independent system to prepare qubits in an initial state either by forcing them to the ground state of the Hamiltonian, a process called "cooling," or by performing a measurement which collapses the state of a qubit to a basis state, e.g., $\mid 0\rangle$. Then, a "qubit conveyor belt" should provide access to the qubits in the initial state whenever they are needed.

Decoherence in its simplest form means that a "pure state" $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ is transformed by interaction with the environment to a "mixed state" with density matrix $\rho = \mid \alpha_0 \mid^2 \mid 0\rangle\langle 0 \mid + \mid \alpha_1 \mid^2 \mid 1\rangle\langle 1 \mid$. The pure state of a qubit is represented by a vector connecting the origin with a point on the surface of the Bloch sphere, while a mixed state of a qubit is represented by a vector connecting the origin to a point inside the Bloch (solid) sphere; the von Neumann entropy of a mixed state is measured by the distance of such a point to the surface of the Bloch sphere, as discussed in Section 2.3.

Since the early days of computing, reliability has been a major concern [80] therefore, it seems reasonable to ask ourselves if a reliable quantum computer could be built at all, knowing that quantum states are subject to decoherence. The initial thought was that a quantum computation could only be carried out successfully if its duration were shorter than the decoherence time of the quantum computer. As we shall see in Section 5.25, the decoherence time ranges from about $10^4$ seconds for the nuclear spin, to $10^{-9}$ seconds for the quantum dot charge. Thus, it seemed very problematic that a quantum computer could be built unless we have a mechanism to periodically deal with errors.

Probabilistic classical computer and communication systems can be error corrected and operate effectively in the presence of noise without requiring an exponential precision while the operation of their analog counterparts is conditioned by a lack of noise and an exponential precision. The analog nature of quantum information and of quantum transformation raised serious questions for the experimental realization of quantum information processing systems.

[8] An SSR is a restriction on the allowed local operations on a system, not on its allowed states, and it is associated with a group of physical transformations [24]. Such restrictions could be imposed by the properties of the underlying theory, or arise due to physical restrictions. The operations it applies include unitary transformations, $\mathbf{O}\rho = U\rho U^\dagger$, and measurements, $\mathbf{O}_r\rho = \mathcal{M}_r\rho\mathcal{M}_r^\dagger$ with $\sum \mathcal{M}_r\mathcal{M}_r^\dagger = 1$.

These questions were answered in [381], [403], [246], [40], [247], [248] showing that quantum computing and communication systems are more similar to probabilistic classical systems than to analog ones.

Now we know that quantum error correcting codes could be used to ensure fault-tolerant quantum computing; quantum error correction allows us to deal algorithmically with decoherence. A quantum error correcting code maps a "logical qubit" onto several "physical qubits" and then encodes these qubits using two classical error correcting codes to deal with bit-flip and phase-flip errors, as discussed in detail in Chapter 5. There is a significant price to pay to achieve fault-tolerance through error correction, the number of qubits required to correct errors is an order of magnitude larger than the number of "useful" qubits.

In 1996, Shor [383] showed how to perform reliable quantum computations when the probability of a qubit or quantum gate error decays polylogarithmically with the size of the computation, a rather unrealistic assumption. The quantum threshold theorem states that arbitrarily long computations can be carried out with high reliability provided that the error rate is below an accuracy threshold according to Knill, Laflamme, and Zurek [247]. In 1999, Aharonow and Ben-Or [5] proved that reliable computing is possible when the error rate is smaller than a constant threshold, but the cost is polylogarithmic in time and space. In practice, error correction is successful for a quantum system whose decoherence time is four to five orders of magnitude larger than the *gate time*, the time required for a quantum gate operation.

In the next section we discuss in more detail the practical requirements regarding universal quantum gates. Here we only note that the fact that any logic function could be implemented using a small set of universal classical gates had a significant impact on current solid-state technology; it allowed us to reduce the size of classical circuits as predicted by Moore's law which states that the number of transistors per chip that yields the minimum cost per transistor doubles every 18 months or so.

An important aspect of a discussion regarding universal quantum gates is that many-body quantum interactions are difficult to control and to analyze. Thus, it is necessary to have a universal set of quantum gates which require at most two-body interactions, and this translates to the ability to implement any unitary transformation using only one-qubit and two-qubit gates. *Any unitary transformation can be approximated arbitrarily well by a quantum circuit consisting of two-qubit* `CNOT` *gates, and one-qubit* `H, T,` *and* `S` *gates.* First, we show that any unitary transformation $\mathbf{U} \in \mathcal{H}_{2^n}$ can be carried out by a set of unitary transformations $U_1, U_2, \ldots U_k, \ldots U_m$, with $m \leq 2^n - 1$, which act only on two or fewer computational basis states. Then, we show that each unitary transformation $U_k,\ 1 \leq k \leq m$, can be expressed exactly as a product of transformations carried out by `CNOT` and one-qubit gates, and finally, we show that the transformation carried out by an one-qubit gate can be approximated arbitrarily well by the transformations carried out by `H, T,` and `S` gates. As an example we show how to implement a Toffoli gate with the set of universal quantum gates described above.

Consider a quantum circuit that consists of $m$ gates $G_k, 1 \leq k \leq m$, which carry out the unitary transformations $\mathcal{G}_k = e^{i\mathbf{H}_k(t)/\hbar}, 1 \leq k \leq m$, with $\mathbf{H}_k(t)$ - the Hamiltonian describing the evolution of the quantum system when implementing the $k$ gate and $\hbar$ - Planck's constant. The simplest modus operandi of the circuit is to consider discrete times $t_1, t_2, \ldots t_k, \ldots t_m, t_{m+1}$ when each Hamiltonian is turned on and off. For example, $\mathbf{H}_k$ should be turned on at $t_k$ and

tuned off before $t_{k+1}$ when $\mathbf{H}_{k+1}$ is turned on.

Recall that the control unit of a classical processor decides which functional units are active, and what specific function each unit is expected to carry out at time $t$. An important question we have to answer is how to control each quantum gate; who plays the role of the "control unit" of a classical computer in a quantum computing setup? The answer is that the individual quantum circuits are controlled by a physical apparatus which regulates the electric and magnetic fields when the quantum gates are implemented on ion-traps or NMR, or which regulates the temperature of the quantum bath for quantum dots, and so on.

We are able to examine the final results as well as partial results of a computation carried out by a classical computer. Classical error correction techniques require the ability to examine some information derived from a sequence of bits stored or transmitted through a communication channel, e.g., the syndrome of linear codes. Therefore, it is natural to require the ability to measure the state of qubits. When the density matrix of a qubit is

$$\rho = p \mid 0\rangle\langle 0 \mid + (1-p) \mid 1\rangle\langle 1 \mid + \alpha \mid 0\rangle\langle 1 \mid + \alpha^* \mid 1\rangle\langle 0 \mid,$$

with $\alpha \in \mathbb{C}$, then an "ideal measurement" should provide an outcome of "0" with probability $p$ and an outcome of "1" with probability $(1-p)$, regardless of the value of $\alpha$, regardless of the state of the neighboring qubits, and without changing the state of the quantum computer. In this case a "nondemolition measurement" leaves the qubit in state with $\rho = \mid 0\rangle\langle 0 \mid$ after reporting the outcome 0 and leaves it in state with $\rho = \mid 1\rangle\langle 1 \mid$ after reporting the outcome 1.

We know that a measurement of a quantum state collapses the state to one of the basis states, the measurement is an irreversible process. This poses serious challenges to the design of quantum algorithms which should avoid irreversible transformations. Quantum error correction requires nondemolition measurements of the error syndrome to preserve the state of the physical qubits.

Lastly, the two additional requirements put forth by DiVincenzo recognize the inherent symbiosis between computation and communication. This is also true for quantum information; indeed, quantum information may be processed or stored at a different location than the one it is collected from, quantum computers may need to exchange information among themselves. There is a general agreement that photons and optical communication channels are ideal for transporting "flying qubits."

We conclude that a quantum computer is an ensemble consisting of a classical and a quantum component and the quantum component must satisfy the requirements outlined by DiVincenzo [131]. A quantum computer consists of quantum circuits and quantum circuits are built by interconnecting quantum gates.

## 1.15 Universal Computers; The Circuit Model of Computation

Informally, a universal computer is a single machine able to perform any physically possible computation. The classical theory of computability, as well as the quantum computability theory, admit the existence of universal computers. Moreover, the concept of universal computer developed in the context of the computability theory has immediate practical consequences, it implies that the components used to build such an instrument must also be universal. Indeed, if we could obtain the solution to a problem using a physical computing instrument, without

having a systematic method to produce that instrument using a set of universal components, than the solution would not necessarily be "computable" in a useful sense [118].

The universality of quantum computers implies *the ability to carry out computations of arbitrary complexity, in other words, computations involving an arbitrary number of operations and an arbitrary amount of storage.* We should be able to maintain the computer in operation for arbitrary long periods of time and provide an arbitrary large number of qubits in a standard state. The last requirement is equivalent to the unlimited "blank tape" of Turing Machines.

A classical computing engine is a deterministic system evolving from an initial/input state to a final/output state. The input state and all states traversed by the system during its dynamic evolution have some canonical labelling. Moreover, the label of each state can be measured by an external observer. The label of the final state is a deterministic function $f$ of the label of the input state; we say that the engine computes a "function" $f$. Two classical computing engines are *equivalent* if they compute the same function $f$ given the same labelling of their input and output states [115].

While early computers could only execute a fixed program, the computers we discuss are universal, they implement an *instruction set* and can carry out any computation described by a *program* expressed as a set of instructions from this set. The most popular architecture of a *universal computer* was proposed by John von Neumann [80]: a stored-program computer with four interconnected components: (i) a control unit, (ii) an arithmetic and logic unit (ALU), (iii) memory, and (iv) an Input/Output unit.

Modern computers process classical information and obey the laws of classical Physics which, ultimately, limit our ability to increase the speed of solid-state circuits and to make them increasingly smaller. The power dissipation of classical circuits increases as $\kappa^\alpha$ where $\kappa$ is the clock rate and $2 \leq \alpha \leq 3$; when we double the speed, the power dissipation may increase by a factor of $2^3 = 8$. Heat dissipation for modern computers operating at clock rates of few GHz is an extremely challenging problem and has forced manufacturers of microprocessors to switch their attention from increasing the clock rate to building multi-core systems.

A classical universal computer is built from *classical gates*, see Figure 9; a gate implements a Boolean function $f : \{0,1\}^n \mapsto \{0,1\}^m$, Figure 9 (a). The classical gates used in existing computers perform irreversible transformations and the transformations carried out by these gates produce a substantial amount of "informational junk." This information must be erased, a dissipative process responsible for a substantial amount of heat; heat removal is a major problem for classical processors.

**Classical gates.** Practical considerations require a computer to be built using gates from a small set, in other words require the existence of one or more universal sets of classical gates. Formally, we say that a set of classical gates is *universal* if every Boolean function can be implemented using only gates from that family.

We can *use a small number of gates to compute any Boolean function.* For simplicity in our presentation we consider a restricted family of Boolean functions with $n$ bits as input and a single output bit: $f : \{0,1\}^n \mapsto \{0,1\}$ and provide a proof by induction that one can construct a circuit for this class of Boolean functions using a small number of different logical gates. First, we consider the case $n = 1$ and realize that we need four circuits to map the two possible inputs to the outputs, Figure 10(a): (i) one that performs the identity transformation $y = x$ - this can be done with a single wire; (ii) one that flips the input $x$, $x = 0 \rightarrow y = 1,\ x = 1 \rightarrow y = 0$ - this can be done with a `NOT` gate; (iii) one that produces

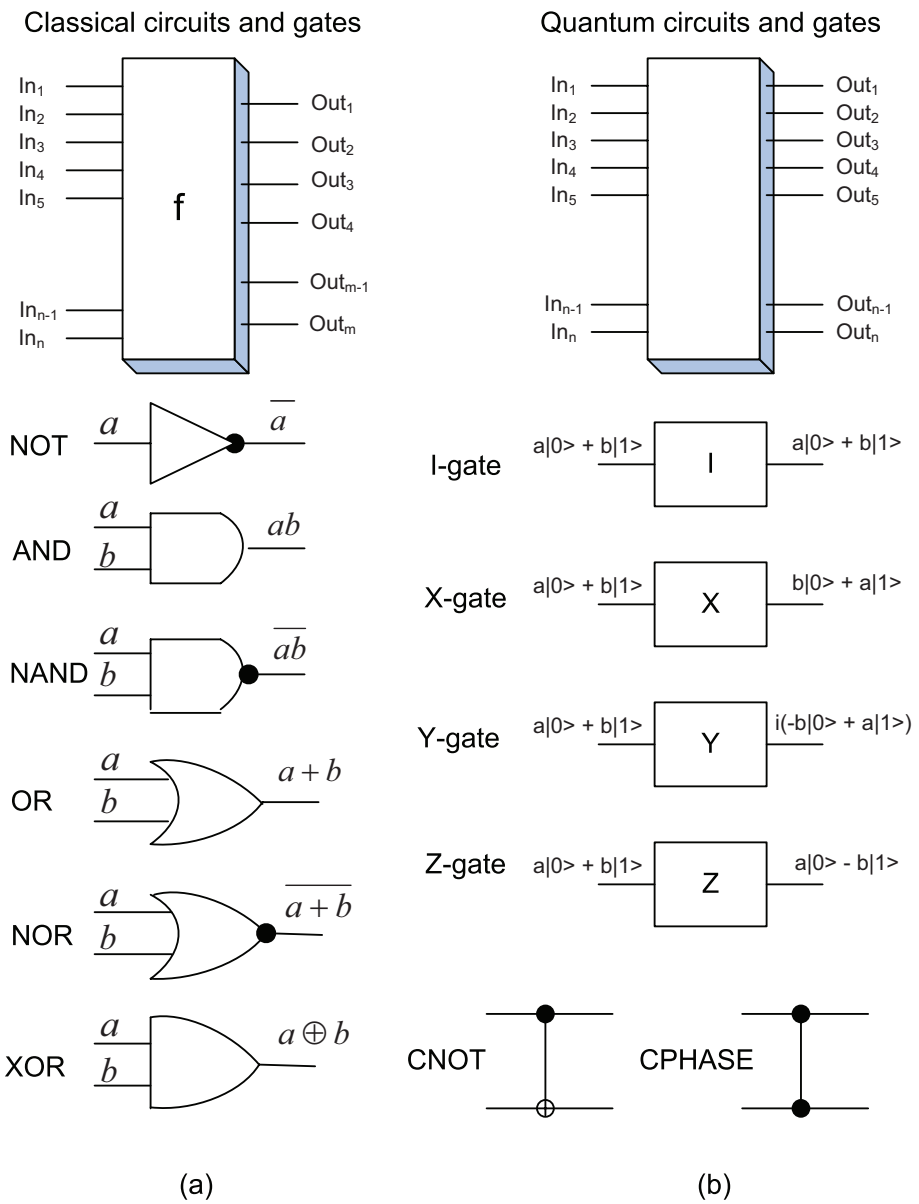



Figure 9: Classical and quantum circuits and gates. (a) A classical circuit implements a multi-output Boolean functions $f : \{0,1\}^n \mapsto \{0,1\}^m$ and is constructed using a finite number of classical gates. $n$, the number of inputs and $m$, the number of outputs of a classical circuit may be, and often are, different, $n \neq m$. The `NOT`, `AND`, `NAND`, `OR`, `NOR`, and `XOR`, classical gates are shown. $a$ and $b$ are Boolean variables, $\bar{a}$ is the negation of $a$; $(a+b)$ is the logical sum $a$OR$b$, also written as $a \vee b$; $ab$ is the logical product $a$AND$b$, also written as $a \wedge b$; $a \oplus b$ is the exclusive OR (XOR) of $a$ and $b$. (b) A quantum circuit performs a unitary operation in the Hilbert space $\mathcal{H}_n$ and consists of a finite collection of quantum gates; each quantum gate implements a unitary transformation on $k$ qubits for a small $k$ and must have the same number of inputs and outputs. The one-qubit gates with for the Pauli transformations, $\sigma_I, \sigma_x, \sigma_y$ and $\sigma_z$ and two-qubit gates the `CNOT` and the `CPHASE` are shown.

$y = 0$ regardless of the input - this can be done with an `AND` gate with the two inputs 0 and $x$; and (iv) one that produces $y = 1$ regardless of the input - this can be done with an `OR` gate with the two inputs 1 and $x$. Then we assume that we can construct the circuit able to compute a function $f_n(b_1, b_2, \ldots, b_n)$ on $n$ bits and we wish to prove that using `NOT, AND,` and `XOR` gates we can construct a circuit able to compute $f_{n+1}(b_0, b_1, b_2, \ldots, b_n)$ on $(n+1)$ bits. Call $f^{(0)}_{n+1} = f_{n+1}(0, b_1, b_2, \ldots, b_n)$ and $f^{(1)}_{n+1} = f_{n+1}(1, b_1, b_2, \ldots, b_n)$; then the circuit in Figure 10(b) will compute $f_{n+1}(b_0, b_1, b_2, \ldots, b_n)$.

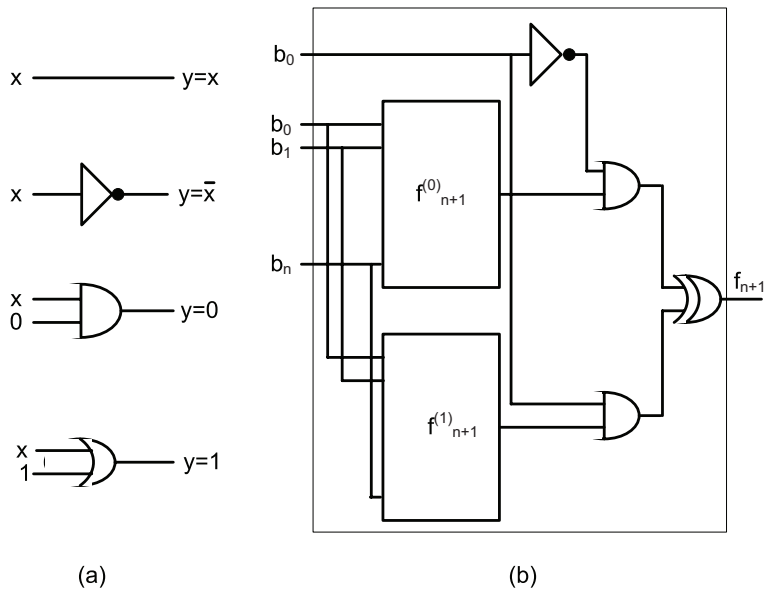



Figure 10: (a) The four circuits mapping an input $x$ to an output $y$. The circuits are able to compute a Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ for n=1: (i) a wire for the identity $y = x$; (ii) a bit flip implemented with a `NOT gate`; (iii) a circuit that produces $y = 0$ regardless of the input, using an `AND` gate with the two inputs 0 and $x$; and (iv) a circuit that produces $y = 1$ regardless of the input, using an `OR` gate with the two inputs 1 and $x$. (b) The circuit to compute a function $f_{n+1}(b_0, b_1, b_2, \ldots, b_n)$ on $(n+1)$ bits given that we can construct the circuit able to compute a function $f_n(b_1, b_2, \ldots, b_n)$ on $n$ bits; the circuit uses `CNOT, AND` and `XOR gates`.

To build a classical circuit implementing a Boolean function we need: (1) wires; (2) gates of several types, `NOT, AND, XOR`; (3) the ability to replicate the input, an operation called fanout; (4) the ability to interchange two bits, an operation called swapping; and last, but not least (5) a supply of ancilla, or auxiliary bits, e.g., the bits initialized to 0 or to 1 for the circuits in Figure 10(a).

**Reversibility and computation.** Logical reversibility - a Boolean function is reversible if there is a one-to-one relationship between input and output. In the early 1960s Landauer established that a necessary condition for a computational process to be physically reversible is that the logical function it implements to be logically reversible [265]; he showed that any irreversible computation may be transformed into a reversible one by embedding it into a larger computation where no information is lost. The next step, Bennett's discovery that

computations can be done reversibly [30] led to the investigation of classical reversible gates. Then, in 1980 Toffoli showed that classical reversible gates can be used to construct classical circuits [422].

For example, instead of the classical `XOR` one could use the `CNOT` (controlled-NOT) gate. The `CNOT` is reversible and has two inputs, $a$ and $b$; the source, or "control" bit, $a$ is not changed and affects the setting of the "target" bit $b$, when $a = 1$ then $b$ is flipped. Toffoli also showed that we can replace an irreversible `NAND` gate with a reversible Toffoli gate. The classical Toffoli gate has three inputs, $(a, b, c)$ and three outputs $(a, b, (c \oplus ab))$ where $\oplus$ stands for binary addition, Figure 11(a). The `NAND` gate transforms its two inputs, $a$ and $b$ to their logical product and negates it; its output is $(\overline{ab})$, an expression also written as $\neg(a \wedge b)$, or as $(a \text{ NAND } b)$.

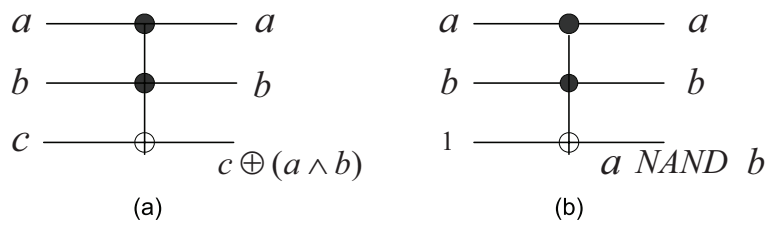



Figure 11: The classical Toffoli gate is universal for reversible Boolean logic. (a) The gate has three inputs, $(a, b, c)$ and three outputs $(a, b, (c \oplus (ab)))$ where $\oplus$ stands for binary addition; $a$ and $b$ are "control" bits and $c$ is a "target" bit. The two control bits are unchanged and the target is flipped when at least one of the control bits is 1. (b) The output of a Toffoli gate with $c = 1$ is $a \wedge b$ or $a$ NAND $b$. The `NAND` gate transforms its two inputs, $a$ and $b$ to their logical product and negates it; its output is $\overline{ab}$, also written as $\neg(a \wedge b)$, or as $(a \text{ NAND } b)$.

The classical Toffoli gate is universal for reversible Boolean logic: indeed, we know that the `NAND` gate is universal, any Boolean function can be implemented using only `NAND`s; we see immediately that the output of a Toffoli gate with $c = 1$ is $a \wedge b$ or $a\text{NAND}b$, Figure 11(b) thus, the Toffoli gate in this configuration simulates a `NAND` gate. This simulation is not very effective as the number of additional bits introduced in a circuit constructed solely with `NAND` gates, increases linearly with the number of gates.

**The circuit model of computation.** The circuit model of computation relates the practical implementation of an algorithm with the circuits to compute Boolean functions; in terms of computational power this model is equivalent to the Turing Machine model. This statement requires some qualifications as the circuit model deals with practical implementation thus, with finite systems, while the Turing Machine model assumes unbounded resources, e.g., an infinite tape. The circuit model should also be able to distinguish between computable and non-computable functions.

So far we have been concerned with the ability of the circuit model of computation to express an algorithm. To construct the circuit for a particular algorithm we have to design a protocol telling us what types of classical gates are needed, how to interconnect them, and how to locate the results of the computations. The need for the circuit design protocol leads to a deeper connection between the circuit computational model and the Turing Machine

computational model. We should find the means to prohibit the ability to "hide" the complexity of an algorithm in the protocol for building the circuit, or even to consider embedding a non-computable function e.g., a function for the "halting problem," in the protocol.

To address the problem of hiding the complexity of an algorithm in the computational system the circuit model of computation uses the concept of uniform circuit family. A circuit in the family, denoted as $\mathcal{C}_n$, has $n$ input bits and any number of auxiliary and output bits; when the input is a string $x$ of $n$ bits the output of the circuit is denoted as $\mathcal{C}_n(x)$. A circuit is consistent if when the input is a string of length $m < n$ then $\mathcal{C}_m(x) = \mathcal{C}_n(x)$. A circuit family is uniform and then it is denoted as $\{\mathcal{C}_n\}$ if an algorithm for a Turing Machine to generate a description of the protocol for the design of the circuit, given $n$, the size of the input, exists. If such an algorithm for a Turing Machine does not exist then the circuit family is called non-uniform. The equivalence of the circuit and the Turing computational models is restricted to uniform circuit families.

We assume that a computing machine $\mathcal{M}$ computes one function $f$ of its input to produce the desired result. To hide the complexity we could first alter the input and then present it to another machine $\mathcal{M}'$ able to produce the same result in polynomial time. To address this problem we consider the input as consisting of two parts: (a) a program/protocol determining which function the computing machine $\mathcal{M}$ will compute and (b) the actual input for the desired function. We require the existence of an algorithm for a Turing Machine to generate a description of the protocol.

The full specification of the state for the Turing Machine computational model as well as for the circuit computational model requires the specification of a set of numbers, all expected to be measurable at any instant of time. Quantum physics excludes the existence of physical systems with this property thus, these computational models are effectively classical and we have to address the question of quantum computational models from a new perspective.

Computational problems are said to be "hard" if the time to solve the problem is extremely long, regardless of the hardware used. When the time $t(n)$ to carry out a computation with $n$ data elements as input satisfies the condition $t(n) \leq \text{Poly}(n)$, where $\text{Poly}(n)$ is a polynomial in $n$, then we have a a polynomial time computation; otherwise it is an exponential computation. Complexity theory, the branch of computer science addressing the question which problems are hard and which ones are easy, defines "easy" problems as *polynomial time* and "hard" problems as *exponential time.*

In Section 1.13 we outlined the arguments that classical computers are not powerful enough to solve efficiently problems such as factorization of large integers, or "efficient" simulation of physical systems at atomic and subatomic level. We also discussed the Church-Turing principle namely that every function which could be regarded as computable is computable by a Turing Machine. A quantitative version of the Church-Turing principle relates the behavior of the abstract model of computation provided by the Turing Machine concept with the physical computing devices used to carry out a computation. This thesis states that: "any physical computing device can be simulated by a Turing Machine in a number of steps polynomial in the resources used by the computing device" [320]. While no one has been able to find counter-examples for this thesis, the search has been limited to systems that are based on the laws of classical Physics. Yet, our universe is essentially quantum mechanical, therefore, there is a possibility that the computing power of quantum computing devices might be greater than the computing power of classical computing devices [382].

## 1.16 Quantum Gates, Circuits, and Quantum Computers

**Quantum gates.** The building blocks of a quantum computer are quantum gates. Each quantum gate implements a unitary transformation on $k$ qubits for a small $k$ Figure 9(b). The fanout - the ability of a logic gate output to drive a number of inputs of other logic gates to form more complex circuits is non-trivial for quantum gates; quantum gates implement unitary thus, reversible transformations, and are required to have the same number of input and output qubits, a strict rule that we do not impose on classical gates which may have different numbers of inputs and outputs.

In 1980, Paul Benioff realized that the Hamiltonian time evolution of an isolated quantum system is reversible and could mimic a reversible Boolean computation [27]. Few years later, in 1985 David Deutsch observed that the linearity of the Schrödinger equation implies that *mapping of the basis states uniquely specifies the dynamics of an arbitrary initial state* [115].

A one-qubit gate carries out a unitary transformation $\mathbf{A}$ of an input state

$$\mid \varphi_{out}\rangle = \mathbf{A} \mid \varphi_{in}\rangle \quad \text{or} \quad \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

**Example.** *One-qubit gates for the Pauli transformations.* The Pauli transformations are given by the set $\mathcal{G} = \{\sigma_I, \sigma_x, \sigma_y, \sigma_z\}$ are carried out by the quantum gates usually denoted as $I, X, Y$, and $Z$, respectively:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**Example.** *The Hadamard gate $H$.* It performs the transformation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and can be used to transform one basis to another

$$(\mid 0\rangle, \mid 1\rangle) \mapsto \left( \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}, \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}} \right) \quad \text{or} \quad \left( \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}, \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}} \right) \mapsto (\mid 0\rangle, \mid 1\rangle).$$

The quantum `CNOT` gate transforms an arbitrary initial state of two qubits expressed in the basis $(\mid 00\rangle, \mid 01\rangle, \mid 10\rangle, \mid 11\rangle)$ as follows:

$$\alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle + \alpha_2 \mid 10\rangle + \alpha_3 \mid 11\rangle \;\mapsto\; \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle + \alpha_3 \mid 10\rangle + \alpha_2 \mid 11\rangle.$$

In this expression $\alpha_0, \alpha_1, \alpha_2$ and $\alpha_3$ are arbitrary complex numbers satisfying the condition $\mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 + \mid \alpha_2 \mid^2 + \mid \alpha_3 \mid^2 = 1$; the matrix describing this unitary transformation is

$$U_{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

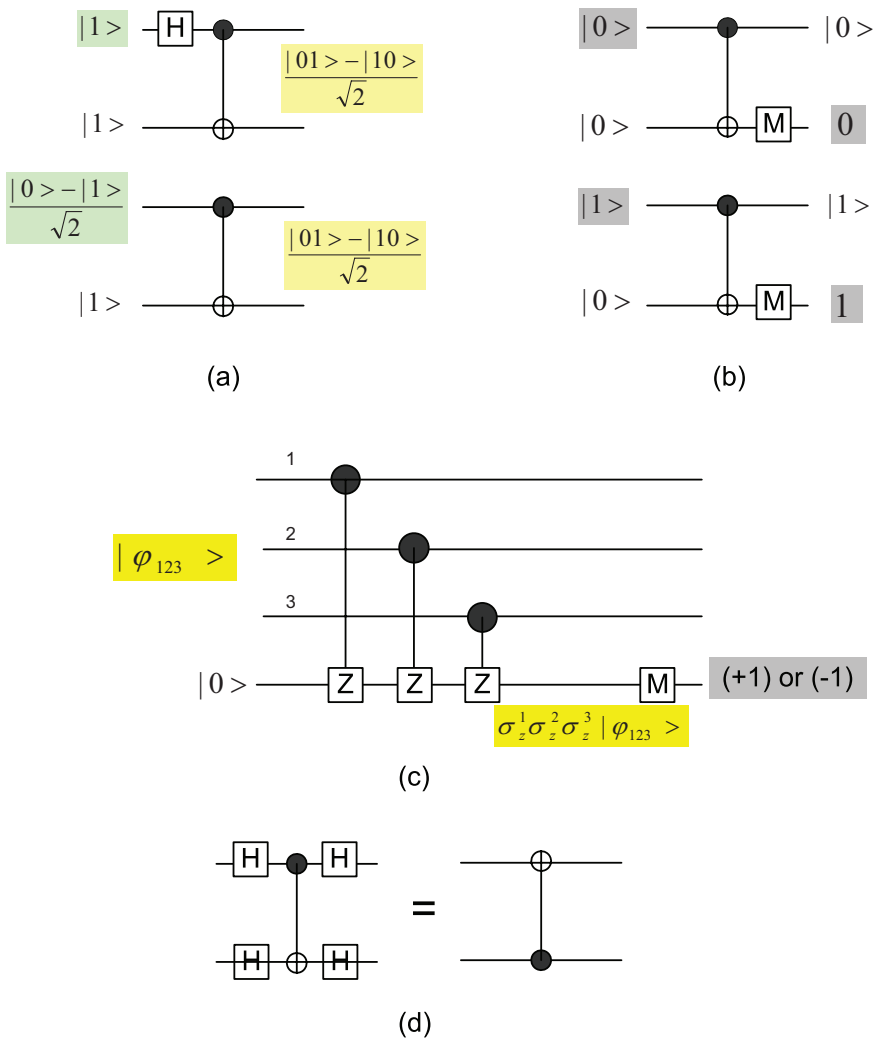



Figure 12: Properties of the quantum `CNOT` gate. (a) It produces perfectly entangled states from non-entangled states. When the control qubit is in state $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$ (this state can be produced by a Hadamard gate with input $|\ 1\rangle$) and the target qubit is in state $|\ 1\rangle$ then the output is in an EPR state, $\frac{|01\rangle-|10\rangle}{\sqrt{2}}$. (b) It can be used for a "non-demolition measurement" of the control qubit. (c) A more sophisticated non-demolition measurement $M$ of the operator $Z^1Z^2Z^3$ applied to the state $|\ \varphi_{123}\rangle$ of a system consisting of three particles, $1, 2$, and $3$. The measurement of the target qubits of the three `CNOT` gates produces the result in an ancilla qubit initially in state $|\ 0\rangle$. (d) When both the control and the target qubits are in a rotated basis, $\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}, \frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$ then the role of the control and target qubit are reversed.

If, for simplicity, we omit the time-ordered product, then the Hamiltonian of the transformation can be written as

$$U_{CNOT} = e^{\frac{i}{\hbar}\int H(t)dt}.$$

There is not a unique solution to this equation, there are many types of Hamiltonians which can be used to implement the CNOT gate [130].

The CNOT gate has special properties [115] some very useful for quantum error correction [130]. First, it can produce perfectly entangled states from non-entangled states, Figure 12(a). Second, it can be used for a "non-demolition measurement" of the control qubit if this qubit is either in state $\mid 0\rangle$ or $\mid 1\rangle$, but not in a superposition state; in Figure 12(b) we see that the target qubit ends up in the same state as the control qubit. Recall that the result of a measurement is an eigenvalue of the operator thus, the result of the measurement of the target qubit is 0 when the control qubit is in state $\mid 0\rangle$ and is 1 when the control qubit is in state $\mid 1\rangle$. Third, the non-demolition measurement $M$ of the operator $Z^1Z^2Z^3$ applied to the state $\mid \varphi_{123}\rangle$ of a quantum system consisting of three particles, $1, 2$, and 3 (control qubits of three CNOT gates) produces the result in an ancilla qubit initially in state $\mid 0\rangle$ as shown in Figure 12(c). Lastly, when both the control and the target qubits are in a rotated basis, $\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}, \frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$, then the role of the source and target qubit are reversed, Figure 12(d).

**Non-demolition measurements.** Such measurements are critical for quantum error correction because they leave the original state of a quantum system unchanged [92]. Circuits similar to the one in Figure 12(c) will be discussed in Section 5.14 but their remarkable properties are outlined now. A critical question for the error correction of a quantum code is the parity of a codeword. In our example we wish to determine the parity of a codeword consisting of the three qubits $1, 2$, and 3. The answer to this question is that the parity of qubit 1 is given by the eigenvalue of the transformation $Z_1$, applied to qubit 1; an eigenvalue of $+1$ corresponds to even parity and $-1$ to odd parity of the qubit. Consequently, the eigenvalue of the product $Z_1Z_2Z_3$ gives the parity of the three-bit codeword; $+1$ implies even parity and $-1$ odd parity. As noted by DiVincenzo [130], prior to the discovery of the properties of the circuit in Figure 12(c) it was thought that a measurement of a multi-particle Hermitian operator would require that each particle be measured separatively and could only be done in a demolishing manner [290] and that would have posed tremendous challenges to quantum error correction. These examples convince us that the CNOT gate plays a very important role in quantum information processing.

**Multi-qubit gates.** Toffoli gate is an example of a three-qubit gate; the unitary transformation performed by the Toffoli gate can be described using the basis states $\mid 000\rangle$, $\mid 001\rangle$, $\mid 010\rangle$, $\mid 011\rangle$, $\mid 100\rangle$, $\mid 101\rangle$, $\mid 110\rangle$, $\mid 111\rangle$ as

$$U_{Toffoli} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Multi-qubit gates require simultaneous access to the state of several qubits; intuitively, we expect that *the larger the number of qubits a quantum gates operates on, the more difficult is to implement that gate;* one-qubit gates are the easiest to implement, two qubit gates are

more complex, and so on. While it may be more expedient to express a transformation using multi-qubit gates such as Toffoli gate, it seems appropriate to ask the question if it is possible to implement multi-qubit gates with simpler, one- and two-qubit gates, a question we shall examine in more detail in the next section when we discuss universality of quantum gates.

As an example, consider the simulation of the three-qubit `controlled-U` gate of Deutsch, Figure 13, where $U$ is a generic unitary transformation and the transformation carried out by this gate is $U_D$

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \quad \text{and} \quad U_D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & u_{11} & u_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & u_{21} & u_{22} \end{pmatrix}.$$

This gate was studied extensively [118, 128, 275, 396] and it was proved that the quantum version of it can be decomposed into simpler parts while the classical one used for Boolean reversible computations cannot. The decomposition of the Deutsch gate as presented in Figure 13 uses two `CNOT` gates and three two-qubit `controlled-V` gates [396].

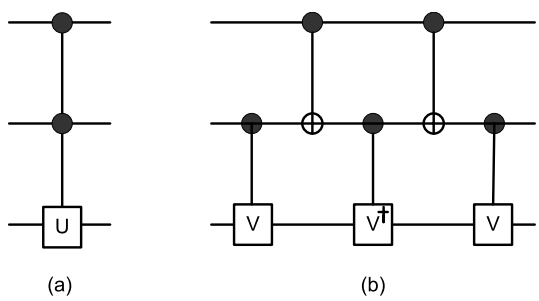



Figure 13: (a) The three-qubit `controlled-U` gate of Deutsch; the unitary transformation $U$ is carried out when both control qubits are set to $|\ 1\rangle$. (b) The decomposition of the three-qubit `controlled-U` using two `CNOT` gates and three two-qubit `controlled-V` gates given by [396] with $V^2 = U$.

**Quantum circuits.** Quantum circuits are collection of quantum gates interconnected by quantum wires. The actual structure of a quantum circuit, the number and the types of gates as well as the interconnection scheme are dictated by the unitary transformation **U** carried out by the circuit. Though in our description of quantum circuits we use the concepts of *input* and *output registers* of qubits we should be aware that physically, the input and the output of a quantum circuit are not separated, as their classical counterparts; this convention allows us to describe the effect of unitary transformation carried by the circuit in a more coherent fashion.

In all descriptions of quantum circuits in addition to gates we see quantum wires that move qubits and allow us to *compose* more complex circuits from simpler ones which, in turn,

are composed of quantum gates. We compose components by connecting the output of one to the input of another; we also compose operations when the results of an operation are used as input to another. The composition does not affect the quantum states. The quantum wires do not perform any transformations in a computational sense; sometimes we can view them as transformation carried out by the Pauli identity operator $\sigma_I$.

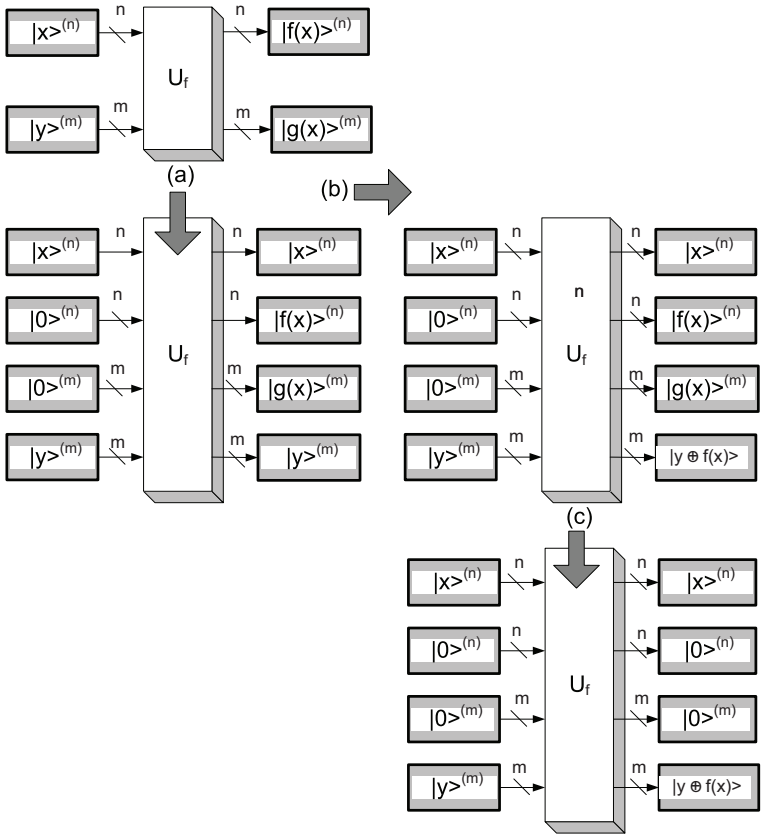



Figure 14: Schematic representation of a reversible quantum gate array; the circuit carries out the unitary transformation $\mathbf{U}_f$ and has two input registers one with $n$ qubits in state $\mid x\rangle^{(n)}$ and the other with $m$ qubits in state $\mid y\rangle^{(m)}$. There are two output registers; after the transformation one will be in state $\mid f(x)\rangle^{(n)}$ and the other in state $\mid g(x)\rangle^{(m)}$. (a) We add two new registers one for the result and the other for the ancilla qubits, both in state $\mid 0\rangle$. (b) We add CNOT gates for bitwise `AND` of $y$ and $f(x)$. (c) We add CNOT gates to reverse the computation and to set the second and third output registers to zero.

**Quantum computers.** Quantum computers perform unitary transformations, the input qubits in state $\mid x\rangle$ should be returned to the same state at the end of the computation. A quantum circuit requires a number of ancilla qubits to store partial results; the ancilla qubits are initially in a well-defined state, usually $\mid 0\rangle$ and must be returned to the same state at the end of the computation. Figure 14 shows a schematic representation of the transformations required by reversibility. The original circuit has two input registers one in state $\mid x\rangle^{(n)}$ and the other in state $\mid y\rangle^{(m)}$ as well as two output registers. As a result of the transformation one output register will be in state $\mid f(x)\rangle^{(n)}$ and the other in state $\mid g(x)\rangle^{(m)}$. First, we add

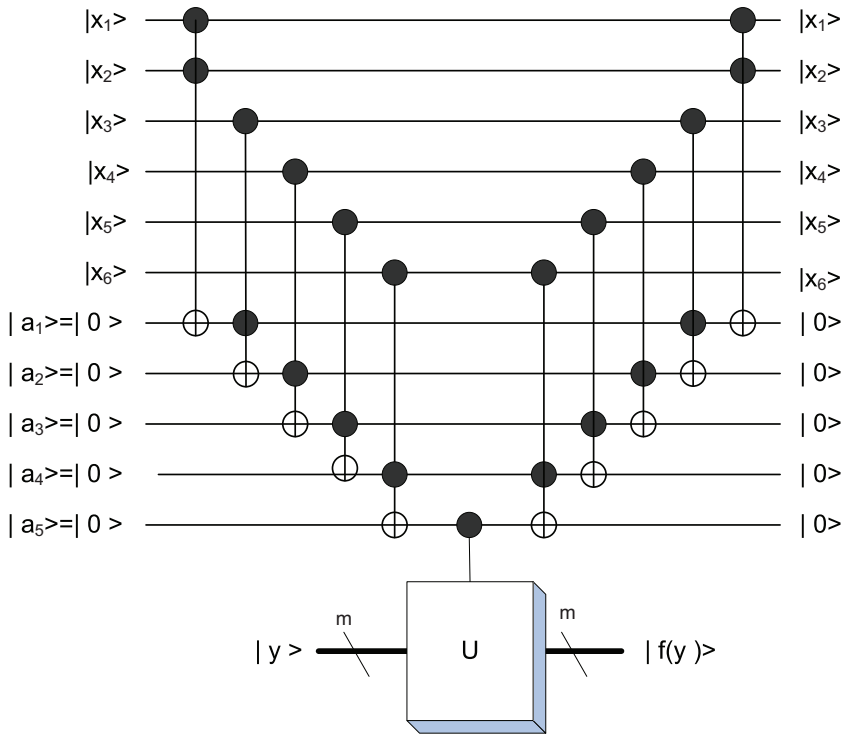



Figure 15: A reversible quantum circuit. The input register $x$ is in state $\mid x\rangle = \mid x_1x_2x_3x_4x_5x_6\rangle$; the register $y$ has $m$ qubits. The unitary transformation $\mathbf{U}$ is applied to $\mid y\rangle$ only if the condition $x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6 = 1$ is satisfied. There are five ancilla qubits used to store partial results; initially they are in state $\mid 0\rangle$ and after the transformation $\mathbf{U}$ they are returned to state $\mid 0\rangle$. The circuit uses ten `Toffoli` gates.

two new registers, one for the result and the other for the ancilla qubits, both in state $\mid 0\rangle$; then, we add CNOT gates for bitwise `AND` of $y$ and $f(x)$, and finally we add CNOT gates to reverse the computation and to return the result and ancilla registers to their initial state.

The quantum circuit in Figure 15 shows a circuit with an input register of six qubits in state $\mid x\rangle$. The unitary transformation $\mathbf{U}$ is applied to the register of qubits in state $\mid y\rangle^{(m)}$. The unitary transformation $\mathbf{U}$ is carried out on the register $\mid y\rangle$ if $x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6 = 1$. The five ancilla qubits used to store partial results are returned to their original state, state $\mid 0\rangle$, at the end of the computation.

In the next section we discuss one of the requirements formulated by DiVincenzo, the universality of quantum gates. The Solovay-Kitaev theorem states that it is possible to approximate with a desired level of accuracy the quantum circuit implementing a quantum algorithm, though no finite set of quantum gates can generate all unitary operations.

## 1.17 Universality of Quantum Gates; Solovay-Kitaev Theorem

Practical considerations related to fault-tolerance restrict the diversity of quantum gates used to express a quantum algorithm and lead us to the question of universality of quantum gates. In this section we discuss the theoretical aspects of the universality of quantum gates.

We start with the observation that one-qubit quantum gates cannot be universal; indeed, they cannot place two initially unentangled qubits in an entangled state. It is also clear that classical gates cannot be universal for quantum computing because they cannot create a superposition of quantum states.

We know from Section 1.15 that a gate operating on $k$ qubits is represented by a unitary transformation on a $2^k$-dimensional vector space. We denote by $SU(d)$ the special unitary group of degree $d$, the multiplicative group of $(d \times d)$ unitary matrices with determinant equal to 1. The $d$-dimensional vector space of $(d \times d)$ unitary matrices is a manifold and can be parameterized by a *continuum* of real parameters. For example, when $k = 1$ we have the set of one-qubit gates characterized by a unitary transformations $\mathbf{U}$ in $SU(2)$ parameterized by the set of continuous parameters $\alpha, \beta$ and $\theta$:

$$\mathbf{U}_2 = \begin{pmatrix} e^{i\alpha}\cos\theta & e^{i\beta}\sin\theta \\ -e^{-i\beta}\sin\theta & e^{-i\alpha}\cos\theta \end{pmatrix}.$$

No finite set of quantum gates can generate all unitary operators, thus, to define universality we should consider the *approximate* simulation of a circuit by another one. Cast in terms of quantum computability, rather than seeking computational universality, we wish to *approximate* any quantum algorithm, without using too many more gates than in its original description.

Barenco and coworkers showed that `CNOT` and one-qubit gates are sufficient to carry out universal quantum computations [16]; Lloyd established that almost any two-qubit gate belongs to a set of universal gates [275]. The general solution to the problem was given by Solovay in 1995 for $SU(2)$, but not published; in 1997 Kitaev published a review paper [240] and outlined a proof for the general case of $SU(d)$ and then Solovay announced that he has generalized the result for $SU(d)$ as well [112]. According to the Solovay-Kitaev (SK) Theorem, if a set $\mathcal{G}$ of one-qubit quantum gates generates a dense subset of $SU(2)$ then it is possible to construct good approximations of any gate using short sequences of gates from the set $\mathcal{G}$ of one-qubit gates; in other words, $\mathcal{G}$ is guaranteed to fill out $SU(2)$ quickly.

The problem of finding efficient approximations of quantum gates can also be cast as *quantum compiling* [203]. In classical computing a compiler transforms a source code to an object code consisting of instructions the target computer was designed to carry out; the translation should take a short time and the object code should run efficiently. By analogy, a quantum compiler expresses a quantum computation as a set of quantum gates, hopefully from a relatively small set. The problem is to define the set of gates, and to design efficient translation algorithms. In this context "efficient" means that the translation takes a short time and, even more importantly, that the resulting string of gates is of minimum length.

A set of gates $\mathcal{B} \subset SU(d)$ called *base gates* is said to be *computationally universal* if given any gate performing the unitary transformation $\mathbf{U} \in SU(d)$ we can find a string consisting of gates from the set $\mathcal{B}$ and their inverses, such that the product of these gates approximates $\mathbf{U}$ with arbitrary precision $\epsilon$; then $\mathcal{B}$ is said to be a dense subset of $SU(d)$.

An important question is the length of the string of basis gates used to approximate $\mathbf{U}$ with precision $\epsilon$. Lloyd's construction requires $\mathcal{O}(e^{1/\epsilon})$ gates [275]. The Solovay-Kitaev algorithm runs in a time polynomial in $\log(1/\epsilon)$ and produces strings of length $\mathcal{O}(\log^c(1/\epsilon))$ with $c$ a constant, $3 \leq c \leq 4$. The constat $c$ cannot be smaller than 1; indeed a ball of radius $\epsilon$ in $SU(d)$ has a volume proportional to $\epsilon^{d^2-1}$. To approximate every element of $SU(d)$ to a

precision $\epsilon$ we need $\mathcal{O}\left((1/e)^{d^2-1}\right)$ different strings of gates. In [203] it is shown that at least for some universal sets of base gates $c = 1$; in other words we need only $\mathcal{O}(\log(1/\epsilon))$ gates to approximate any gate to a precision $\epsilon$.

Given an arbitrary projector $\mathbf{P}$ and two states $\mid \varphi\rangle$ and $\mid \tilde{\varphi}\rangle$ such that $\| (\mid \varphi\rangle - \mid \tilde{\varphi}\rangle) \| \leq \epsilon$ then the probabilities of the two states satisfy the inequality

$$\| \mathbf{P} \mid \varphi\rangle \|^2 - \| \mathbf{P} \mid \tilde{\varphi}\rangle \|^2 = (\| \mathbf{P} \mid \varphi\rangle \| + \| \mathbf{P} \mid \tilde{\varphi}\rangle \|)(\| \mathbf{P} \mid \varphi\rangle \| - \| \mathbf{P} \mid \tilde{\varphi}\rangle \|) \leq 2\epsilon.$$

A similar inequality holds for mixed state with density matrices $\rho$ and $\tilde{\rho}$.

The $Solovay - Kitaev$ theorem guarantees that different sets of universal gates can simulate each other exceedingly fast and to a very good approximation:

**Theorem.** *$\mathcal{G} \subset SU(d)$ is a universal family of gates, where $SU(d)$ is a group of operator in a d-dimensional Hilbert space if:*

1. *$\mathcal{G}$ is closed under the inverse operation, $\forall g \in \mathcal{G} \Leftrightarrow g^{-1} \in \mathcal{G}$, and*
2. *$\mathcal{G}$ generates a dense subset of $SU(d)$*

*then there exists a finite sequence of gates from $\mathcal{G}$ such that*

$$\forall \mathbf{U} \in SU(d),\ \epsilon > 0,\ \ \exists g_1, g_2, \ldots, g_q \in \mathcal{G} : \| \mathbf{U} - \mathbf{U}_{g_1, g_2, \ldots, g_q} \| \leq \epsilon \quad \text{and} \quad q = \mathcal{O}(\log^2 1/\epsilon)$$

where $\| \mathbf{U} \ \|$ is the norm of the linear operator $\mathbf{U}$ and $\mathbf{U}_{g_1, g_2, \ldots, g_q}$ is an implementation of $\mathbf{U}$ using only $q$ gates from the set $\mathcal{G}$.

A proof for the general case $SU(d)$ is sketched in [240]. An algorithm for compiling an arbitrary one-qubit gate to a sequence of gates from a fixed and finite set is described in [112]. The algorithm is based on the SK Theorem and runs in $\mathcal{O}(\log^{2.71}(1/\epsilon))$ time and produces an output sequence of $\mathcal{O}(\log^{3.97}(1/\epsilon))$ quantum gates. The algorithm can be used to compile Shor's algorithm which uses rotations of $\pi/2^k$ using a target set consisting of Hadamard $H$ gates and $\pi/8$ one-qubit gates and `CNOT` gates.

An important class of quantum circuits are the stabilizer circuits used for quantum error correcting codes. Such circuits consist only of Hadamard, phase, and measurement one-qubit gates and `CNOT` gates. Unitary stabilizer circuits are also known as Clifford group circuits. It turns out that such circuits can be simulated efficiently using classical computers, a very important consideration for the investigation of the fault-tolerance of quantum circuits [1].

Recall that a quantum circuit maps $n$ input qubits to precisely $n$ output qubits, a necessary, but not a sufficient condition for reversibility. The transfer matrix $U$ of a quantum circuit is a unitary operator, $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$; when applied to an input register with $n$ qubits in state $\mid \psi^{(n)}\rangle$, the circuit produces $n$ qubits in state $\mid \varphi^{(n)}\rangle = U \mid \psi^{(n)}\rangle$. It is easy to show that any $n$-qubit quantum circuit $A$ can be simulated by another circuit $B$ constructed with only `CNOT` and `H,S` and `T` one-qubit gates which carry out the following transformations:

$$U_{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \; S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \; T = \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\pi/4) \end{pmatrix}.$$

with $S = T^2$ and $S^2 = Z$. The process of constructing the circuit $B$ consisting of only `CNOT` and `H,S` and `T` one-qubit gates that simulates the unitary transformation $A$ of $n$ qubits consists of three stages:

1. Decompose the unitary transformation $A$ as a product of unitary transformations $U_k, 1 \leq k \leq 2^n$ which affect at most two computational basis sates.

2. Rearrange the basis states so that each unitary transformation $U_k$ affects only one qubit thus, can be carried out by either a two-qubit `CNOT` gate or by one-qubit gates.

3. Approximate the transformation carried out by any one-qubit gate by the transformation carried out by the three gates in the set {`H, T, S`}.

## 1.18 Quantum Computational Models and Quantum Algorithms

A *quantum computational model* specifies the resources needed for a quantum computer, as well as the means to specify and control a quantum computation. To process quantum information, each computational model requires several steps from the set: free-time quantum evolution, controlled-time evolution, preparation, and measurement.

Several quantum computational models have been proposed: the quantum Turing Machine (QTM), quantum circuit, topological quantum computer, adiabatic, one-way quantum computer, and the measurements models, see Figure 16.

**The quantum Turing Machine computational model.** This model is a generalization of the classical Turing Machine model; it was the first quantum computational model introduced by Benioff in 1980 [27] and further developed by Bernstein and Vazirani in 1997 [51]. Though based on quantum kinematics and dynamics, Benioff's model was classical in the sense that it required specification of the state as a set of numbers measurable at any instant of time. Then, in 1982, Richard Feynman introduced a "universal quantum simulator" consisting of a lattice of spin systems with near-neighbor interactions; Feynman's model could simulate any physical system with a finite-dimensional state space but did not include a mechanism to select arbitrary dynamic laws.

**The quantum circuit model.** The standard model for a quantum computer is related to the classical circuit model. It was proposed by David Deutsch [116] and further developed by Andrew Yao [465]; this model applies only to uniform families of quantum circuits.

To answer the question how close is the quantum circuit model to the classical one we should remember the procedure discussed in Section 1.15 for constructing the Boolean circuit implementing a function $f : \{0,1\}^n \mapsto \{0,1\}$ and the five elements required for the

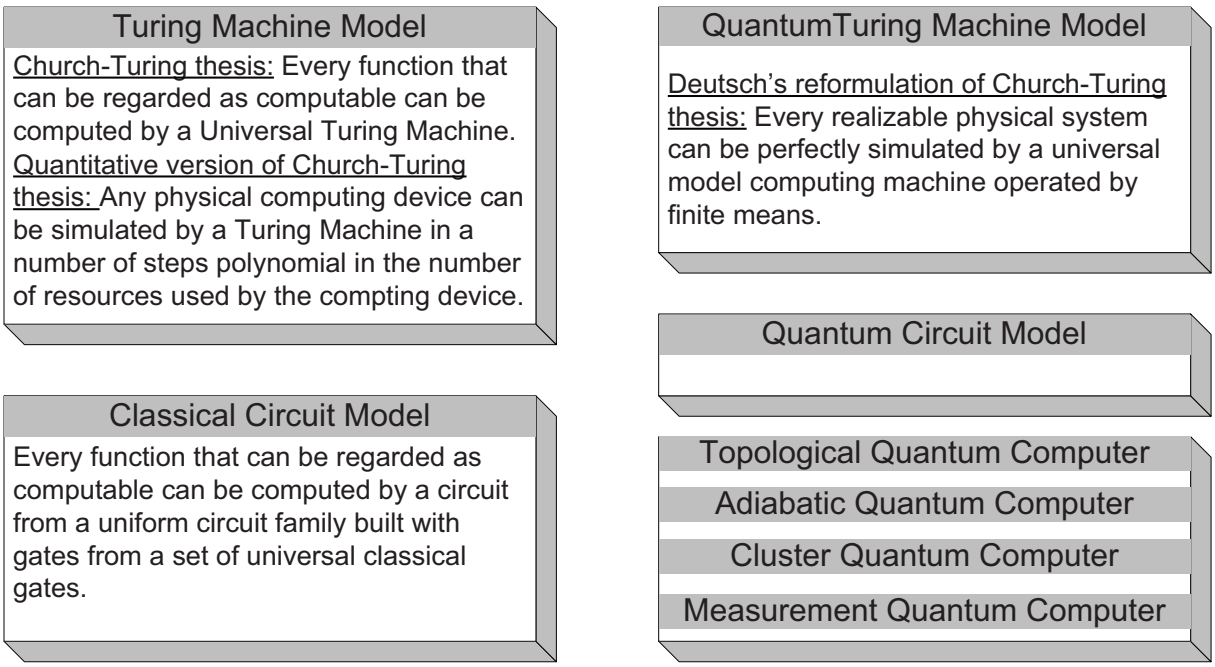



Figure 16: Classical and quantum computational models.

construction of the circuit. Some of the challenges posed by the quantum circuit model are: the no-cloning theorem prohibits fanout; the decoherence of quantum states makes even the implementation of quantum wires non-trivial; we can only approximate the function of a one-qubit gates with gates from a small set of universal quantum gates.

To capture the finiteness of resources required by a physical realization of a computing device David Deutsch restated the Church-Turing hypothesis [116] as: "every realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means." The much stronger formulation of the Church-Turing hypothesis is not satisfied by a Turing Machine $\mathcal{T}$ operating in the realm of classical Physics; indeed, the set of states of a classical physical system form a continuum due to the continuity of classical dynamics. The classical Turing Machine $\mathcal{T}$ cannot simulate every classical dynamics system because there are only countable ways of preparing the input for $\mathcal{T}$.

In the quantum circuit model *a quantum computation is carried out by quantum circuits that transform information under the control of external stimuli.* A quantum circuit operating on $n$ qubits performs a unitary operation in the Hilbert space $\mathcal{H}_{2^n}$ and consists of a finite collection of quantum gates; each quantum gate implements a unitary transformation on a small number $k$ of qubits, Figure 9 (b). In the quantum circuit model, a computation involving $n$ qubits consists of three stages:

1. Initialization stage, when the quantum computer is prepared in a basis state in $\mathcal{H}_{2^n}$;

2. Processing stage, when a sequence of one- and two-qubit gates are applied to the qubits;

3. Readout stage, when a measurement of a subset of qubits in the computational basis allows us to obtain the classical information revealing the result of the computation.

The quantum circuit model involves an *open* quantum system; the role of the control unit of the von Neumann architecture is played by a classical system that controls the temperature, the pressure, the magnetic filed, the electric field, or other parameters specific to the physical realization of the quantum computing device.

A quantum computing engine $\mathcal{Q}$ consists of two components: a finite processor consisting of $n$ two-state observables (qubits) $p_i \in \mathcal{H}_{2^n}$ and an infinite memory consisting of an infinite sequence of two-state observables (qubits) $m_i$. The computation consists of steps of finite duration, $t_s$; during each step the processor and a finite segment of the memory interact. The state of $\mathcal{Q}$ is a unit vector in a Hilbert space spanned by simultaneous eigenvectors $\mid b_i\rangle$ forming computational basis states. The computation begins at $t = 0$ and we need to specify the state $\mid \psi\rangle$ only at times that are multiples of $t_s$. The system dynamics are specified by a unitary operator $\mathbf{U}$:

$$\mid \psi(kt_s)\rangle = \mathbf{U}^k \mid \psi(0)\rangle.$$

While a classical Turing Machine signals the end of a computation when two consecutive states are identical, two states of a reversible computing engine $\mathcal{Q}$ can never be identical. Thus, the processor of the quantum computing engine $\mathcal{Q}$ must have an internal qubit $p_e$ initially in the state $\mid 0\rangle$ and set to the state $\mid 1\rangle$ when the computation has finished. This qubit must be periodically observed from the outside without affecting the state of $\mathcal{Q}$; this is possible because this qubit contains only classical information.

We say that a quantum "program" is valid if the state of $p_e$ goes to $\mid 1\rangle$ in a finite time; this statement is analogous to the one for a classical Turing Machine $\mathcal{T}$ in which case a"program" is valid if the machine halts after a finite number of steps. We now realize that a classical Turing Machine is similar to a quantum computing engine $\mathcal{Q}$ whose evolution in time ensures that at the end of each computational step it remains in a computational basis state provided that it started in one.

In summary, according to the circuit model a quantum "program" is a sequence of quantum gates applied to a group of qubits at a time. The function of a multi-qubit gate can be simulated with one-qubit and two-qubit gates and multi-qubit gates are more difficult to realize physically.

**The topological quantum computer model.** In 1997 Alexei Kitaev [241] proposed anyonic quantum computing as an inherently fault-tolerant quantum computing model. Kitaev showed that a system of non-abelian anyons can simulate efficiently a quantum circuit; in Kitaev's scheme measurements are necessary to simulate some quantum gates. Preskill elaborated on the idea of topological quantum computing showing that it provides elegant means to ensure fault-tolerance [342]. Freedman, Larsen, and Wang showed that the measurements can be postponed until the final readout of the results [158]; then they proved that a system of anyons can be simulated effectively by a quantum circuit thus, demonstrating that the topological quantum computing model has a similar computing power as the quantum circuit model [159]. Mochon gave a constructive proof that anyonic magnetic charges with fluxes in a non-solvable finite group can perform universal quantum computations; the gates are based on elementary operations of braiding, fusion, and vacuum pair creation [296].

Recall from Section 1.13 that a Fractional Quantum Hall Effect, FQHE, occurs when a two-dimensional electron gas placed in a strong magnetic field, at very low temperature, behaves

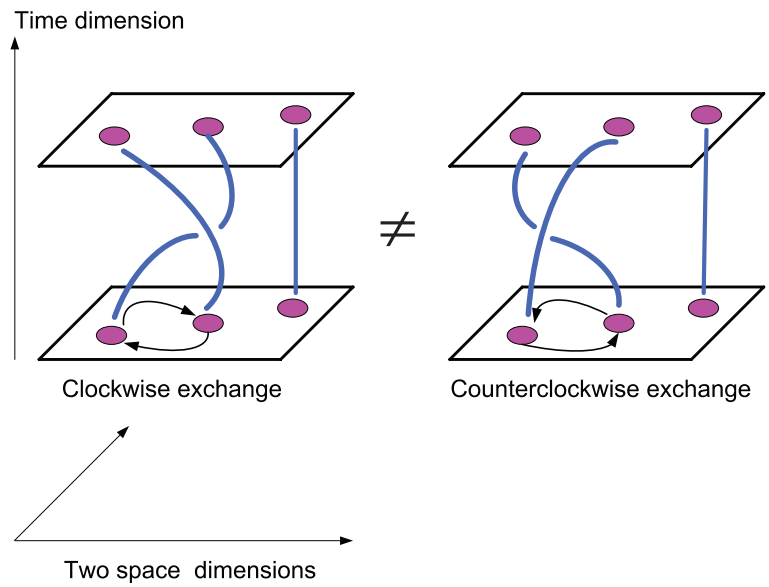



Figure 17: Particle world-lines form braids in a space with $(2+1)$ dimensions. (Left) The braids corresponding to the clockwise exchange of two particles. $\mid \varphi_{init}\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ is the initial state and $\mid \varphi_{final}\rangle = \gamma_0 \mid 0\rangle + \gamma_1 \mid 1\rangle$ is the final state of the system. The symmetry group generated by the exchanges of 2 particles is in an one-to-one correspondence with the braid group $B_2$ of $2 \times 2$ unitary matrices. (Right) The braids corresponding to the counterclockwise exchange of the two particles.

as a system of anyons, particles with a fractional charge, e.g., $e/3$ where $e$ is the electric charge of an electron. In a nut shell, *a topological quantum computer braids world-lines by swapping the positions of anyons.* This terse statement needs a fair amount of explaining. A *braid*, or a strand, connects two objects; informally, world-lines are representations of particles as they move through time and space. Particle world-lines form braids in a space with $(2+1)$ dimension (two space- and one time-dimension), Figure 17. The symmetry group generated by the exchanges of $n$ particles is in a one-to-one correspondence with the braid group.

*The braid group.* $B_n$, the braid group on $n$ strands, is an infinite group with an intuitive geometric representation, it shows how $n$ strands can be laid out to connect two groups of $n$ objects, Figure 18. The braid group is a multiplicative group; the product of two elements is constructed geometrically by laying the two elements next to one another in the order dictated by the product and connecting the two objects in the middle. To construct the inverse of an element, the left hand set becomes the right hand set and viceversa. The identity element connects objects with the same index from the left and right sets and the strands are straight lines. Matrices form a non-Abelian representation of the braid group.

The symmetry group generated by the exchanges of any 2 particles are in an one-to-one correspondence with the braid group $B_2$ of $2 \times 2$ unitary matrices. If $\mid \varphi_{init}\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ denotes the initial state and $\mid \varphi_{final}\rangle = \gamma_0 \mid 0\rangle + \gamma_1 \mid 1\rangle$ denotes the final state of the system in Figure 17 (Left) then two states are related by the relation:

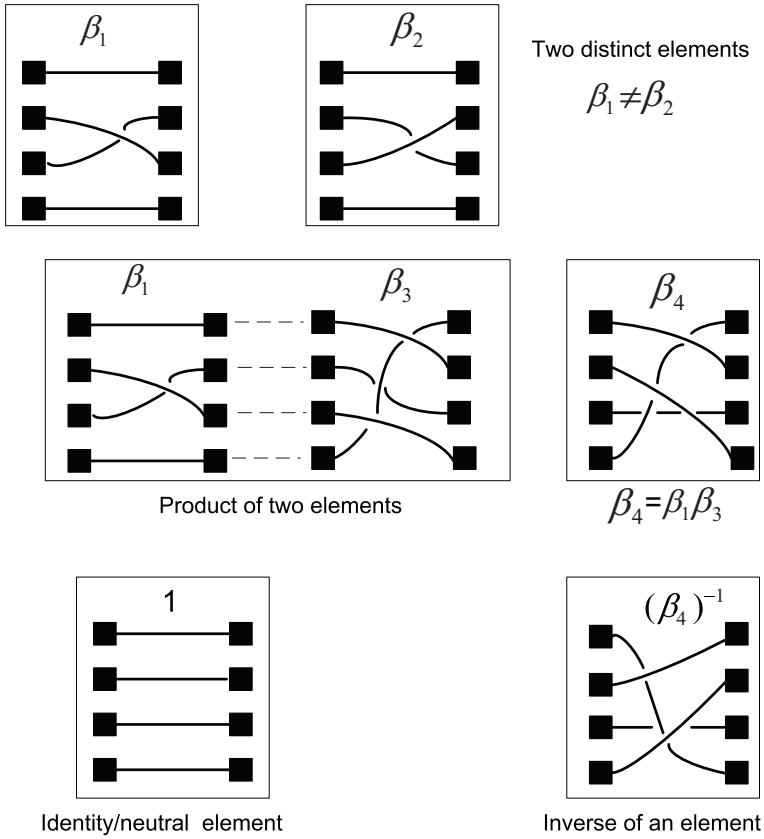



Figure 18: The braid group $B_4$. $\beta_1, \beta_2, \beta_3, \beta_4$ are elements of $B_4$; each element consists of two sets of 4 objects (in our case squares), one on the left and one on the right, connected by 4 strands. Each set of 4 squares is arranged vertically; a strand connects one object from the left set with an object from the right set. A strand moves from left to right and knots are not allowed; for example the four strands of $\beta_1$ connect the objects: $(1, \text{left}) \rightarrow (1, \text{right})$, $(2, \text{left}) \nearrow (3, \text{right})$,$(3, \text{left}) \searrow (2, \text{right})$, and $(4, \text{left}) \rightarrow (4, \text{right})$. (Top) $\beta_1 \neq \beta_2$ because the positions of the strands connecting objects 2 and 3 are different, $(2, \text{left}) \nearrow (3, \text{right})$ is above in $\beta_1$ and $(2, \text{left}) \searrow (3, \text{right})$ is below in $\beta_2$. (Middle) The braid group is a multiplicative group; the product of two elements is constructed by laying the two elements next to one another in the order dictated by the product and connecting the two objects in the middle. (Bottom) The construction of the inverse of an element: the left set becomes the right set and viceversa. The identity element connects objects with the same index from the left and right sets and the strands are straight lines; it is easy to see that $(\beta_4)(\beta_4)^{-1} = 1$.

$$\begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \quad \text{or} \quad \mid \varphi_{final} \rangle = A \mid \varphi_{init} \rangle, \quad \text{with} \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

This relation resembles the one describing the transformation carried out by a one-qubit quantum gate and hints to the fact that the topological quantum computer model and the quantum circuit model are closely related; in fact they are equivalent [159].

**The adiabatic quantum computing model.** The adiabatic quantum computing model was proposed in 2000 by Farhi *et al.* [146] who suggested an algorithm to solve optimization problems such as SATISFIABILITY (SAT); there is now evidence that this algorithm takes an exponential time for some NP-complete problems. The interest in the adiabatic quantum computing was renewed in 2005 when Aharonov *et al.* proved that it is equivalent to the quantum circuit model [3].

An adiabatic process is a *quasi-static* thermodynamic process when no heat is transferred; the opposite of an adiabatic process is an isothermal process when heat is transferred to maintain the temperature constant. An *adiabatic evolution of a quantum system means that the Hamiltonian is slowly varying;* recall from Section 1.6 that the Hamiltonian operator corresponds to the total energy of the quantum system. The Hamiltonian is a Hermitian operator and its eigenvector corresponding to the smallest eigenvalue, i.e., to the lowest total energy of the system, is called the ground state of the system. A *local* Hamiltonian describes a quantum system where the interactions occur only among a constant, and rather small, number of particles.

The adiabatic approximation is a standard method to derive approximate solutions to the Schrödinger equation when the Hamiltonian is slowly varying. This method is based on a simple idea: if the quantum system is prepared in a ground state and the Hamiltonian varies slowly enough, then, as the time goes on, the system will stay in a state close to the ground state of the instantaneous Hamiltonian. This idea is captured by the Adiabatic Theorem of Born and Fock [66]: "A physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough, and if there is a gap between the eigenvalues (corresponding to this eigenstate) and the rest of the Hamiltonian's spectrum."

Consider a quantum system in state $\mid \psi(t)\rangle \in \mathcal{H}_n$ with a Hamiltonian $\mathbf{H}(t)$. The evolution of the system is described by Schrödinger equation

$$i\frac{d}{dt} \mid \psi(t)\rangle = \mathbf{H}(t) \mid \psi(t)\rangle.$$

Assume that the Hamiltonian $\mathbf{H}(t)$ is slowly varying

$$\mathbf{H}(t) = \tilde{\mathbf{H}}(t/T)$$

where $T$ controls the rate of variation of $\mathbf{H}(t)$ and $\tilde{\mathbf{H}}(t/T)$ belongs to a smooth one-parameter family of Hamiltonians $\tilde{\mathbf{H}}(s), 0 \leq s \leq 1$. The instantaneous eigenstates $\mid i; s\rangle$ and eigenvalues $E_i$ of $\tilde{\mathbf{H}}(s)$ are defined as

$$\mathbf{H}(s) \mid i; s\rangle = E_i \mid i; s\rangle.$$

The eigenvalues of $\tilde{\mathbf{H}}(s)$ are ordered

$$E_0(s) \leq E_1(s) \leq \ldots E_n(s).$$

Call $\mid \psi_0\rangle = \mid i = 0; s = 0\rangle$ the ground state of $\tilde{\mathbf{H}}(0)$. The Adiabatic Theorem says that if the gap between the lowest energy levels $E_1(s) - E_0(s) > 0,\ \forall\ 0 \leq s \leq 1$, then the state, $\mid \psi(t)\rangle$, of the system after an evolution described by the Schrödinger equation, will be very close to the ground state of the Hamiltonian $\mathbf{H}(t)$ for $0 \leq t \leq T$ when $T$ is large enough.

The adiabatic quantum computer evolves between an initial state with the Hamiltonian $\mathbf{H}_{init}$ and a final state with the Hamiltonian $\mathbf{H}_{final}$. The input data and the algorithm are encoded as the ground state of $\mathbf{H}_{init}$ and the result of the computation is the ground state of $\mathbf{H}_{final}$. The running time of the adiabatic computation is determined by the minimal spectral gap of all Hamiltonians of the form

$$\mathbf{H}(s) = (1-s)\mathbf{H}_{init} + s\mathbf{H}_{final}, \quad 0 \leq s \leq 1.$$

These Hamiltonians lie on the straight line connecting $\mathbf{H}_{init}$ and $\mathbf{H}_{final}$ [3]. The ground state of the Hamiltonian $\mathbf{H}_{final}$ for the optimization algorithm in [146] was a classical state in the computational basis and $\mathbf{H}_{final}$ was a diagonal matrix as the solution of a combinatorial optimization problem. This restriction was removed by Aharonov *et al.* [3] who require only that the Hamiltonians be *local*. This condition resembles the one imposed on the quantum circuit model, namely, that the quantum gates operate on a constant number of qubits.

**The one-way quantum computer model.** In 2001, Raussendorf and Briegel proposed a model based on a special type of entangled states, cluster states [350]. A cluster state refers to a family of quantum states of $n$-qubit two- or three-dimensional lattice, or even a more general graph, in which each vertex corresponds to a qubit. A specific preparation procedure is applied, e.g., vertices are connected using controlled-Z two-qubit gates.

To process quantum information in this network it is sufficient to measure the qubits in certain order, in a certain basis. In a two-dimensional lattice the quantum information propagates horizontally by measuring the qubits on the wire, while qubits on the vertical connection implement two-qubit gates. The bases used for measurement depend on the results of previous measurements. Any quantum circuit can be implemented on a cluster state [350].

**The quantum measurement model.** In this model introduced by Nielsen [309] no coherent dynamical operations are allowed; the model allows only three operations: (i) Preparation of qubits in the initial state, $\mid 0\rangle$; (ii) Storage of qubits; and (iii) Projective measurements (see Section 2.7) of up to four qubits at a time in arbitrary bases.

**Quantum algorithms** Once we are convinced that given a function $f$ we are able to assemble together a quantum circuit capable of evaluating this function we switch our attention to quantum algorithms. Recall that a computational problem is considered tractable if an algorithm to solve it in a number of steps and requiring storage space polynomial in the size of the input exists. There are classically intractable problems, such as the Travelling Salesman Problem which are proven to belong to the complexity class *non-deterministic polynomial* ($\mathcal{NP}$). The expectation that quantum algorithms could lead to efficient solution of "hard" computational problems proved to be justified.

A number of "toy problems" such as Deutsch problem, provided a first glimpse of hope. Then, in 1994, Peter Shor found a polynomial time algorithm for the factorization of $n$-bit numbers on quantum computers [380]. His discovery generated a wave of enthusiasm for quantum computing, for two major reasons: the intrinsic intellectual beauty of the algorithm and the fact that efficient integer factorization has important applications.

The security of widely used cryptographic protocols is based on the conjectured difficulty of the factorization of large integers. Like most factorization algorithms, Shor's algorithm reduces the factorization problem to the problem of finding the period of a function, but

uses quantum parallelism to find a superposition of all values of the function in one step. Then the algorithm calculates the quantum Fourier transform of the function, which sets the amplitudes to multiples of the fundamental frequency, the reciprocal of the period. To factor an integer, Shor's algorithm measures the period of the function[9].

In 1996, Grover described a quantum algorithm for searching an unsorted database $\mathcal{D}$ containing $N$ items in a time of order $\sqrt{N}$ [183]; on a classical computer the search requires a time of order $N$.

It seems natural to start our analysis of quantum algorithms with the very first algorithms developed to reveal the power of quantum parallelism. These algorithms for "toy" problems proposed by Deutsch, Jozsa, Bernstein and Vazirani, and Simon will be discussed before an in depth analysis of phase estimation and of Grover search algorithm.

The algorithms discussed in the next section are constructed around black boxes and measurements performed on the output of these black boxes. The algorithms of Shor and Grover use white boxes implementing the quantum Fourier transform (QFT) or the quantum Hadamard transform and determine the period of a function, or the angle of the rotation of a state vector.

## 1.19 Deutsch, Deutsch-Jozsa, Bernstein-Vazirani, and Simon Oracles

In this section we show that one can construct quantum circuits capable of providing simple answers, in an *efficient* manner, to questions that require very elaborate computations, very much like oracles do[10]. Classical solutions to such problems require either an exponential number of time steps, or an exponential number of copies of a circuit able to compute the value of a function. The term "efficient" means that the quantum solution requires a single copy of the quantum circuit and either a single, or a linear number of time steps.

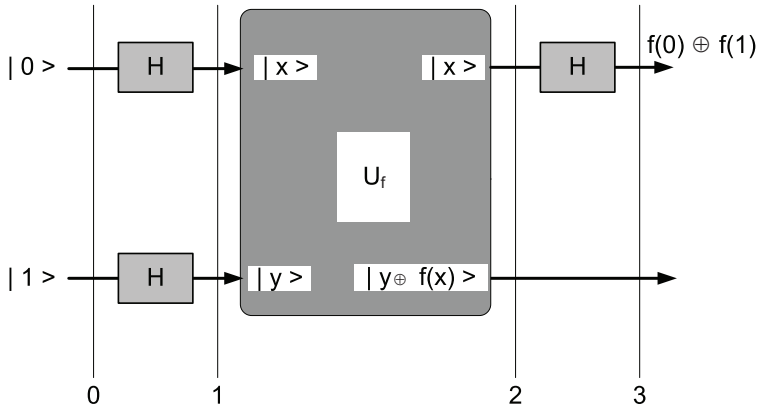



Figure 19: A quantum circuit for solving Deutsch problem.

[9]A powerful version of the technique used by Shor is the *phase-estimation algorithm* of Kitaev [239]

[10]The name "oracle" comes form the Latin verb "orare" meaning "to speak;" an oracle is a source of wise pronouncements or prophetic opinions with some connection to an infallible spiritual authority.

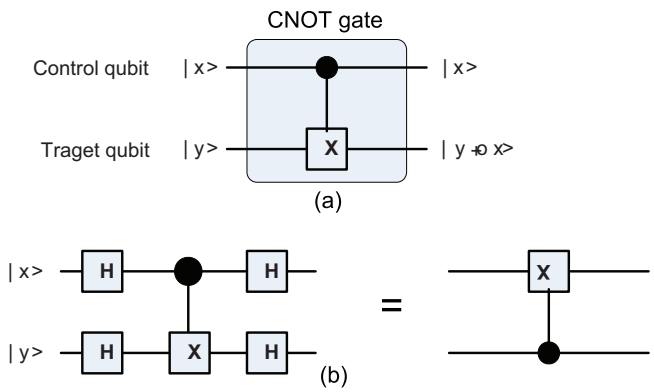



Figure 20: (a) A `CNOT` gate. (b) `Hadamard` gates on the input and output lines of a `CNOT` gate allow us to swap control and target qubits.

In the realm of computing, an oracle is an abstraction for a black box that can follow a very elaborate procedure to answer a very complex question with a "yes" or "no" answer. Quantum computing has a natural affinity for oracles because the result of a quantum computation is probabilistic and it is a superposition of all possible results.

Several increasingly more complex oracles used in quantum computing are:

- *Deutsch oracle.* Given a function $f : \{0,1\} \mapsto \{0,1\}$ the oracle decides if the function is constant, $f(0) = f(1)$, or balanced, $f(0) \neq f(1)$.

- *Deutsch-Jozsa oracle.* Given a function $f : \{0,1\}^n \mapsto \{0,1\}$ the oracle decides if the function is constant or balanced. The function is balanced if for half of the arguments it is equal to 0 and for the other half is equal to 1. If the function is neither constant, nor balanced, the oracles answer is meaningless.

- *Bernstein-Vazirani oracle.* Given a function $f : \{0,1\}^n \mapsto \{0,1\}$ of the form $f(x) = a \cdot x$, where $a$ is a constant vector of 0s and 1s and $a \cdot x$ is the scalar product of the vectors $a$ and $x$, the oracle determines the value of $a$ in one time step.

- *Simon oracle.* Given a function $f : \{0,1\}^n \mapsto \{0,1\}^{n-1}$ which is $2 \mapsto 1$ and periodic with period $a$, the oracle returns the period in $\mathcal{O}(n)$ measurements.

To illustrate the feasibility of quantum oracles we present two quantum circuits, one implementing the oracle for the Deutsch problem and the other for the Bernstein-Vazirani problem. Let us first discuss briefly Deutsch oracle. Recall that we could use two copies of a classical circuit to compute in a single time step $f(0)$ and $f(1)$ and decide if the function $f$ is balanced or constant, assuming that we can ignore the time to compare $f(0)$ and $f(1)$. Alternatively, we can use a single copy of the circuit built to compute a binary function $f$ of a binary argument $x$, $f(x)$, but in this case we need two time steps, one to compute $f(0)$ and the other to compute $f(1)$ .

The quantum circuit in Figure 19, see [284], shows the quantum circuit for Deutsch problem. It is easy to see that the quantum circuit produces $f(0) \oplus f(1)$. We calculate the state of the system after each stage, $\mid \xi_i \rangle$, $1 \leq i \leq 3$ and we conclude that

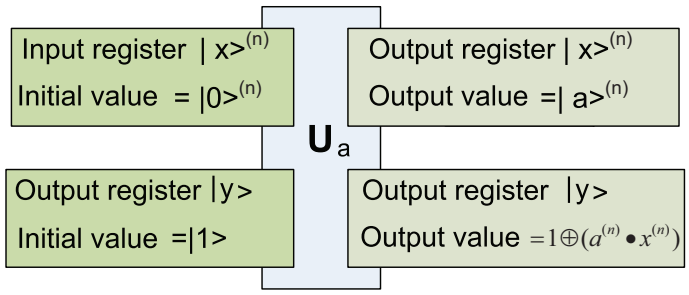



Figure 21: The black box of the Bernstein-Vazirani oracle performs a unitary transformation, $U_a$. The expression $f(x) = a \cdot x$ contains an $n$-bit vector hardwired into the circuit. The oracle has as input a register of $n$ qubits, $\mid x\rangle^{(n)}$, initially in state $\mid 0\rangle^{(n)}$ and a register $\mid y\rangle$ of one qubit, initially set to state $\mid 1\rangle$. The oracle returns in the output register $\mid x\rangle^{(n)}$ the $n$ qubit vector $\mid a\rangle^{(n)}$ and in the output register $\mid y\rangle$ the expression $\mid 1 \oplus (a^{(n)} \cdot x^{(n)})\rangle = \mid 1 \oplus \sum_{i=1}^{n} a_i x_i\rangle$.

$$\mid \xi_3\rangle = \pm \mid f(0) \oplus f(1)\rangle \left[\frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}\right].$$

This expression tells us that by measuring the first output qubit of the circuit in Figure 19 we are able to determine $f(0) \oplus f(1)$ after performing a single evaluation of the function. If the outcome of a measurement of the first qubit is equal to 0 we conclude that the function is constant and if it is 1 we conclude that the function is balanced.

It is relatively easy to construct a circuit for the more general oracle capable of solving Deutsch-Jozsa problem. We use as a model the circuit in Figure 19 and the input is a register $x^{(n)}$ of $n$ qubits.

Before discussing the Bernstein-Vazirani oracle we review the effect of adding two `Hadamard` gates to the control and target qubits of a `CNOT` gate, Figure 20; when we add two `Hadamard` gates on each input and output line of a `CNOT` gate we swap the control and the target qubits. A detailed analysis of this circuits can be found in [284].

Now we turn our attention to the Bernstein-Vazirani problem. The quantum circuit in Figure 21 computes the binary value of the scalar product

$$z = (a \cdot x) = \sum_{i=1}^{n} a_i x_i$$

of two $n$-dimensional binary vectors:

$$x = (x_n x_{n-1} \ldots x_1), \quad x_i \in \{0, 1\}, \ 1 \leq i \leq n$$

and

$$a = (a_n a_{n-1} \ldots a_1), \quad a_i \in \{0, 1\}, \ 1 \leq i \leq n.$$

Thus,

$$a = a_n \times 2^{n-1} + a_{n-1} \times 2^{n-2} + \ldots a_2 \times 2^1 + a_1 \times 2^0.$$

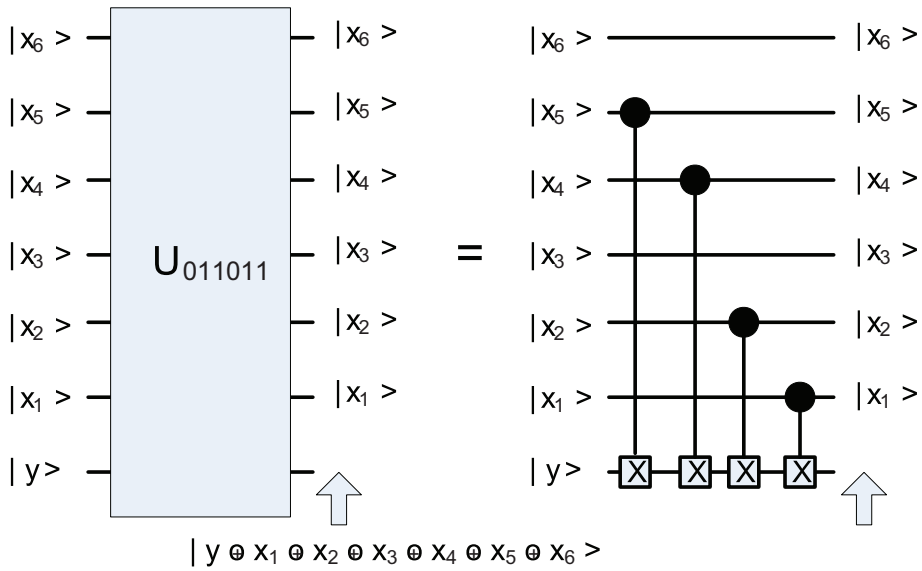



Figure 22: The Bernstein-Vazirani oracle for $n = 6$ and $a = 27$. (Left): the black box. (Right): the actual circuit with 4 `CNOT` gates. The output qubit $\mid y\rangle$ is initially in state $\mid 1\rangle$ and becomes $\mid 1 \oplus a_6x_6 \oplus a_5x_5 \oplus a_4x_4 \oplus a_3x_3 \oplus a_2x_2 \oplus a_1x_1\rangle$. Here $\oplus$ is addition modulo 2. In binary $a = a_6 \times 2^5 + a_5 \times 2^4 + a_4 \times 2^3 + a_3 \times 2^2 + a_2 \times 2^1 + a_1 \times 2^0$, or $a = 011011$.

Initially, the control qubits $\mid x\rangle^{(n)}$ are in state $\mid 0\rangle^{(n)}$ and the target qubit $\mid y\rangle$ is in state $\mid 1\rangle$. As a result of the computation, the target qubit is the binary complement, $z$, of the inner product $(a \cdot x) = \sum_{i=1}^{n} a_i x_i$. If the inner product has an even number of terms equal to 1 then $z = 0$; if the inner product has an odd number of terms equal to 1 then $z = 1$. The reason for choosing the initial state of the target qubit to be $\mid 1\rangle$ and obtain the complement of the inner product $(a \cdot x)$ will become apparent later when we examine the actual transformation carried out by the circuit. A classical solution to the problem of determining the constant vector $a = (a_1, a_2, \ldots, a_n)$ requires $n$ computations; a quantum circuit implementing the oracle requires only one copy of the circuit and one time step.

**Example.** *A Bernstein-Vazirani oracle for $n = 6$ and $a = (a_6a_5a_4a_3a_2a_1) = 011011$.*

$$a = 0 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 27.$$

The oracle and the inner working of circuits for this example are shown in Figures 22 and 23, respectively. The circuit on top of Figure 23 has two `Hadamard` gates on each control qubit and two `Hadamard` gates on the target line; this circuit performs the same transformation as the one in Figure 22, because the Hadamard transformation is unitary, $\mathbf{HH}^\dagger = \mathbf{HH} = \mathbf{I}$. The circuit in the middle of Figure 23 has two `Hadamard` gates on each control and target line of every `CNOT`; Figure 20 shows that in this case we swap control and target qubits. The circuit at the bottom of Figure 23 reflects the reversal of the role of control and target qubits; when the input register is $\mid 000000\rangle$ the output register will be $\mid 011011\rangle$, the binary expression of constant $a = 27$; the control qubit will be equal to $\mid 1\rangle$. We want the input target qubit to be equal to 1 to allow the flipping of the qubits corresponding to a 1 in the binary expression

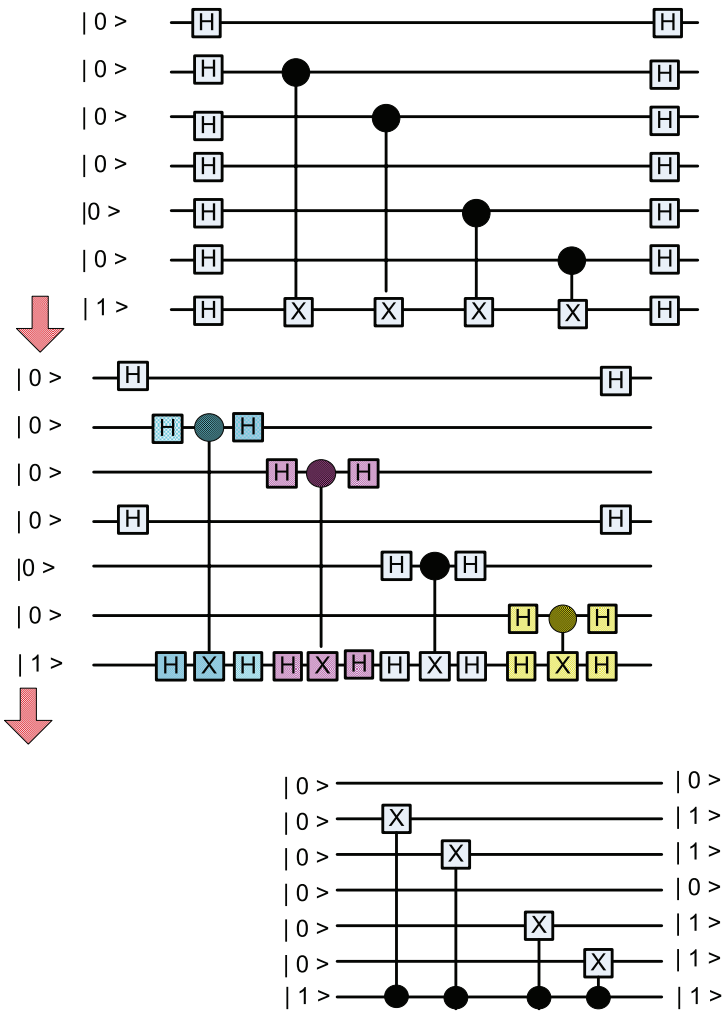



Figure 23: The inner working of the Bernstein-Vazirani oracle in Figure 22. Top: we add two `Hadamard` gates on each line without altering the function. Middle: we add two `Hadamard` gates on each control and target line of every `CNOT` and thus, we swap the control and target qubit of every `CNOT` gate. Bottom: The target and control qubits of each `CNOT` gate are swapped. Now all `CNOT` gates share a common control qubit (the original target qubit) initially in state $|\,1\rangle$. Thus, the output 6-bit register will contain the binary expression of $a = 27$, the vector $|\,011011\rangle$.

of the constant vector $a$.

Given a black box which carries out the transformation $f_a : \{0,1\}^n \mapsto \{0,1\}^{n-1}$ the problem addressed by Simon oracle is to determine the period $a$ with a minimum number of evaluations of function $f_a$. The function $f_a$ is said to be periodic for bitwise modulo 2 addition, an operation denoted as $\oplus$, if there exists an integer $a \neq 0$ such that

$$f_a(x) = f_a(x \oplus a), \quad \forall\ 0 \leq x \leq 2^n - 1.$$

Then $a$ is called the period of $f_a$. If the function is periodic with period $a$, then given two

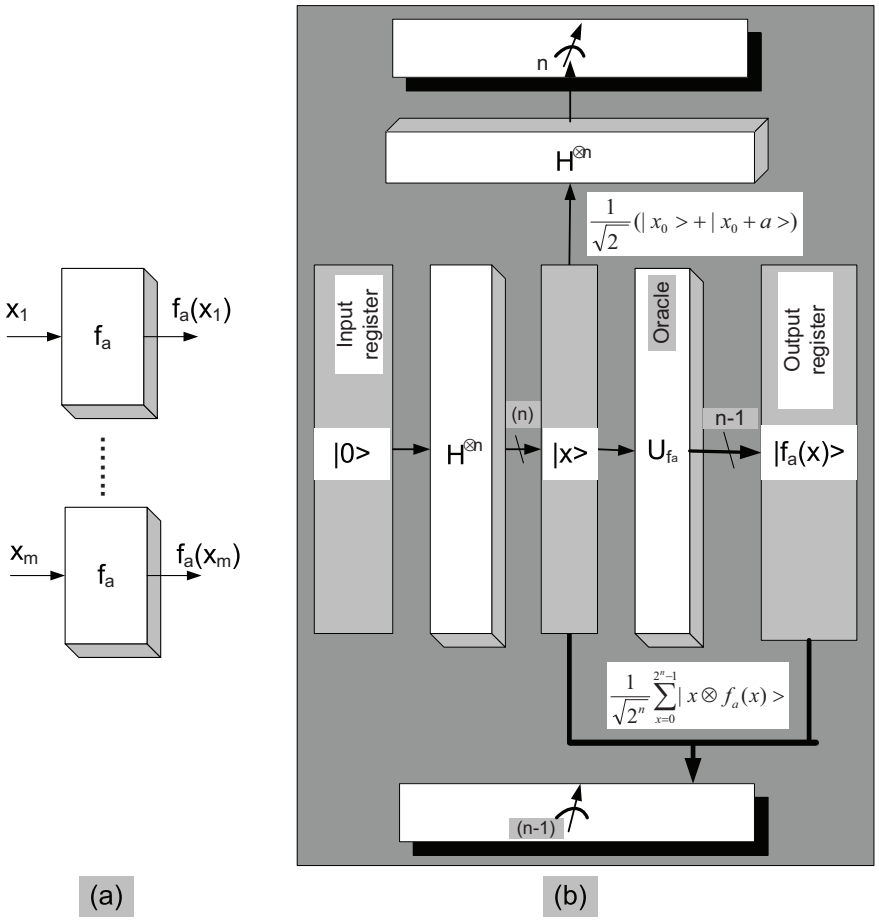



Figure 24: Simon's problem: find the period $a \neq 0$ of the function $f_a : \{0,1\}^n \mapsto \{0,1\}^{n-1}$. (a) A classical solution requires an exponential number of trials. There are $2^n$ possible values of $a$ and after $m$ evaluations of $f_a(x_i)$ we have eliminated $m(m-1)/2$ possible values of $a$. (b) The quantum solution: apply a Walsh-Hadamard transform to the input register on $n$ qubits in state $\mid 0\rangle^{(n)}$ to create an equal superposition state; then the oracle performs the unitary transformation $\mathbf{U}_f$. Now the joint state of the input and the output registers of the oracle is $\mid \varphi\rangle$; we measure the output register only. As a result of this measurement the input register will be left in the superposition state $1/\sqrt{2}\,(\mid x_0\rangle + \mid x_0 \oplus a\rangle)$ with $x_0$ the value of the argument $x$ corresponding to the outcome of the measurement we have observed, $f_a(x_0)$. Finally, apply a Walsh-Hadamard transform to the input register and then perform a measurement. Repeat this procedure $\mathcal{O}(n)$ times to determiner the $n$ bits of the period $a = (a_1 a_2 \ldots a_n)$.

values of the argument, $x_i, x_j, \quad f_a(x_i) = f_a(x_j) \implies x_i = x_j \oplus a$. Addition modulo 2 is associative and $x_i \oplus x_i = 0$. Thus, $x_i = x_j \oplus a \implies a = x_i \oplus x_j$; this suggests a method to determine $a$: compute $x_i \oplus x_j, \quad \forall x_j \quad 0 \leq x_j \leq 2^n - 1, x_j \neq x_i$ for every argument $x_i$. As we can see in Figure 24(a) after $m$ evaluations of $f_a(x_i)$ we have eliminated $\binom{m}{2} = m(m-1)/2$ possible values of $a$. The classical solution requires $\mathcal{O}(2^n)$ operations while Simon oracle finds the answer in $\mathcal{O}(n)$ operations.

At the heart of the solution provided by Simon oracle is a measurement of the quantum

state of a register of multiple qubits. Born's rule states that a measurement of one qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ produces the outcome 0 with probability $p_0 = \mid \alpha_0 \mid^2$ and 1 with probability $p_1 = \mid \alpha_1 \mid^2$. Now we outline an extension of the Born rule for a register of multiple qubits discussed in more detail in Section 2.5. First, we express the state of a register of $(n+1)$ qubits, $\mid \psi\rangle^{(n+1)}$, as a superposition of states of two groups of qubits, the first consists of a single qubit, that could be described in the orthonormal base $[\mid 0\rangle, \mid 1\rangle]$, and the second group of $n$ qubits:

$$\mid \psi\rangle^{(n+1)} = \alpha_0 \mid 0\rangle \otimes \mid \psi_0\rangle^{(n)} + \alpha_1 \mid 1\rangle \otimes \mid \psi_1\rangle^{(n)}.$$

If we only measure the single qubit the outcome is either 0 or 1. The extended Born rule states that when the outcome of the measurement is 0 then the state of the $n$ qubits in the second group is $\mid \psi_0\rangle^{(n)}$ and when the outcome of the measurement is 1 then the state of the $n$ qubits is $\mid \psi_1\rangle^{(n)}$.

After this brief digression we return to the problem solved by Simon oracle and consider the quantum circuit in Figure 24(b). This circuit uses the Walsh-Hadamard transform discussed in detail in Section 1.21; for now we only mention that it is the result of applying a Hadamard transformation to each one of the $n$ qubits of an input register.

First, we apply a Walsh-Hadamard transform to the input register on $n$ qubits in state $\mid 0\rangle^{(n)}$ and create an equal superposition state

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} \mid x\rangle.$$

This state is then transformed by the oracle to the state $\mid f_a(x)\rangle$. The state of the qubits in the output register of the oracle is an equal superposition of all $2^{n-1}$ values of the function $f_a(x)$; recall that each value $f_a(x_i)$ corresponds to a pair of values of the argument, namely $x_i$ and $x_i \oplus a$. The joint state of the input and the output registers of the oracle is

$$\mid \varphi\rangle = \left[\frac{1}{2^n} \sum_{x=0}^{2^n-1} \mid x\rangle\right] \otimes \mid f_a(x)\rangle = \frac{1}{2^n} \sum_{x=0}^{2^n-1} \mid x\rangle \otimes \mid f_a(x)\rangle.$$

Now we apply a measurement to the qubits in the output register only; all $2^{n-1}$ values of the function $f_a$ are equally likely to be the outcome of this measurement. Let $x_0$ be the value of the argument $x$ corresponding to the outcome of the measurement we have observed, $f_a(x_0)$. Then, according to the generalized Born rule, the input register will be left in a superposition corresponding to the two arguments $\mid x_0\rangle$ and $\mid x_0 \oplus a\rangle$, namely in the state

$$\frac{1}{\sqrt{2}} \left(\mid x_0\rangle + \mid x_0 \oplus a\rangle\right).$$

Unfortunately, from this superposition we cannot identify the two integers $x_0$ and $x_0 \oplus a$. If we could, then computing the period would be trivial: $x_0 \oplus (x_0 \oplus a) = a$. Nevertheless, we can use our old trick, apply the Walsh-Hadamard transformation to the input register. Recall that the Hadamard gate transform a single qubit as follows:

$$\mathbf{H} \mid x\rangle = \frac{\mid 0\rangle + (-1)^x \mid 1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \sum_{y=0,1} (-1)^{x \cdot y} \mid y\rangle.$$

Similarly, the Walsh-Hadamard transform of $n$ qubits is

$$\mathbf{H}^{\otimes n} \mid x\rangle^{(n)} = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} \mid y\rangle^{(n)}.$$

When we apply the Walsh-Hadamard transformation to the qubits in the input register we obtain

$$\mathbf{H}^{\otimes n} \left[ \frac{1}{\sqrt{2}} (\mid x_0\rangle + \mid x_0 \oplus a\rangle) \right] = \frac{1}{\sqrt{2^{n+1}}} \sum_{y=0}^{2n-1} \left( (-1)^{x_0 \cdot y} + (-1)^{(x_0 \oplus a) \cdot y} \right) \mid y\rangle^{(n)}.$$

Call $\mathcal{Y}_0$ the subset of $y$ such that $a \cdot y = 0$ and $\mathcal{Y}_1$ the subset of $y$ such that $a \cdot y = 1$. We observe that

$$\sum_{y \in \mathcal{Y}_1} \left[ (-1)^{x_0 \cdot y} + (-1)^{(x_0 \oplus a) \cdot y} \right] \mid y\rangle^{(n)} = 0.$$

Indeed, when $a \cdot y = 1$ then $(-1)^{(x_0 \oplus a) \cdot y} = (-1)^{x_0 \cdot y} (-1)^{a \cdot y} = -(-1)^{x_0 \cdot y}$. It follows that the input register is in state

$$\frac{1}{\sqrt{2^n + 1}} \sum_{y \in \mathcal{Y}_0} \left[ (-1)^{x_0 \cdot y} + (-1)^{x_0 \cdot y} \right] \mid y\rangle^{(n)} = \frac{1}{\sqrt{2^{n-1}}} \sum_{y \in \mathcal{Y}_0} (-1)^{x_0 \cdot y} \mid y\rangle^{(n)}.$$

Now we carry a measurement of the $n$ qubits in the input register and obtain with equal probability one of the values of $y \in \mathcal{Y}_0$, in other words a value of $y$ such that $y \cdot a = 0$. But:

$$y \cdot a = 0 \quad \Longrightarrow \quad \sum_{i=0}^{2^n-1} y_i a_i = 0 \mod 2.$$

Every value of $y$, with the exception of $y = 0$, allows us to establish a relation among the $n$ bits of $a$. Thus, the oracle allows us to determine $a$ after $\mathcal{O}(n)$ operations.

We have concluded the analysis of the first group of quantum algorithms and continue our presentation with the discussion of more sophisticated quantum algorithms. Useful information about the transformation carried out by a quantum algorithm is often encoded in the phases of the basis states. We discuss next a very useful procedure to extract this information through quantum phase estimation.

## 1.20 Quantum Phase Estimation

Quantum phase estimation is the process of determining the phase of the basis states after a transformation in a Hilbert space. Several quantum algorithms and the quantum Fourier transform use quantum phase estimation. We start our discussion with an example which illustrates the basic idea of quantum phase estimation; the problem we pose is to determine *the order of an element* in $GF(q)$, a finite field with $q$ elements. Factoring an integer $N$ can be reduced to the following problem: "given an integer $a$ find the smallest positive integer $r$ with the property that $a^r \equiv 0 \mod N$;" also, order finding is important for coding theory as

we shall see in Chapter 4. Section 4.4 provides an in-depth discussion of finite fields, here we simply state that integers modulo a prime number form an algebraic structure with a finite number of elements and two operations, addition and multiplication, with standard properties. For example, the integers modulo 5 form a finite field with 5 elements, $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$; $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$ is the set of nonzero elements in $\mathbb{Z}_5$. The order of an element $a_i \in \mathbb{Z}_5^*$ is the smallest integer $r$ such that $a_i^r = 1$. In our example, the orders of the 4 elements $\{1, 2, 3, 4\}$ in $\mathbb{Z}_5^*$ are, respectively, $\{1, 4, 4, 2\}$; indeed,

$$1^1 = 1 \mod 5, \quad 2^4 = 1 \mod 5, \quad 3^4 = 1 \mod 5, \quad 4^2 = 1 \mod 5.$$

The elements of $\mathbb{Z}_5^*$ can be expressed in a canonical base with $\mid j\rangle$ the binary expression of the element $j$, e.g., $\mid 3\rangle = \mid 011\rangle$. We are looking for a linear transformation $\mathbf{U}_a$ of the basis vectors with the property

$$\mathbf{U}_a \mid j\rangle = \mid aj\rangle, \qquad \mathbf{U}_a^2 \mid j\rangle = \mathbf{U}_{a^2} \mid j\rangle = \mid a^2 j\rangle, \qquad \ldots \qquad , \mathbf{U}_a^{2^n} \mid j\rangle = \mathbf{U}_{a^{2^n}} \mid j\rangle = \mid a^n j\rangle.$$

If $r$ is the order of $a$ then $\mathbf{U}_a^r = \mathbf{U}_{a^r} = \mathbf{I}$; this implies that the eigenvalues of $\mathbf{U}_a$ are of the form $e^{2\pi i(k/r)}$ with $k$ an integer and the eigenvectors of $\mathbf{U}_a$ are

$$\mid \varphi_k \mid = \sum_{j=0}^{r-1} e^{-2\pi i \frac{jk}{r}} \mid a^j\rangle.$$

We see that the information about $r$ is encoded in the phase $\omega_k = -2\pi i(jk/r)$ of the eigenvector $\mid \varphi_k \mid$. This justifies our interest in quantum phase estimation.

**Eigenvalue kickback.** We wish to construct a quantum circuit to determine the phase of a unitary transformation. The circuit in Figure 25 illustrates the mapping carried out by a controlled gate; $f(x)$ is the transformation applied to the target qubit and we see that the function $f(x)$ determines the phase of the control qubit:

$$\begin{aligned} \mid x\rangle \otimes (\mid 0\rangle - \mid 1\rangle) &\mapsto \mid x\rangle(\mid f(x) \oplus 0\rangle - \mid f(x) \oplus 1\rangle) \\ &\mapsto \mid x\rangle(\mid f(x)\rangle - \mid f(x) \oplus 1\rangle) \\ &\mapsto (-1)^{f(x)} \mid x\rangle(\mid 0\rangle - \mid 1\rangle) \end{aligned}$$

Indeed, if $f(x) = 0$ then $\mid x\rangle(\mid 0\rangle - \mid 1\rangle) \mapsto \mid x\rangle(-1)^0(\mid 0\rangle - \mid 1\rangle)$ as $0 \oplus 0 = 0$; similarly, when $f(x) = 1$ then $\mid x\rangle(\mid 0\rangle - \mid 1\rangle) \mapsto \mid x\rangle(-1)^1(\mid 0\rangle - \mid 1\rangle)$ as $1 \oplus 0 = 1$.

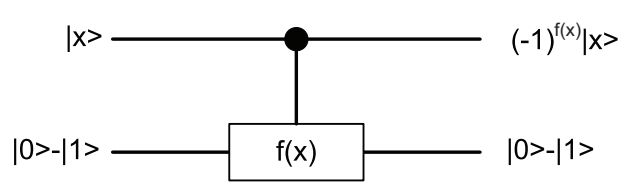



Figure 25: Eigenvalue kickback. The function $f(x)$ determines the phase of the control qubit.

Next we examine `controlled-U` circuits which apply unitary transformations $\mathbf{U}$ and $\mathbf{U}^k$ to their eigenvector $\mid \varphi\rangle$ with eigenphase $\omega$:

$$\mathbf{U} \mid \varphi\rangle = e^{2i\pi\omega} \mid \varphi\rangle \quad \text{and} \quad \mathbf{U}^k \mid \varphi\rangle = e^{2i\pi\omega k} \mid \varphi\rangle.$$

The intuition is that $\mathbf{U}^k$, *$k$ successive applications of the transformation* $\mathbf{U}$ *which rotates the eigenvector with an angle $\omega$, result in a rotation with an angle $\omega_k = \omega k$.* In Figure 26 we see that

$$\begin{array}{lcl}
\mathbf{U}(\mid 1\rangle \otimes \mid \varphi\rangle) & = & e^{2\pi i\omega} \mid 1\rangle \otimes \mid \varphi\rangle \\
\mathbf{U}(\mid 0\rangle \otimes \mid \varphi\rangle) & = & \mid 0\rangle \otimes \mid \varphi\rangle \\
\mathbf{U}((\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) \otimes \mid \varphi\rangle) & = & (\alpha_0 \mid 0\rangle + e^{2\pi i\omega}\alpha_1 \mid 1\rangle) \otimes \mid \varphi\rangle \\
\mathbf{U}^k(\mid 1\rangle \otimes \mid \varphi\rangle) & = & e^{2\pi i\omega k} \mid 1\rangle \otimes \mid \varphi\rangle \\
\mathbf{U}^k(\mid 0\rangle \otimes \mid \varphi\rangle) & = & e^{2\pi i\omega k} \mid 0\rangle \otimes \mid \varphi\rangle \\
\mathbf{U}^k((\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) \otimes \mid \varphi\rangle) & = & (\alpha_0 \mid 0\rangle + e^{2\pi i\omega k}\alpha_1 \mid 1\rangle) \otimes \mid \varphi\rangle
\end{array}$$

The `controlled-U` can be any unitary transformation including a controlled phase shift described by

$$\mathbf{R}_m = \begin{pmatrix} 1 & 0 \\ 0 & e^{\frac{2\pi i}{2^m}} \end{pmatrix}$$

If the phase is $\omega = 2^{-m}$ then the controlled phase shift $\mathbf{R}_m^{-1}$ rotates the eigenvector $\mid \varphi\rangle$ with an angle $(2\pi i)/2^m$ ; when the control qubit is $\mid 1\rangle$ then the eigenvalue is equal to 1, as shown in Figure 26.

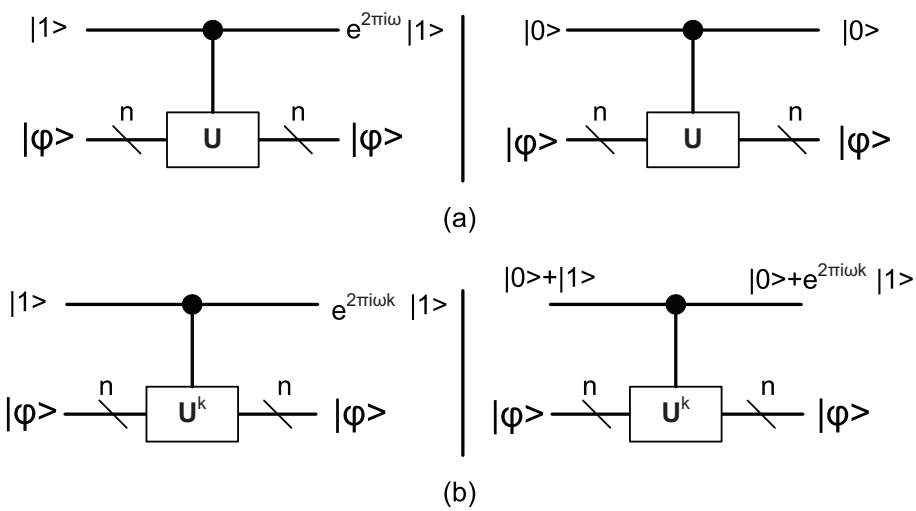



Figure 26: The eigenvalue of linear transformations for `controlled-`$U$ operations with an eigenvector as input. (a) The unitary transformation $\mathbf{U}$ applied to an eigenvector $\mid \varphi\rangle \in \mathcal{H}_{2^n}$ with an eigenphase $\omega$. (b) The unitary transformation $\mathbf{U}^k$.

After a phase shift we apply to the control qubit a Hadamard gate which carries out a transformation described by:

$$\mathbf{H}\frac{\mid 0\rangle + (-1)^j \mid 1\rangle}{\sqrt{2}} = \mid j\rangle \qquad \begin{cases} \mathbf{H}\frac{|0\rangle+|1\rangle}{\sqrt{2}} = \mid 0\rangle & \text{when} \quad j = 0 \\ \mathbf{H}\frac{|0\rangle-|1\rangle}{\sqrt{2}} = \mid 1\rangle & \text{when} \quad j = 1. \end{cases}$$

**Estimation of the eigenphase.** The problem we address now is to determine the real number $\omega \in (0, 1)$, the eigenphase of an eigenvector $\mid \varphi\rangle$ expressed as

$$\mid \varphi\rangle = \sum_{k=0}^{2^n-1} e^{2\pi i\omega k} \mid k\rangle.$$

We observe that

$$\sum_{k=0}^{2^n-1} e^{2\pi i\omega k} \mid k\rangle = \Big(\mid 0\rangle + e^{2\pi i(2^{n-1}\omega)} \mid 1\rangle\Big) \otimes \Big(\mid 0\rangle + e^{2\pi i(2^{n-2}\omega)} \mid 1\rangle\Big) \otimes \ldots \otimes \big(\mid 0\rangle + e^{2\pi i\omega} \mid 1\rangle\big).$$

We can express $\omega$ in binary as:

$\omega = \frac{\omega_1}{2} + \frac{\omega_2}{2^2} + \frac{\omega_3}{2^3} + \ldots + \frac{\omega_j}{2^j} + \ldots$ or in a compact form as $\omega = 0.\omega_1\omega_2\omega_3\ldots\omega_j\ldots$.

It is easy to see that

$$\begin{array}{lcl} 2\omega & = & \omega_1.\omega_2\omega_3\ldots\omega_j\ldots \\ 2^2\omega & = & \omega_1\omega_2.\omega_3\ldots\omega_j\ldots \\ \vdots & & \\ 2^{n-1}\omega & = & \omega_1\omega_2\ldots\omega_{n-1}.\omega_n\ldots \end{array}$$

Then we can express

$$\begin{array}{rl} \sum_{k=0}^{2^n-1} e^{2\pi i\omega k} \mid k\rangle = & \big(\mid 0\rangle + e^{2\pi i(0.\omega_n\omega_{n+1}\ldots)} \mid 1\rangle\big) \otimes \big(\mid 0\rangle + e^{2\pi i(0.\omega_{n-1}\omega_n\ldots)} \mid 1\rangle\big) \otimes \ldots \\ \cdots\otimes & \big(\mid 0\rangle + e^{2\pi i(0.\omega_2\omega_3\ldots)} \mid 1\rangle\big) \otimes \big(\mid 0\rangle + e^{2\pi i(0.\omega_1\omega_2\ldots)} \mid 1\rangle\big). \end{array}$$

**Example.** *We wish to estimate:* $\omega = 0.\omega_1\omega_2\omega_3$. A Hadamard gate transforms an input state $(\mid 0\rangle + e^{2\pi i(0.\omega_1)})/\sqrt{2}$ to $\mid \omega_1\rangle$, Figure 27(a)-(top); when we measure the output of the Hadamard gate we obtain the real value $\omega_1$, Figure 27(b)-(top).

The circuit in Figure 27(a)-(bottom) carries out the transformation

$$\frac{\mid 0\rangle + e^{2\pi i(0.\omega_3)} \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + e^{2\pi i(0.\omega_2\omega_3)} \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + e^{2\pi i(0.\omega_1\omega_2\omega_3)} \mid 1\rangle}{\sqrt{2}} \mapsto \mid \omega_1\rangle \otimes \mid \omega_2\rangle \otimes \mid \omega_3\rangle.$$

A measurement after a controlled-phase shift is equivalent to a measurement followed by a controlled phase-shift if the outcome (classical information) is equal to 1. The circuit in Figure 27(b)-(bottom) adds a measurement as the final stage of the circuit in Figure 27(a)-(bottom) and carries out the transformation which produces the real values $\omega_1, \omega_2$ and $\omega_3$

$$\frac{\mid 0\rangle + e^{2\pi i(0.\omega_3)} \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + e^{2\pi i(0.\omega_2\omega_3)} \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + e^{2\pi i(0.\omega_1\omega_2\omega_3)} \mid 1\rangle}{\sqrt{2}} \mapsto \omega_1\omega_2\omega_3.$$

We return now to the problem of finding the order of the elements of $\mathbb{Z}_5^*$, expressed as $r = 2k_2 + k_1$. We consider the unitary transformation denoted as $\mathbf{U}_2$ and its successive applications $\mathbf{U}_2^2$, $\mathbf{U}_2^3$ and $\mathbf{U}_2^4$. The basis vectors are transformed as follows:

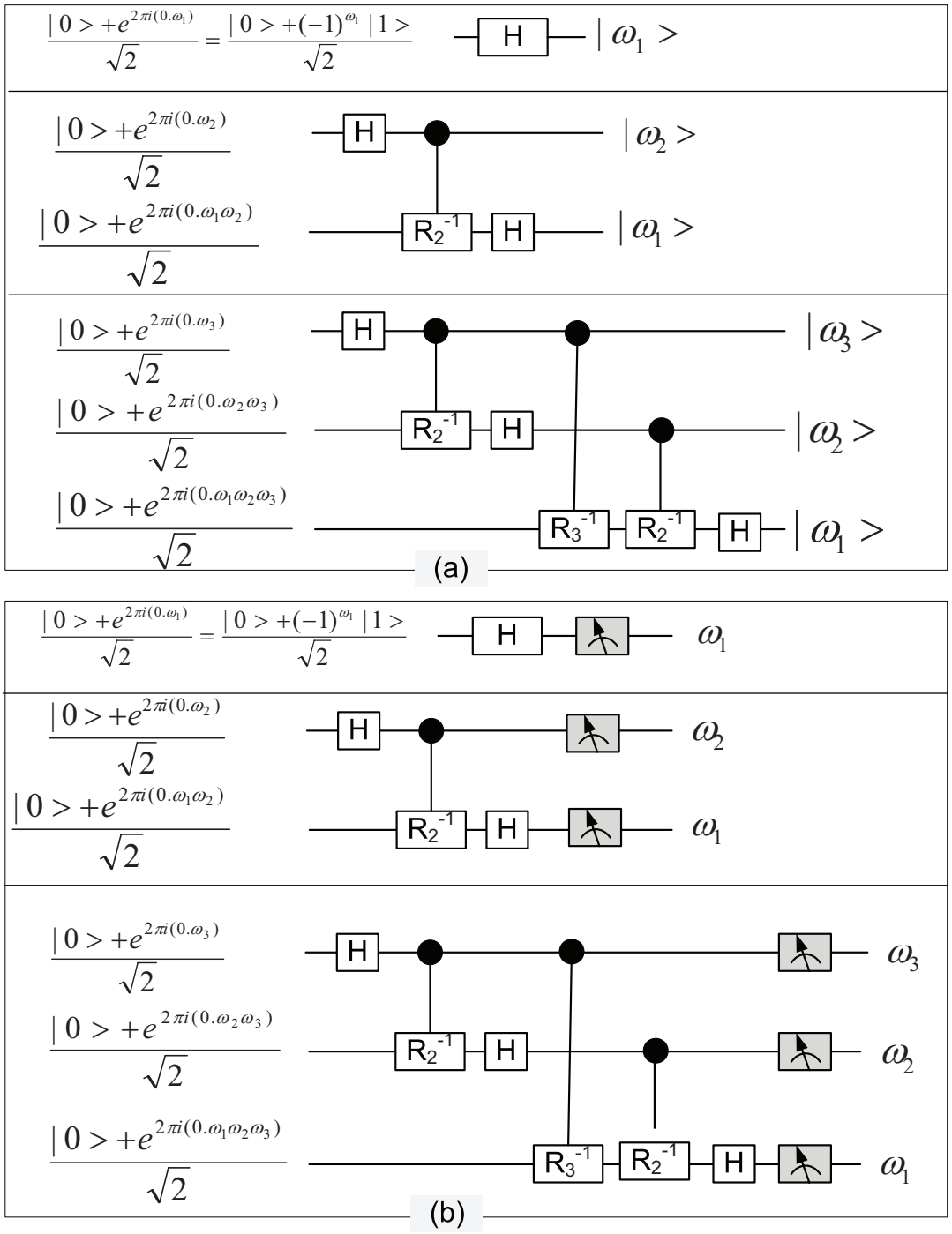



Figure 27: Quantum circuits to determine $\omega = 0.\omega_1\omega_2\omega_3$. (a) Controlled-phase shifts $\mathbf{R}_2$ and $\mathbf{R}_3$ followed by a Hadamard transform $\mathbf{H}$ produce: (top) $|\,\omega_1\rangle$; (middle) $|\,\omega_2\rangle$ and $|\,\omega_1\rangle$; (bottom) $|\,\omega_3\rangle$, $|\,\omega_2\rangle$, and $|\,\omega_1\rangle$. (b) A measurement after a controlled-phase shift is equivalent to a measurement followed by a controlled phase-shift if the outcome (classical information) is equal to 1. The outcome of the measurements are: (top) $\omega_1$; (middle) $\omega_2$ and $\omega_1$; (bottom) $\omega_3$, $\omega_2$, and $\omega_1$.

| | $\mathbf{U}_2$ | $\mathbf{U}_2^2$ | $\mathbf{U}_2^3$ | $\mathbf{U}_2^4$ |
|---|---|---|---|---|
| $\mid 001\rangle \mapsto$ | $\mid 010\rangle$ | $\mid 100\rangle$ | $\mid 011\rangle$ | $\mid 001\rangle$ |
| $\mid 010\rangle \mapsto$ | $\mid 100\rangle$ | $\mid 011\rangle$ | $\mid 001\rangle$ | $\mid 010\rangle$ |
| $\mid 011\rangle \mapsto$ | $\mid 001\rangle$ | $\mid 010\rangle$ | $\mid 100\rangle$ | $\mid 011\rangle$ |
| $\mid 100\rangle \mapsto$ | $\mid 011\rangle$ | $\mid 001\rangle$ | $\mid 010\rangle$ | $\mid 100\rangle$ |

The eigenvectors of $\mathbf{U}_2$ are:

$$\mid \varphi_j\rangle = \sum_{k=0}^{3} e^{-2\pi i \frac{jk}{4}} \mid 2^k \mod 5\rangle$$

or

$$\begin{aligned}
\mid \varphi_0\rangle &= \mid 001\rangle + \mid 010\rangle + \mid 100\rangle + \mid 011\rangle,\\
\mid \varphi_1\rangle &= \mid 001\rangle + e^{-2\pi i \frac{1}{4}} \mid 010\rangle + e^{-2\pi i \frac{2}{4}} \mid 100\rangle + e^{-2\pi i \frac{3}{4}} \mid 011\rangle,\\
\mid \varphi_2\rangle &= \mid 001\rangle + e^{-2\pi i \frac{2}{4}} \mid 010\rangle + e^{-2\pi i \frac{2\times 2}{4}} \mid 100\rangle + e^{-2\pi i \frac{3\times 2}{4}} \mid 011\rangle,\\
\mid \varphi_3\rangle &= \mid 001\rangle + e^{-2\pi i \frac{3}{4}} \mid 010\rangle + e^{-2\pi i \frac{2\times 3}{4}} \mid 100\rangle + e^{-2\pi i \frac{3\times 3}{4}} \mid 011\rangle.
\end{aligned}$$

Then

$$\tfrac{1}{2}(\mid \varphi_0\rangle + \mid \varphi_1\rangle + \mid \varphi_2\rangle + \mid \varphi_3\rangle) = \mid 001\rangle.$$

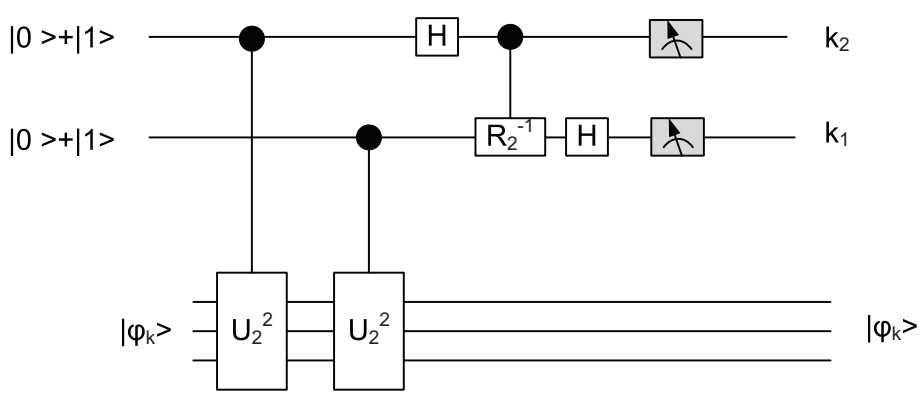



Figure 28: Quantum circuit to determine the order of an element in $\mathbb{Z}_5^*$; $r = 2k_2 + k_1$

The information about the order $r$ is encoded in the phase $\omega_j = -2\pi i(jk/r)$ of the eigenvector $\mid \varphi_j\rangle$

$$\begin{array}{l|l}
\mathbf{U}_2 \mid \varphi_0\rangle = \mid \varphi_0\rangle & \mathbf{U}_2 (\mid 0\rangle + \mid 1\rangle) \mid \varphi_0\rangle = (\mid 0\rangle + \mid 1\rangle) \mid \varphi_0\rangle\\
\mathbf{U}_2 \mid \varphi_1\rangle = e^{2\pi i \frac{1}{4}} \mid \varphi_1\rangle & \mathbf{U}_2 (\mid 0\rangle + \mid 1\rangle) \mid \varphi_1\rangle = \left(\mid 0\rangle + e^{2\pi i \frac{1}{4}} \mid 1\rangle\right) \mid \varphi_1\rangle\\
\mathbf{U}_2 \mid \varphi_2\rangle = e^{2\pi i \frac{2}{4}} \mid \varphi_2\rangle & \mathbf{U}_2 (\mid 0\rangle + \mid 1\rangle) \mid \varphi_2\rangle = \left(\mid 0\rangle + e^{2\pi i \frac{2}{4}} \mid 1\rangle\right) \mid \varphi_2\rangle\\
\mathbf{U}_2 \mid \varphi_3\rangle = e^{2\pi i \frac{3}{4}} \mid \varphi_3\rangle & \mathbf{U}_2 (\mid 0\rangle + \mid 1\rangle) \mid \varphi_3\rangle = \left(\mid 0\rangle + e^{2\pi i \frac{3}{4}} \mid 1\rangle\right) \mid \varphi_3\rangle.
\end{array}$$

We also see that

$$\begin{aligned}
\mathbf{U}_2^2\left(\mid 0\rangle + \mid 1\rangle\right) \mid \varphi_0\rangle &= \left(\mid 0\rangle + \mid 1\rangle\right) \mid \varphi_0\rangle \\
\mathbf{U}_2^2\left(\mid 0\rangle + \mid 1\rangle\right) \mid \varphi_1\rangle &= \left(\mid 0\rangle + e^{2\pi i \frac{1}{4} 2} \mid 1\rangle\right) \mid \varphi_1\rangle \\
\mathbf{U}_2^2\left(\mid 0\rangle + \mid 1\rangle\right) \mid \varphi_2\rangle &= \left(\mid 0\rangle + e^{2\pi i \frac{2}{4} 2} \mid 1\rangle\right) \mid \varphi_2\rangle \\
\mathbf{U}_2^2\left(\mid 0\rangle + \mid 1\rangle\right) \mid \varphi_3\rangle &= \left(\mid 0\rangle + e^{2\pi i \frac{3}{4} 2} \mid 1\rangle\right) \mid \varphi_3\rangle.
\end{aligned}$$

The circuit to find the order of the elements of $\mathbb{Z}_5^*$ is depicted in Figure 28.

We mentioned that the circuits for quantum phase estimation are also used for the quantum Fourier transform (QFT). Recall that the classical Discrete Fourier transform (DFT) performs the following mapping

$$j = (0, 0, \ldots, 0, 1, 0 \ldots, 0) \mapsto \left(1, e^{2\pi i \frac{j}{n}}, e^{2\pi i \frac{2j}{n}}, \ldots, e^{2\pi i \frac{(n-1)j}{n}}\right).$$

Similarly, QFT discussed in the next section transforms the basis vectors in $\mathcal{H}_{2^n}$ as follows

$$\mid j\rangle \mapsto \sum_{k=0}^{2^n-1} e^{2\pi i \frac{j}{2^n} k} \mid k\rangle.$$

We see immediately that the circuits in Figure 27 implement in fact an inverse QFT. There is yet another connection between phase estimation and QFT; we have assumed that the phase is of the form $\omega = j/2^k$ with $j$ an integer so the next question is if we can approximate $\omega$ by $\tilde{\omega}$ when this condition is not satisfied. The answer is that the QFT will transform the state as follows

$$\sum_{k=0}^{2^n-1} e^{2\pi i \omega k} \mid k\rangle \mapsto \mid \tilde{\omega}\rangle = \sum_j \alpha_j \mid j\rangle.$$

The error is bounded

$$\text{Prob}\left(\mid \frac{j}{2^n} - \omega \mid \leq \frac{1}{2^n}\right) \geq \frac{8}{\pi^2} \quad \text{and} \quad \mid \alpha_j \mid = \mathcal{O}\left(\frac{1}{\mid j/2^n - \omega \mid}\right).$$

We introduce next two transforms of interest for quantum computing: the Walsh-Hadamard transform used by several quantum algorithms, including Grover search algorithm, and the quantum Fourier transform used by several algorithms including Shor quantum factoring algorithm..

## 1.21 Walsh-Hadamard and Quantum Fourier Transforms

In this section we first introduce Hadamard matrices and then discuss the Walsh-Hadamard transform and the quantum Fourier transform (QFT).

**Hadamard matrix**. A Hadamard matrix of order $n$ is an $n \times n$ matrix with elements $h_{ij}$ either $+1$ or $-1$; a Hadamard matrix of order $2n$ is a $2n \times 2n$ matrix:

$$H(n) = [h_{ij}],\ 1 \le i \le n,\ 1 \le j \le n \quad \text{and} \quad H(2n) = \begin{pmatrix} H(n) & H(n) \\ H(n) & -H(n) \end{pmatrix}.$$

$H(n)$ has the following properties:

1. The product of a Hadamard matrix and its transpose satisfies the following relations

$$H(n)H(n)^T = nI_n \qquad H(n)^T H(n) = nI_n.$$

2. The exchange of rows or columns transforms one Hadamard matrix to another one.

3. The row vectors $\{h_1, h_2, \ldots, h_n\}$ of the matrix $H(n)$, are pairwise orthogonal

$$h_k \cdot h_l = 0 \qquad \forall (k,l) \in \{1, n\}.$$

4. The multiplication of rows or columns by $-1$ transforms one Hadamard matrix to another one.

5. If $n = 2^{q-1}$ then $H(2n)$ can be expressed as the tensor product of $q$ matrices of size $2 \times 2$

$$H(2n) = H(2) \otimes H(2) \otimes \ldots H(2) \otimes H(2).$$

Recall that the tensor product of the $p \times q$ matrix $A = [a_{ij}]$ with the $r \times s$ matrix $B = [b_{kl}]$ is the $(p \cdot r) \times (q \cdot s)$ matrix $C = [c_{mn}]$ with $c_{mn} = a_{ij} \cdot b_{kl}$ where $i, j, k, l, m, n$ are the only integers that satisfy the relations $m = (i-1)r + k$ and $n = (j-1)s + l$. In other words, $C$ is the matrix obtained by replacing the element $a_{ij}$ of matrix $A$ by the $r \times s$ matrix $a_{ij} \cdot B$.

Let us consider the first property: $H(n)$ and $H(n)^T$ can be written as

$$H(n) = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \qquad H(n)^T = \begin{pmatrix} h_1^T & h_2^T & \ldots & h_n^T \end{pmatrix}.$$

The elements of $h_k$ and $h_k^T$ are identical: $h_{ki} = h_{ik}$.
Then,

$$H(n)H(n)^T = \begin{pmatrix} h_1 \cdot h_1 & h_1 \cdot h_2 & \ldots & h_1 \cdot h_n \\ h_2 \cdot h_1 & h_2 \cdot h_2 & \ldots & h_2 \cdot h_n \\ \ldots & \ldots & \ldots & \ldots \\ h_n \cdot h_1 & h_n \cdot h_2 & \ldots & h_n \cdot h_n \end{pmatrix} = \begin{pmatrix} n & 0 & \ldots & 0 \\ 0 & n & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & n \end{pmatrix} = nI_n.$$

Now multiply the previous equation with $H(n)^{-1}$

$$H(n)^{-1}H(n)H(n)^T = nH(n)^{-1}I_n$$

It follows that

$$H(n)^T = nH(n)^{-1} \quad \Rightarrow \quad H(n)^T H(n) = nH(n)^{-1}H(n) \quad \Rightarrow \quad H(n)^T H(n) = nI_n.$$

Properties (2), (3), (4), and (5) follow immediately. The Hadamard matrices used in quantum computing are normalized

$$H(n) = 2^{-n/2}[h_{ij}].$$

**Example.** *Hadamard matrices of order* $2, 4$ *and* $8$ *are:*

$$H(2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad H(4) = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$H(8) = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}.$$

**The Walsh-Hadamard transform.** The Walsh-Hadamard transform performs a randomization operation, but it is perfectly reversible. Given $n$ qubits the transform allows us to construct a quantum mechanical system with $N = 2^n$ states. The Walsh-Hadamard transform, $\mathbf{H}^{\otimes n}$, rotates each of the $n$ qubits independently

$$\mathbf{H}^{\otimes n} = \mathbf{H}(N) \quad \text{with} \quad N = 2^n.$$

The Walsh-Hadamard transform is used to create an equal superposition of qubits when the input is in state $\mid 00\ldots0\rangle$. For example

$$\mathbf{H}(4) \mid 00\rangle = \frac{\mid 00\rangle + \mid 01\rangle + \mid 10\rangle + \mid 11\rangle}{2}$$

and

$$\mathbf{H}(8) \mid 000\rangle = \frac{\mid 000\rangle + \mid 001\rangle + \mid 010\rangle + \mid 011\rangle + \mid 100\rangle + \mid 101\rangle + \mid 110\rangle + \mid 111\rangle}{2\sqrt{2}}.$$

If we start in an arbitrary state $\mid \psi\rangle = \sum_{k=0}^{N-1} \alpha_k \mid k\rangle$ and apply the Walsh-Hadamard transform then the system reaches the state

$$\mid \varphi\rangle = \mathbf{H}^{\otimes n} \mid \psi\rangle = \sum_{k=0}^{N-1} \beta_k \mid k\rangle \quad \text{with} \quad \beta_k = \sum_{j=0}^{N-1} (-1)^p \alpha_j,$$

with $p = 0, 1$. For example, in $\mathcal{H}_4$ when $\mid \psi\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle + \alpha_2 \mid 10\rangle + \alpha_3 \mid 11\rangle$ we have

$$\beta_0 = \alpha_0+\alpha_1+\alpha_2+\alpha_3,\ \beta_1 = \alpha_0-\alpha_1+\alpha_2-\alpha_3,\ \beta_2 = \alpha_0+\alpha_1-\alpha_2-\alpha_3,\ \beta_3 = \alpha_0-\alpha_1-\alpha_2+\alpha_3.$$

Recall that the probability of projecting a superposition state

$$\mid \psi\rangle = \sum_{k=0}^{N-1} \alpha_k \mid k\rangle$$

onto the basis state $\mid k\rangle$ is equal to the square of the absolute value of the probability amplitude of that basis state, $p_k = \mid \alpha_k \mid^2$. The probability amplitude $\alpha_k$ is a complex number characterized by its magnitude and phase. The state $\mid \psi\rangle$ is characterized by two $N$-dimensional vectors: the probability amplitude, a vector of complex numbers, and the probability, a vector of real numbers

$$(\alpha_0, \alpha_1, \ldots, \alpha_{N-1}) \quad \text{and} \quad (\mid \alpha_0 \mid^2,\ \mid \alpha_1 \mid^2,\ \ldots,\ \mid \alpha_{N-1} \mid^2), \ \text{respectively.}$$

Two states may have the same probability vectors but different probability amplitude vectors. For example, the states

$$\mid \psi_1\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \qquad \text{and} \qquad \mid \psi_2\rangle = \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}$$

have different probability amplitude vectors $(1/\sqrt{2}, 1/\sqrt{2})$ and $(1/\sqrt{2}, -1/\sqrt{2})$, respectively, but have the same probability vector $(1/2, 1/2)$.

It follows that the Walsh-Hadamard transform of a superposition state of $n$ qubits preserves the magnitude of the probability amplitude of individual components.

**Quantum Fourier transform**. QFT is a unitary operator that transforms the vectors of an orthonormal basis

$$\{\mid 0\rangle, \mid 1\rangle \ldots, \mid j\rangle, \ldots, \mid k\rangle, \ldots, \mid N-1\rangle\} \in \mathcal{H}_N \quad \text{with} \quad N = 2^n$$

as follows:

$$\mid j\rangle \ \mapsto \ \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{i2\pi jk/N} \mid k\rangle, \quad i = \sqrt{-1}$$

QFT transforms a state $\mid \psi\rangle$ of a quantum system to another state $\mid \varphi\rangle$:

$$\mid \psi\rangle = \sum_{j=0}^{N-1} \alpha_j \mid j\rangle \quad \mapsto \quad \mid \varphi\rangle = \sum_{k=0}^{N-1} \beta_k \mid k\rangle.$$

The probability amplitude $\beta_k$ is the *Discrete Fourier transforms* (DFT) of the probability amplitude $\alpha_j,\ 0 \le j \le N-1$

$$\beta_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \alpha_j e^{i2\pi jk/N}.$$

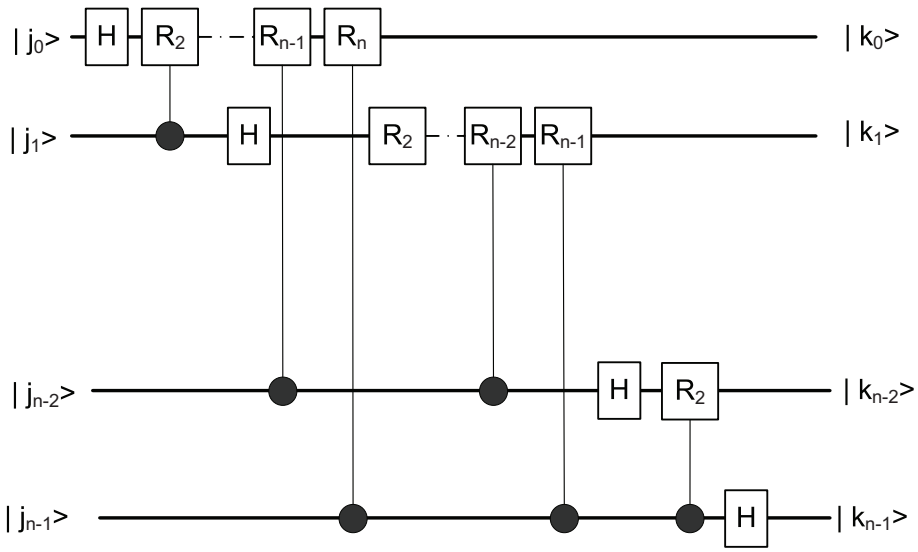



Figure 29: A circuit for quantum Fourier transform.

The binary representation of integers $j$ and $k$ is
$j = j_0 2^{n-1} + j_1 2^{n-2} + \cdots + j_{n-2} 2^1 + j_{n-1} 2^0$ and $k = k_0 2^{n-1} + k_1 2^{n-2} + \cdots + k_{n-2} 2^1 + k_{n-1} 2^0$.
Then, the definition of the QFT can be rewritten as

$$\mid j_0 j_1 \ldots j_{n-1}\rangle \;\mapsto\; \frac{1}{2^{n/2}} \sum_{k_0=(0,1)} \sum_{k_1=(0,1)} \cdots \sum_{k_{n-1}=(0,1)} e^{i2\pi j \sum_{m=0}^{n-1} k_m 2^{-m}} \mid k_0 k_1 \ldots k_m \ldots k_{n-1}\rangle,$$

$$\mid j_0 j_1 \ldots j_{n-1}\rangle \;\mapsto\; \frac{1}{2^{n/2}} \sum_{k_0=(0,1)} \sum_{k_1=(0,1)} \cdots \sum_{k_{n-1}=(0,1)} \bigotimes_{m=0}^{n-1} e^{i2\pi j k_m 2^{-m}} \mid k_m\rangle,$$

$$\mid j_0 j_1 \ldots j_{n-1}\rangle \;\mapsto\; \frac{1}{2^{n/2}} \bigotimes_{m=0}^{n-1} \{ \sum_{k_m=(0,1)} e^{i2\pi j k_m 2^{-m}} \mid k_m\rangle \}.$$

The bit $k_m$ may only take two values, 0 and 1. We note that $e^{i2\pi j k_m} = 1$ when $k_m = 0$, thus:

$$\mid j_0 j_1 \ldots j_{n-1}\rangle \;\mapsto\; \frac{1}{2^{n/2}} \bigotimes_{m=0}^{n-1} \{\mid 0\rangle + e^{i2\pi j 2^{-m}} \mid 1\rangle\}.$$

But

$$\mid 0\rangle + e^{i2\pi j 2^{-m}} \mid 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ e^{i2\pi j 2^{-m}} \end{pmatrix} = \begin{pmatrix} 1 \\ e^{i2\pi (j/2^m)} \end{pmatrix}$$

The transformation of the input $\mid j\rangle$ can be rewritten as

$$\mid j\rangle \;\mapsto\; \frac{1}{2^{n/2}} \begin{pmatrix} 1 \\ e^{i2\pi (j/2^0)} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ e^{i2\pi (j/2^1)} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ e^{i2\pi (j/2^2)} \end{pmatrix} \ldots \otimes \begin{pmatrix} 1 \\ e^{i2\pi (j/2^{n-1})} \end{pmatrix}$$

If we denote by $S_{p,q},\ 0 \leq p, q \leq n-1,\ \ p \neq q$, the joint transformation of qubits $p$ and $q$ as

$$S_{p,q} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\pi/2^{p-q}} \end{pmatrix}$$

we obtain an equivalent expression for the QFT that leads immediately to the quantum circuit for QFT in Figure 29. The QFTs of the $n$ qubits are given by the following transformations:

$$\begin{array}{lcr} \mid k_0\rangle & = & S_{0,n-1}\ S_{0,n-2} \ldots S_{0,4}\ S_{0,3}\ S_{0,2}\ S_{0,1}\ (H \mid j_0\rangle) \\ \mid k_1\rangle & = & S_{1,n-1}\ S_{1,n-2} \ldots S_{1,4}\ S_{1,3}\ S_{1,2}\ (H \mid j_1\rangle) \\ \mid k_2\rangle & = & S_{2,n-1}\ S_{2,n-2} \ldots S_{2,4}\ S_{2,3}\ (H \mid j_2\rangle) \\ \vdots & & \\ \mid k_{n-3}\rangle & = & S_{n-3,n-1}\ S_{n-3,n-2}\ (H \mid j_{n-2}\rangle) \\ \mid k_{n-2}\rangle & = & S_{n-2,n-1}\ (H \mid j_{n-2}\rangle) \\ \mid k_{n-1}\rangle & = & (H \mid j_{n-1}\rangle) \end{array}$$

All $S_{p,q}$ transform qubit $p$ as a target and use $q$ as a control qubit. $S_{p,q}$ does not change the value of a qubit, it only changes the phase. If both qubits $p$ and $q$ are equal to 1 then $S_{p,q}$ adds the value $\pi$ to the phases of qubit $p$. The one-qubit gates $R_m$ in Figure 29 represent the unitary transformation

$$R_m = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^m} \end{pmatrix}$$

This transformation is followed by a bit reversal. A qubit $\mid b\rangle$ is transformed to a qubit $\mid \bar{b}\rangle$ with $\bar{b}$ the bit reversal of $b$. If $b = b_0 2^{n-1} + b_1 2^{n-2} + \ldots b_{n-3} 2^2 + b_{n-2} 2 + b_{n-1}$ then $\bar{b} = b_{n-1} 2^{n-1} + b_{n-2} 2^{n-2} + b_{n-3} 2^{n-3} + \ldots b_2 2^2 + b_1 2 + b_0$.

The circuit shown in Figure 29 does not carry out the bit reversal.

## 1.22 Quantum Parallelism and Reversible Computing

We take a closer look at the transformations of quantum states required by a quantum computation. A classical computation evaluates a function $f$ for a particular value of the argument $x$; the outcome of this evaluation is $y = f(x)$. The picture is slightly different for a quantum computation which evaluates a function $f(\mid x\rangle)$, with $\mid x\rangle \in \mathcal{H}_{2^n}$. The unitary transformation $\mathbf{U}_f$ required by $f$, maps computational basis states to computational basis states. $\mathbf{U}_f$ is reversible; this means that at the end of the computation the $n$-qubit input register is in its initial state $\mid x\rangle$.

The quantum circuit for $\mathbf{U}_f$ should have access to an $m$-qubit register of ancilla qubits in state $\mid a\rangle$ to store partial results of the computation. The circuit in Figure 15 illustrates the use of ancilla qubits to store partial results for a reversible controlled gate with multiple control qubits. To simplify the presentation in Figure 30(a) we grouped together the ancilla qubits and the second input register shown in Figure 14 from Section 1.15. $\mathbf{U}_f$ carries out the following transformation

$$\mathbf{U}_f \left( \mid x\rangle^{(n)} \otimes \mid a\rangle^{(m)} \right) = \mid x\rangle^{(n)} \otimes \mid a \oplus f(x)\rangle^{(m)}.$$

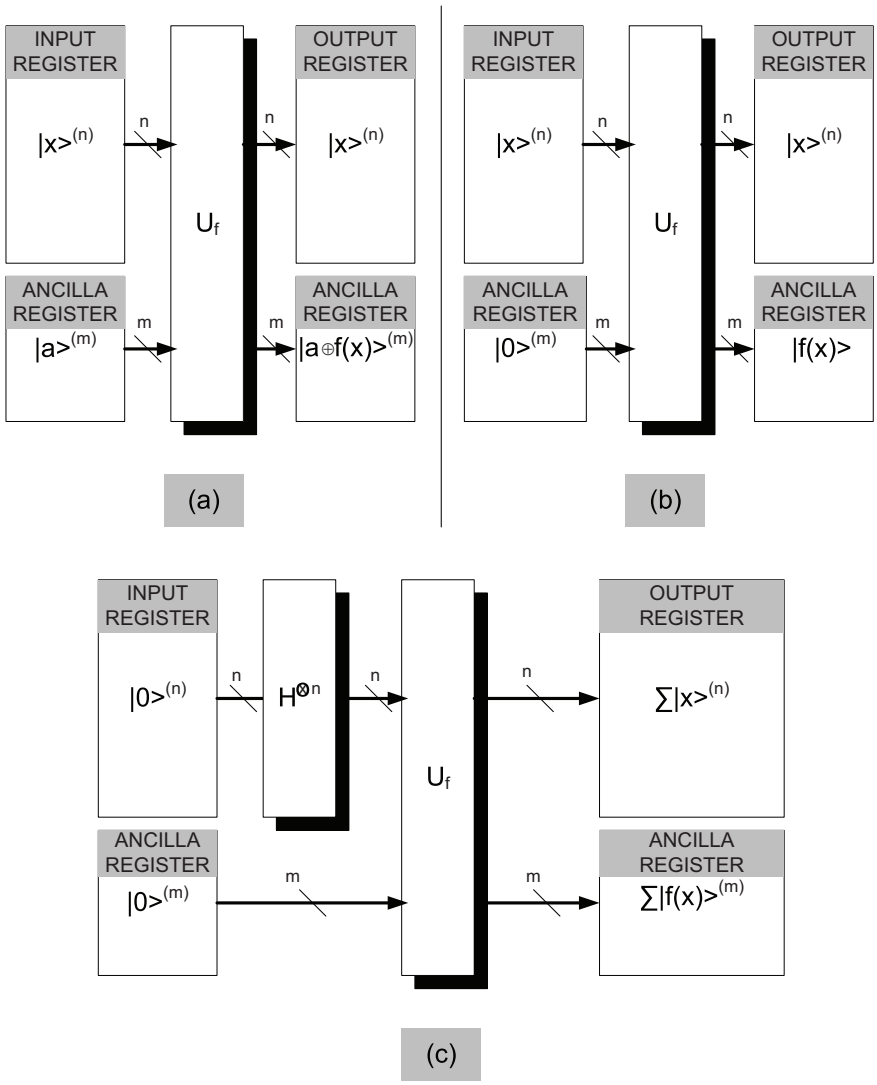



Figure 30: Quantum parallelism. (a) The unitary transformation $\mathbf{U}_f$ is applied to an input register of $n$ qubits in state $\mid x\rangle^{(n)}$; the transformation also requires a register of $m$ ancilla qubits in state $\mid a\rangle^{(m)}$. After the transformation the state of the ancilla qubits is $\mid a \oplus f(x)\rangle^{(m)}$. (b) When $\mid a\rangle^{(m)} = \mid 0\rangle^{(m)}$ the state of the ancilla qubits is $\mid f(x)\rangle^{(m)}$. (c) When $\mid x\rangle^{(n)} = \mid 0\rangle^{(n)}$ and we apply a Walsh-Hadamard transform to the qubits in the input register then the output is a superposition of the values of $f(x)$ for all values of the argument, $0 \leq x \leq 2^n - 1$.

The ancilla qubits must be returned to their initial state at the end of the computation, after recording the result of the computation. When the ancilla qubits are initially in state $\mid 0\rangle^{(m)}$, Figure 30(b), then $\mid a \oplus f(x)\rangle^{(m)} = \mid f(x)\rangle^{(m)}$ and the unitary transformation $\mathbf{U}_f$ returns $\mid f(x)\rangle^{(m)}$:

$$\mathbf{U}_f \left( \mid x\rangle^{(n)} \otimes \mid 0\rangle^{(m)} \right) = \mid x\rangle^{(n)} \otimes \mid f(x)\rangle^{(m)}.$$

In Section 1.21 we have learned that when we apply a Walsh-Hadamard transform to a register of $n$ qubits in state $\mid 0\rangle^{(n)}$ we obtain an equal superposition state, a state in which

the probability amplitudes of different states are equal to $\frac{1}{\sqrt{2^n}}$:

$$\mathbf{H}^{\otimes n}\left(\mid 0\rangle^{(n)}\right) = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} \mid i\rangle$$

We apply the Walsh-Hadamard transform, $\mathbf{H}^{\otimes n}$, to the input register and the identity transformation, $\mathbf{I}^{\otimes m}$ to the register of ancilla qubits:

$$\mathbf{U}_f\left(\mathbf{H}^{\otimes n} \otimes \mathbf{I}^{\otimes m}\right)\left(\mid 0\rangle^{(n)} \otimes \mid 0\rangle^{(m)}\right) = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} \mathbf{U}_f\left(\mid i\rangle^{(n)} \otimes \mid 0\rangle^{(m)}\right).$$

We have shown earlier that $\mathbf{U}_f\left(\mid x\rangle^{(n)} \otimes \mid 0\rangle^{(m)}\right) = \mid x\rangle^{(n)} \otimes \mid f(x)\rangle^{(m)}$ thus:

$$\mathbf{U}_f\left(\mathbf{H}^{\otimes n} \otimes \mathbf{I}^{\otimes m}\right)\left(\mid i\rangle^{(n)} \otimes \mid 0\rangle^{(m)}\right) = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} \mid i\rangle^{(n)} \otimes \mid f(i)\rangle^{(m)}.$$

The result of the transformation $\mathbf{U}_f$ is a superposition of the $2^n$ values of the function $f(x)$ for each of the $2^n$ possible values of the argument $x$, Figure 30(c). If $f$ is a one-to-one function then $m$ should be equal to $n$; sometimes, we require only a *Yes/No* answer and $m = 1$, e.g., in the case of several algorithms discussed in Section 1.19.

The outcome of a measurement of the output ancilla qubits reveals only one of the values of the function, e.g., $f(x_i)$, but we have no way either to discover the argument $x_i$, or to force the system to produce $f(x_0)$ for a particular argument $x_0$. Therefore, we have to defer the measurement of the output ancilla qubits; first, we should amplify the amplitude of the projection on the particular basis state corresponding to the argument of interest, or, carry out additional computations using as input this superposition state, and only then carry out the measurement to access the result of the computation.

Algorithmically, this means that we have to work harder to get the desired result, but there is a silver lining: the transformation $\mathbf{U}_f$ does the work of a classical system consisting of $2^n$ copies of the circuit evaluating in parallel the function $f : \{0,1\}^n \mapsto \{0,1\}^m$ for all possible values of its argument $x$. This property of quantum systems, referred to as quantum parallelism, is one of the reasons for the excitement triggered by the possibility of building a quantum computer. Think about a computation with $n = 100$; classically, to evaluate the value of a function $f(x)$ for all possible arguments $x \in \{0,1\}^{100}$ we would need $2^{100} \approx 10^{33}$ copies of the circuit, or, equivalently, $10^{33}$ time steps, with a single copy of the circuit; a quantum computer can realize this monumental task with a single copy of the circuit and in one time step. A critical facet of any quantum algorithm is to amplify the amplitude of the desired solution and we shall see this idea very clearly in our discussion of Grover search algorithm in Section 1.23.

Quantum, as well as classical algorithms start from an initial state and then cause a set of state transformations of the quantum or of the classical device, which eventually lead to the desired result. Indeed, the first step for any quantum computation is to initialize the system to a state that we can easily prepare; then we carry out a sequence of unitary transformations that cause the system to evolve towards a state which provides the answer to the computational problem.

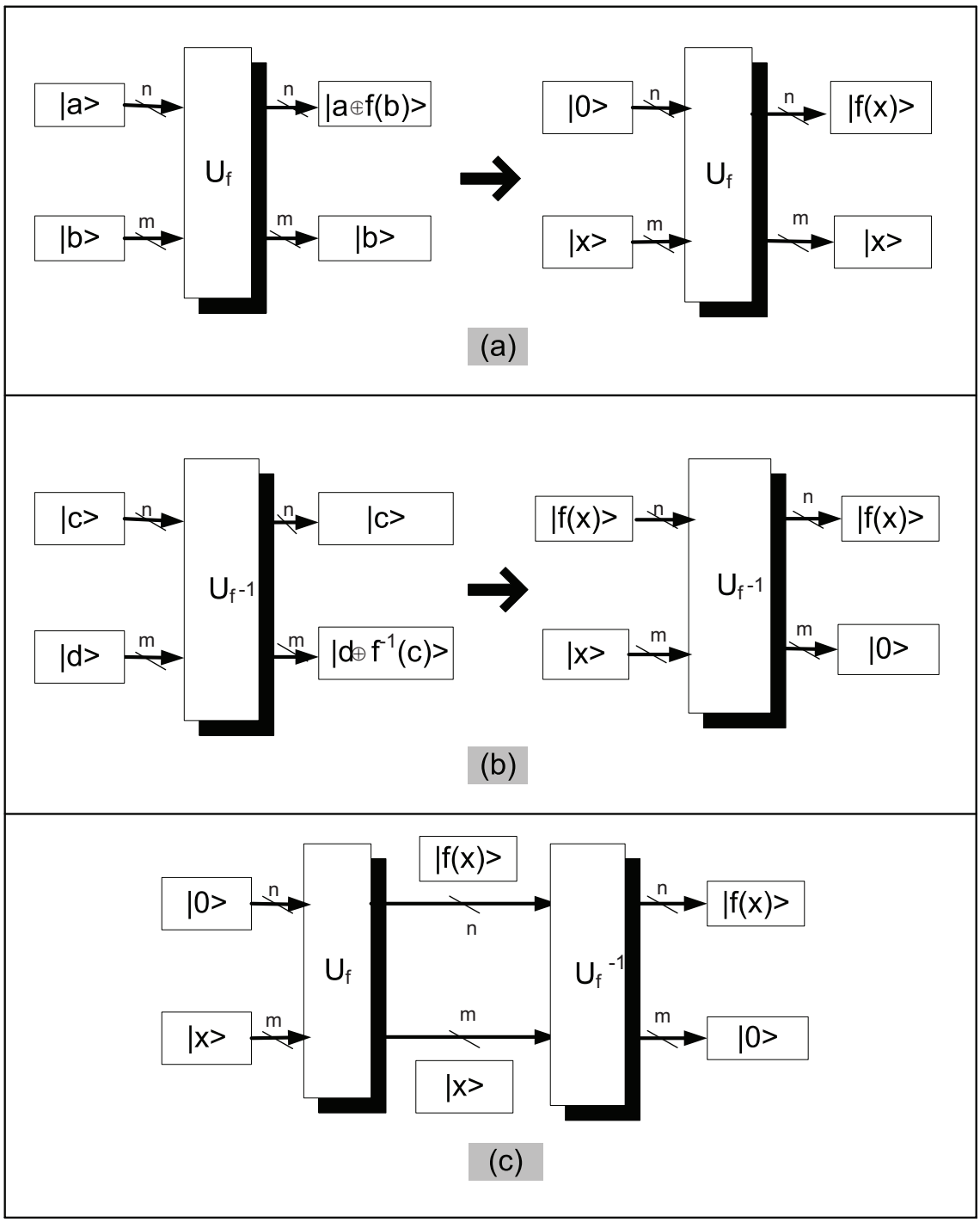



Figure 31: Reversible computation. (a) The quantum circuit $U_f$ maps its input $\mid a\rangle \otimes \mid b\rangle$ to $(\mid a \oplus f(b)\rangle) \otimes \mid b\rangle$; when $a = 0$ and $b = x$ then $\mathbf{U}_f(\mid 0\rangle \otimes \mid x\rangle) = \mid f(x)\rangle \otimes \mid x\rangle$. (b) $U_{f^{-1}}$ maps its input $\mid c\rangle \otimes \mid d\rangle$ to $\mid c\rangle \otimes (\mid d \oplus f^{-1}(c)\rangle)$; when $c = f(x)$ and $d = x$ then $\mathbf{U}_{f^{-1}}(\mid f(x)\rangle \otimes \mid x\rangle) = \mid f(x)\rangle \otimes \mid 0\rangle$. (c) If the two circuits work in tandem implementing the linear transformation $\mathbf{U}_f \mathbf{U}_{f^{-1}}$ they map the tensor product $\mid 0\rangle \otimes \mid x\rangle$ to $\mid f(x)\rangle \otimes \mid x\rangle$.

A quantum operation is a rotation of the state $\mid \psi\rangle$ in N-dimensional Hilbert space. Thus, the ultimate challenge is to build up powerful N-dimensional rotations as sequences of one and two dimensional rotations. For any quantum algorithm there are multiple paths leading from the initial to the final state and there is a degree of interference among these paths. The amplitude of the final state thus, the probability of reaching the desired final state, depends on the interference among these paths. This justifies the common belief that quantum algorithms are very sensitive to perturbations and one has to be extremely careful when choosing the transformations the quantum mechanical system is subjected to. Recently, Grover showed that the search algorithm is extremely robust and instead of using the Walsh-Hadamard transform one can in principle use almost any unitary transformation [186].

If we know how to compute efficiently a function $f(x)$ and its inverse $f^{-1}(x)$ then we can efficiently map any input state $\mid x\rangle$ to $\mid f(x)\rangle$ and the computation is reversible. If $U_f$ is a quantum circuit with input as well as output consisting of two registers of $n$ and, respectively, $m$ qubits, then $\mathbf{U}_f[\mid a\rangle \otimes \mid b\rangle] = (\mid a \oplus f(b)\rangle) \otimes \mid b\rangle$; when $a = 0$ and $b = x$ the circuit transforms its input $\mid 0\rangle \otimes \mid x\rangle$ to $\mid f(x)\rangle \otimes \mid x\rangle$ as shown in Figure 31(a).

If $U_{f^{-1}}$ is a quantum circuit with input as well as output consisting of two registers of $n$ and, respectively, $m$ qubits, then $\mathbf{U}_{f^{-1}}[\mid c\rangle \otimes \mid d\rangle] = \mid c\rangle \otimes (\mid d \oplus f^{-1}(c)\rangle)$; when $c = f(x)$ and $d = x$ the circuit transforms its input $\mid f(x)\rangle \otimes \mid x\rangle$ to $\mid f(x)\rangle \otimes \mid 0\rangle$ as shown in Figure 31(b); indeed $x \oplus f^{-1}(f(x)) = x \oplus x = 0$. When the two circuits work in tandem then they implement the linear transformation $\mathbf{U}_f \mathbf{U}_{f^{-1}}$ mapping the tensor product $\mid 0\rangle \otimes \mid x\rangle$ to $\mid f(x)\rangle \otimes \mid x\rangle$, as shown in Figure 31(c).

This brief discussion reinforces the realization that the development of quantum algorithms requires a different thinking than the one required for the development of classical algorithms. We cannot replicate the ability of a classical computation to reveal immediately the result of the evaluation of a function $f(x)$; the superposition state resulting from the application of the unitary transformation $\mathbf{U}_f$ reveals only the form of the solution dictated by the shape of the function $f$. This partially explains why so few quantum algorithms have been developed so far, a topic addressed by Shor in [387].

## 1.23 Grover Search Algorithm

We now discuss the quantum search algorithm of Grover, a discovery impressive not only due to its simplicity and elegance, but also to the range of potential applications. Preskill called Grover's algorithm "perhaps the most important new development" in quantum computing. "If quantum computers are being used 100 years from now, I would guess they will be used to run Grover algorithm or something like it," Preskill says. The quantum search illustrates the quintessence of a quantum algorithm: take advantage of quantum parallelism to create a superposition of all possible values of a function and then amplify, i.e., increase the probability amplitude of the solution. It also illustrates the contrast between classical and quantum algorithm strategies: *a classical search algorithm continually reduces the amplitude of non-target states, while a quantum search algorithm amplifies the amplitude of the target states.* In this context, to amplify means to increase the probability.

**Introduction.** Grover describes the series of steps that led to his algorithm in [188]; he starts by discretizing the Schrödinger equation and considers a unidimensional case. Grover

argues that when we start with a uniform linear superposition and let it evolve in the presence of a potential function it will gravitate towards points at which the potential is lower; therefore, if we want to design an algorithm to reach specific marked states, these marked states should correspond to a lower potential. The algorithm should implement several iterations of the transformations obtained from the evolution of Schrödinger equation.

The main idea of the quantum search algorithm [184, 185, 186, 187] is to rotate the state vector in a two-dimensional Hilbert space defined by an initial and a final (target) state vector. The algorithm is iterative and each iteration causes the same amount of rotation. A more recent version of the algorithm, the *fixed-point quantum search* discovered by Grover [190] could be used for quantum error correction [356], as discussed in Section 5.24.

The speedup of Grover algorithm is achieved by exploiting both quantum parallelism and the fact that, according to quantum theory, a probability is the square of an amplitude. Bennett and his co-workers [42] and Zalka [466] showed that Grover algorithm is optimal. No classical or quantum algorithm can solve this problem faster than time of order $\sqrt{N}$.

Grover search algorithm can be applied directly to a wide range of problems, see for example [165]. Even problems not generally regarded as searching problems, can be reformulated to take advantage of quantum parallelism and entanglement, and lead to algorithms which show a square root speedup over their classical counterparts [278].

**The intuition.** We consider a search space $\mathcal{Q} = \{q_i\}$ consisting of $N = 2^n$ items; each item $q_i, 1 \leq i \leq N$, is uniquely identified by a binary $n$-tuple $i$, called *the index* of the item. We assume that $M \leq N$ items satisfy the requirements of a query and we wish to identify one of them.

The classic approach is to repeatedly select an item $q_i$, decide if the item is a solution to the query, and if so, terminate the search. If there is a single solution ($M = 1$) then a classical search algorithm requires $\mathcal{O}(N)$ iterations; in the worst case, we need to examine all $N$ elements; if we repeat the experiment many times then, on average, we will end up examining $N/2$ elements before finding the desired one.

In this section we provide an intuitive explanation for two transformations at the heart of Grover algorithm, the <u>inversion about the mean</u> in classical context and the <u>phase inversion</u>.

Given a set of integers $\mathcal{Q} = \{q_i\}$ with the mean $\bar{q} = 1/N \sum_{i=1}^{N} q_i$, the inversion about the mean, denoted as $P$, transforms $q_i$ to $q_i'$:

$$P : q_i \mapsto q_i' \quad \text{with} \quad q_i' = \bar{q} + (\bar{q} - q_i) = 2\bar{q} - q_i.$$

In the example depicted in Figures 32 (a) and (b) the mean is $\bar{q} = 49$ and the individual integers are transformed as follows:

$$34 \mapsto 98 - 34 = 64, \quad 66 \mapsto 98 - 66 = 32, \quad 47 \mapsto 98 - 47 = 51, \quad 63 \mapsto 98 - 63 = 35,$$
$$54 \mapsto 98 - 54 = 44, \quad 28 \mapsto 98 - 28 = 70, \quad 42 \mapsto 98 - 42 = 56, \quad 58 \mapsto 98 - 58 = 40.$$

The inversion about the mean can be reformulated if we represent the transformation $P$ as matrix of state transitions, the elements of the set $\mathcal{Q}$ as a column vector $Q$, and the set of new values as the column vector $Q'$:

$$Q' = 2PQ - Q = (2P - I)Q$$

with

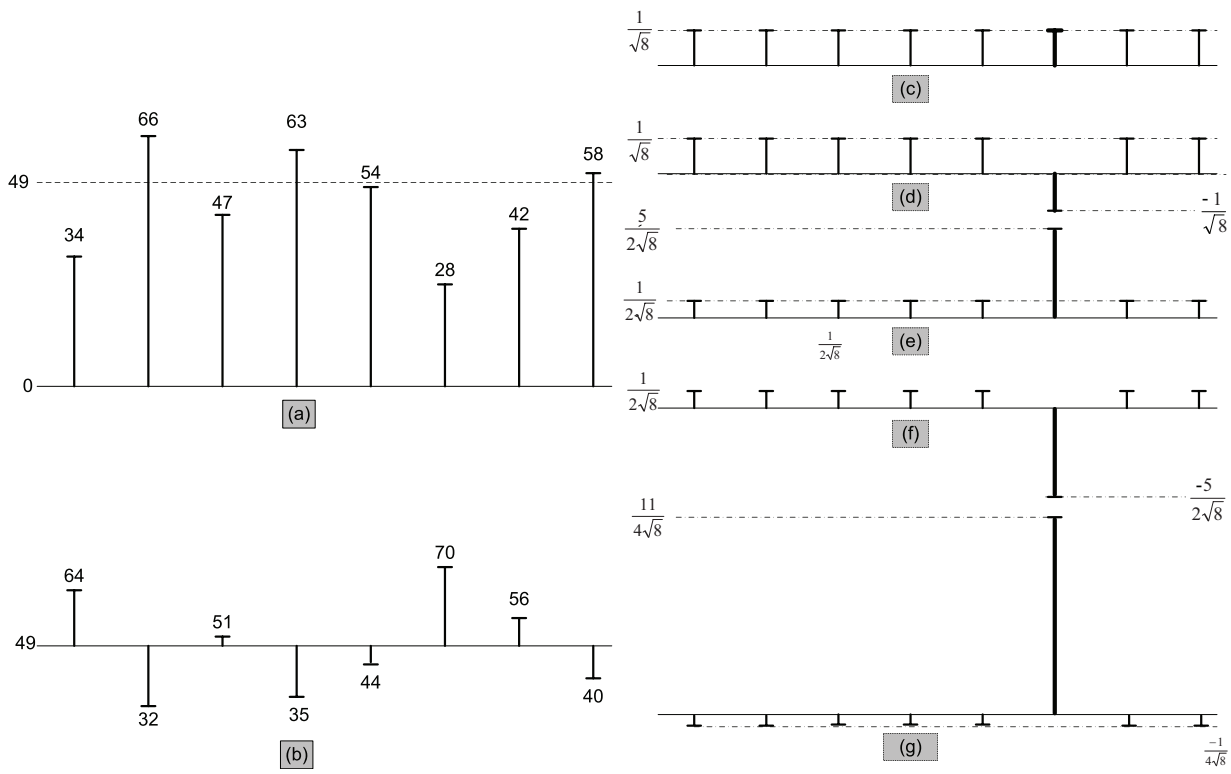



Figure 32: Graphic illustration of the inversion about the mean and the phase inversion; the amplitude of vertical bars is proportional to the modulus of the numbers. (a) A set of eight integers $\mathcal{Q} = \{34, 66, 47, 63, 54, 28, 42, 48\}$ with average $\bar{q} = 1/8 \sum_{i=1}^{8} q_i = 49$; (b) Inversion about the mean of the set $\mathcal{A}$; integer $q_i$ becomes $q_i' = \bar{q} + (\bar{q} - q_i) = 2\bar{q} - q_i$. (c) A set of eight complex numbers of equal amplitude $q_i = 1/\sqrt{8}$. (d) Phase inversion of the 6-th item, the one we search for. (e) Inversion about the mean from step (d): $\bar{q} = 3/(4\sqrt{8})$; it follows that: $q_1' = q_2' = q_3' = q_4' = q_5' = q_7' = q_8' = 1/(2\sqrt{8})$ and $q_6' = 5/(2\sqrt{8})$. (f) A second phase inversion of the item we search for. (g) A second inversion about the new mean from step (e): $\bar{q}' = 1/(8\sqrt{8})$; now $q_1'' = q_2'' = q_3'' = q_4'' = q_5'' = q_7'' = q_8'' = -1/(4\sqrt{8})$ and $q_6'' = 11/(4\sqrt{8})$.

$$P = \frac{1}{N} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}, \quad Q' = \begin{pmatrix} q_1' \\ q_2' \\ \vdots \\ q_N' \end{pmatrix}.$$

Next we consider the case when we label the $N$ items with complex numbers of equal amplitude $q_i = 1/\sqrt{N}$ and require that after each transformation the sum of the squares of the moduli of the complex numbers is one; we also assume that there is an oracle capable to invert the phase of the complex number identifying the item we search for. The inversion about the mean amplifies the amplitude of the item we search for and reduce the amplitude of all other items. When $q_1 = q_2 = q_3 = q_4 = q_5 = q_7 = q_8 = 1/(2\sqrt{8})$ the average amplitude after the first phase inversion, Figure 32 (c) and (d), is

$$\bar{q} = \frac{7(\frac{1}{\sqrt{8}}) - \frac{1}{\sqrt{8}}}{8} = \frac{3}{4\sqrt{8}}.$$

Then, the inversion about $\bar{q}$, Figure 32 (e), leads to

$$q_1' = q_2' = q_3' = q_4' = q_5' = q_7' = q_8' = 2\left(\frac{3}{4\sqrt{8}}\right) - \frac{1}{\sqrt{8}} = \frac{1}{2\sqrt{8}}$$

and

$$q_6' = 2\left(\frac{3}{4\sqrt{8}}\right) + \frac{1}{\sqrt{8}} = \frac{5}{2\sqrt{8}}.$$

A second phase inversion of the item we search for, Figure 32 (f), followed by a second inversion about the new mean

$$\bar{q}' = \frac{7(\frac{1}{2\sqrt{8}}) - \frac{5}{2\sqrt{8}}}{8} = \frac{1}{8\sqrt{8}}$$

leads to the situation depicted in Figure 32 (g). We notice the reduction of the amplitudes of the complex numbers labelling the items which do not match our search

$$q_1'' = q_2'' = q_3'' = q_4'' = q_5'' = q_7'' = q_8'' = 2\left(\frac{1}{8\sqrt{8}}\right) - \frac{1}{2\sqrt{8}} = -\frac{1}{4\sqrt{8}}$$

and the further amplification of the amplitude of the item we search for

$$q_6'' = 2\left(\frac{1}{8\sqrt{8}}\right) + \frac{5}{2\sqrt{8}} = \frac{11}{4\sqrt{8}}.$$

The example in Figures 32 (c), (d), (e), (f), and (g) allows us to refocus the discussion from classical to quantum search; we now view the complex numbers $q_i, 1 \leq i \leq 8$, as the probability amplitudes of the basis states of the canonical base in $\mathcal{H}_8$. The original state of the system is

$$\mid \varphi\rangle = q_1 \mid 000\rangle + q_2 \mid 001\rangle + q_3 \mid 010\rangle + q_4 \mid 011\rangle + q_5 \mid 100\rangle + q_6 \mid 101\rangle + q_7 \mid 110\rangle + q_8 \mid 111\rangle.$$

The final state after two iterations of phase inversion followed by inversion about the mean is

$$\mid \varphi\rangle'' = q_1'' \mid 000\rangle + q_2'' \mid 001\rangle + q_3'' \mid 010\rangle + q_4'' \mid 011\rangle + q_5'' \mid 100\rangle + q_6'' \mid 101\rangle + q_7'' \mid 110\rangle + q_8'' \mid 111\rangle.$$

The item we search is labelled 101 and we have amplified the probability amplitude of the state $\mid 101\rangle$ from $1/8 = 0.125$ to $(11/4\sqrt{8})^2 = 121/128 = 0.97227$ in two iterations; thus, we have increased the probability that a measurement of the state $\mid \varphi\rangle''$ produces the result 101.

**Phase inversion and the oracle.** For simplicity, we assume that the database $\mathcal{D}$ contains $N = 2^n$ items and only one item satisfies the query. The basic idea of the quantum search is to associate the index of an item $0 \leq x \leq 2^n - 1$ with the corresponding basis state $\mid x\rangle$ of a Hilbert space $\mathcal{H}_{2^n}$. The canonical base in $\mathcal{H}_{2^n}$ is $\{\mid 0\rangle, \mid 1\rangle, \mid 2\rangle, \ldots, \mid 2^n - 1\rangle\}$.

The quantum search algorithm requires an <u>oracle</u> to identify the index $x_0$ of the item we are searching for. To mark the desired item the oracle performs a phase inversion. To understand the inner working of the oracle we consider a function $f(x)$ such that

$$f(x) = \begin{cases} 0 & \text{if } x \neq x_0 \quad \Rightarrow \quad x \text{ is not a solution to the query} \\ 1 & \text{if } x = x_0 \quad \Rightarrow \quad x \text{ is a solution to the query.} \end{cases}$$

A black box performing a unitary transformation $\mathbf{U}_f$ for a function $f : \{0,1\}^n \mapsto \{0,1\}$ is shown in Figure 33(a). The black box used as an oracle accepts as input $n$ qubits representing the basis state $\mid x\rangle$ corresponding to the index $x$ and an *oracle qubit*, $\mid y\rangle$; the oracle qubit is flipped if $f(x) = 1$, Figure 33(b). The oracle recognizes the solution, inverts its phase and produces at the output $(-1)^{f(x)}[\mid x\rangle \otimes \mid y\rangle]$. If the oracle qubit is set $\mid y\rangle = \mid 1\rangle$, after the Hadamard gate $H$ is applied to it the qubit state becomes $\frac{|0\rangle - |1\rangle}{\sqrt{2}}$. When the oracle recognizes the solution the oracle qubit is flipped to the state $\frac{|1\rangle - |0\rangle}{\sqrt{2}}$ or $-(\frac{|0\rangle - |1\rangle}{\sqrt{2}})$. The output of the oracle becomes $[(-1)^{f(x)} \mid x\rangle] \otimes (\mid 0\rangle - \mid 1\rangle)/\sqrt{2}$. The basis state corresponding to the index $x_0$ of the item we are searching for is branded as the solution by the oracle and its phase is inverted; the oracle qubit remains unchanged. To take advantage of the quantum parallelism we apply as the input to the oracle an equal superposition of all basis states

$$\mid \psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} \mid x\rangle.$$

The oracle carries a phase inversion denoted as $\mathbf{O}_{x_0}$, in other words, it inverts the phase of the basis state corresponding to $x_0$ and leaves all other basis states unchanged. The $(n+1)$ qubit output of the oracle is

$$[\mathbf{O}_{x_0} \mid \psi\rangle] \otimes \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}.$$

**Inversion about the mean.** For $\mid \psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} \mid x\rangle$, an equal superposition state, the operator describing the conditional phase shift, or inversion about the mean, is

$$\mathbf{D}_\psi = 2\mathbf{P}_\psi - \mathbf{I} = 2 \mid \psi\rangle\langle\psi \mid - \mathbf{I},$$

or

$$\mathbf{D}_\psi = 2 \begin{pmatrix} \frac{1}{\sqrt{N}} \\ \frac{1}{\sqrt{N}} \\ \vdots \\ \frac{1}{\sqrt{N}} \\ \vdots \\ \frac{1}{\sqrt{N}} \end{pmatrix} (\frac{1}{\sqrt{N}} \frac{1}{\sqrt{N}} \cdots \frac{1}{\sqrt{N}} \cdots \frac{1}{\sqrt{N}}) - \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & & & & & \\ 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix}.$$

The matrix representation of the operator $\mathbf{D}_\psi$ is

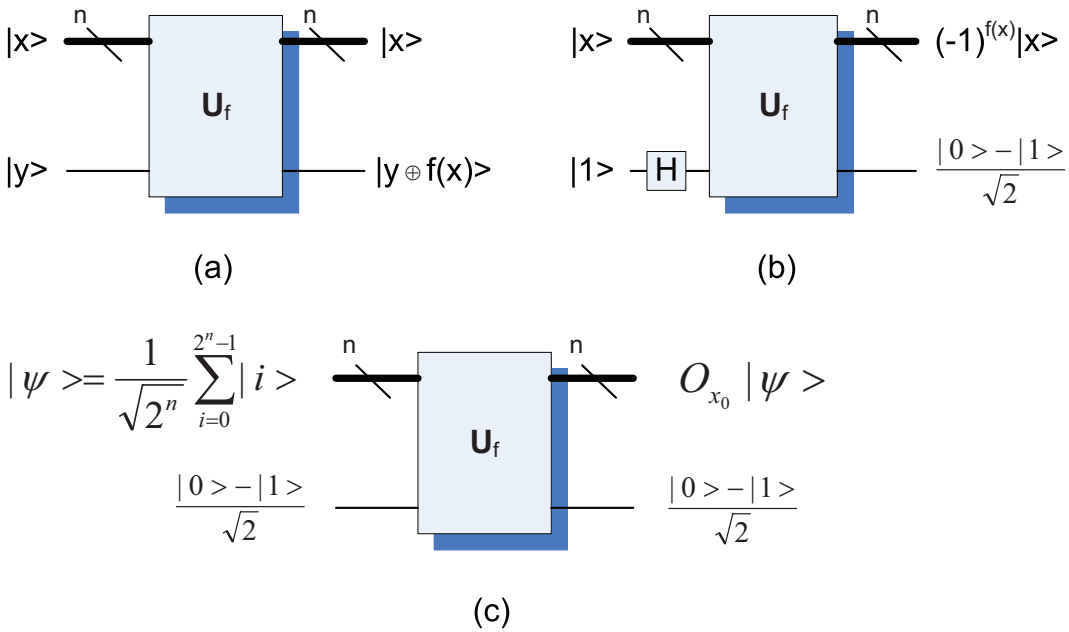



Figure 33: The oracle for Grover quantum search algorithm. (a) A black box performing a unitary transformation $\mathbf{U}_f$ of its $n$-qubit input $\mid x\rangle$ with $f : \{0,1\}^n \mapsto \{0,1\}$. (b) The black box used as an oracle carries out a phase shift of the input $\mid x\rangle$ when $x = x_0$ and $\mathbf{O}_{x_0} \mid x_0\rangle \rightarrow (-1)^{f(x_0)} \mid x_0\rangle$; here, $x_0$ is the index of the item we are searching for. (c) When we apply $\mid \psi\rangle$, a superposition of all basis states, as the input, then the oracle carries a phase inversion denoted as $\mathbf{O}_{x_0}$ and produces $\mathbf{O}_{x_0} \mid \psi\rangle$ at the output; the oracle qubit remains unchanged.

$$\mathbf{D}_\psi = \begin{pmatrix} \frac{2}{N}-1 & \frac{2}{N} & \dots & \frac{2}{N} & \dots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N}-1 & \dots & \frac{2}{N} & \dots & \frac{2}{N} \\ \vdots & & & & & \\ \frac{2}{N} & \frac{2}{N} & \dots & \frac{2}{N}-1 & \dots & \frac{2}{N} \\ \vdots & & & & & \\ \frac{2}{N} & \frac{2}{N} & \dots & \frac{2}{N} & \dots & \frac{2}{N}-1 \end{pmatrix}.$$

It is easy to see that $\mathbf{D}_\psi^\dagger = [(\mid \psi\rangle\langle\psi \mid -\mathbf{I}]^\dagger = \mathbf{D}_\psi$ thus, the inversion about mean is unitary

$$\mathbf{D}_\psi \mathbf{D}_\psi^\dagger = \mathbf{D}_\psi^\dagger \mathbf{D}_\psi = \mathbf{I}.$$

To convince ourselves that individual basis states suffer an inversion about the mean we consider a system in state $\mid \xi\rangle \in \mathcal{H}_N$

$$\mid \xi\rangle = \sum_{i=0}^{N-1} a_i \mid i\rangle.$$

Then,

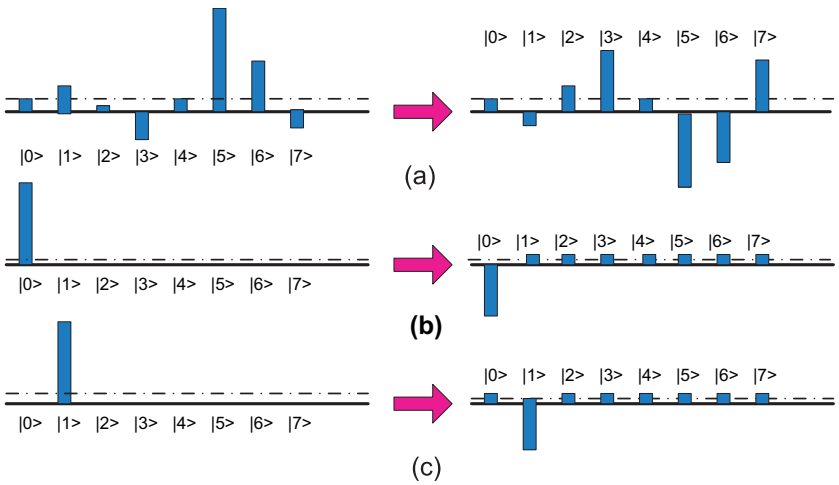



Figure 34: (a) An example of inversion about the mean in $\mathcal{H}_{2^3}$. (b) and (c) The inversion about mean of basis vectors $\mid 0\rangle$ and $\mid 1\rangle$.

$$\mid \zeta\rangle = \mathbf{D}_\psi \mid \xi\rangle = \sum_{k=0}^{N-1} b_k \mid k\rangle.$$

The probability amplitude $b_k$ is

$$b_k = 2\bar{a} - a_k, \quad 0 \leq k \leq N-1 \quad \text{with} \quad \bar{a} = \frac{1}{N}\sum_{i=0}^{N-1} a_i.$$

Jozsa [225] observes that any state $\mid \xi\rangle$ can be expressed as a sum of components parallel and orthogonal to $\mathbf{D}_\psi \mid \xi\rangle$, namely $\mid \alpha\rangle$ and $\mid \beta\rangle$. He shows that given any state $\mid \xi\rangle$, $\mathbf{D}_\psi \mid \xi\rangle$ preserves the two-dimensional space spanned by $\mid \xi\rangle$ and $\mid \psi\rangle$.

As a result of the inversion about the mean the probability amplitudes are redistributed. According to Grover, [183]: "after the inversion, the amplitude in each state increases/decreases; the amplitude is as much below/above $\bar{a}$ as it was above/below $\bar{a}$ before the inversion." If the probability amplitudes of all but one basis states are positive, after we apply the inversion about mean the probability amplitude of the negative basis state is amplified and becomes positive, Figure 34. Intuitively, the inversion about mean increases the amplitude of the target state by $\frac{2}{\sqrt{N}}$ at each iteration.

**Grover search algorithm**. Figure 35 illustrates the transformations for Grover quantum search algorithm: we apply an $n$-dimensional Walsh-Hadamard transform to create an equal superposition of the indices of all items in the database; then we perform $\sqrt{N}$ Grover iterations, $\mathbf{G}$. A Grover iteration $\mathbf{G}$ consists of the transformation performed by the oracle, $\mathbf{O}$, followed by an inversion about the mean, $\mathbf{D}_\psi$

$$\mathbf{G} = \mathbf{D}_k\mathbf{O} = (2 \mid \psi\rangle\langle\psi \mid -\mathbf{I})\mathbf{O}.$$

At each iteration the phase of the "solution(s)" is rotated with an angle of $\theta$ radians; the mean

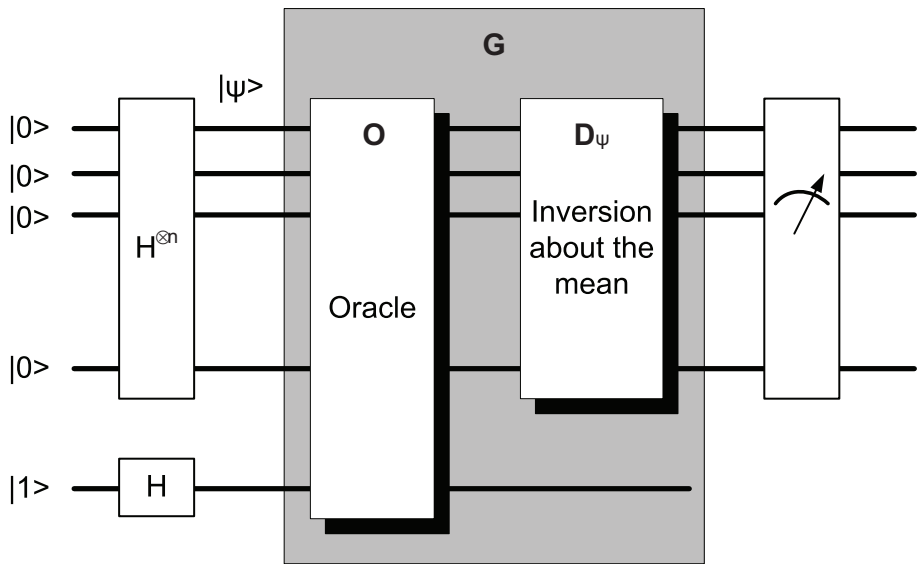



Figure 35: The schematic representation of the transformations for Grover quantum search algorithm. An $n$-dimensional Walsh-Hadamard transform creates $\mid \psi\rangle$, an equal superposition of the indices of all items in the database. Then we perform $\sqrt{N}$ Grover iterations, **G**. A Grover iteration **G** consists of the transformation performed by the oracle, **O**, followed by an inversion about the mean, $\mathbf{D}_\psi$. Finally, we measure the result.

value of the probability amplitudes of the superposition state is the same, but the amplitude of the " solution" is amplified.

One may question why do we need to carry out several iterations once the oracle is able to identify the solution. Recall that originally the input to the oracle is an equal superposition of $2^n$ basis vectors, each with a probability amplitude of $1/\sqrt{2^n}$. For example, if $n = 30$ then the probability amplitude of each of the $2^{30} \approx 10^9$ basis vectors in the Hilbert space $\mathcal{H}_{2^n}$ is very low, of the order $10^{-4}$. It is hard to distinguish the solution, which has a very small negative probability amplitude thus, we need to perform several iterations and amplify the amplitude of the solution.

**Geometric interpretation of one Grover iteration.** First, we review a property of the groups of transformations related to the symmetry operators: the product of two reflections in a two dimensional Euclidian plane is a rotation, see Section 1.9. Consider two lines $L1$ and $L2$ at an angle $\gamma$, Figure 36(a); the reflection $L2'$ of line $L2$ about line $L1$ is shown in Figure 36(b). Then in Figure 36(c) we see a reflection of $L2'$ about $L2$; the product of the two reflections is a rotation of $L2$ with an angle $2\gamma$.

We consider a slightly modified search problem and assume that there are $M \leq N$ items in the database $\mathcal{D}$ that satisfy the search criteria. The $N$ items in the database $\mathcal{D}$ are partitioned in two disjoint subsets: $\mathcal{S}$ the set of items that satisfy the search criteria and $\bar{\mathcal{S}}$ the set of items that do not satisfy the search criteria, Figure 37(a):

$$\mathcal{D} = \mathcal{S} \cup \bar{\mathcal{S}}, \quad \mid \mathcal{D} \mid = N, \quad \mid \mathcal{S} \mid = M, \quad \text{and} \quad \mid \bar{\mathcal{S}} \mid = N - M.$$

We define two states: $\mid \alpha\rangle$ - an equal superposition of the $N-M$ basis states $\mid x\rangle$ corresponding

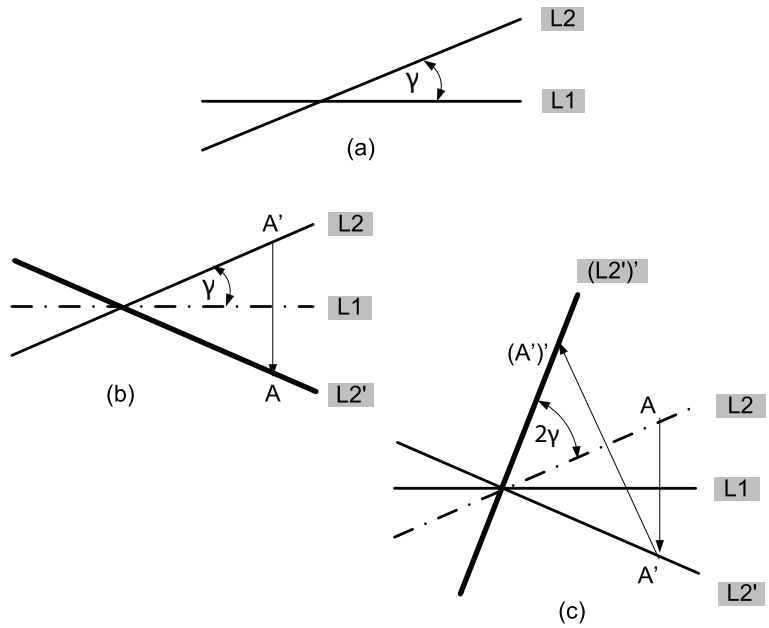



Figure 36: (a) Two lines $L1$ and $L2$ at an angle $\gamma$ in a two-dimensional Euclidean plane. (b) $L2'$ is the reflection of line $L2$ about $L1$. (c) $(L2')'$ is the reflection of line $(L2)'$ about $L2$. The two successive reflections correspond to a rotation with an angle $2\gamma$.

to indices $x$ of items that *do not satisfy the query,* $x \in \bar{\mathcal{S}}$; and $\mid \beta\rangle$ - an equal superposition of the $M$ basis states $\mid x\rangle$ corresponding to indices $x$ of items which *satisfy the query*, $x \in \mathcal{S}$

$$\mid \alpha\rangle = \frac{1}{\sqrt{N-M}} \sum_{x \in \bar{\mathcal{S}}} \mid x\rangle \quad \text{and} \quad \mid \beta\rangle = \frac{1}{\sqrt{M}} \sum_{x \in \mathcal{S}} \mid x\rangle.$$

The two states are orthogonal, $\langle \alpha \mid \beta\rangle = 0$; we can then express any state $\mid \psi\rangle$ as the superposition of the two states, Figure 37(b), as

$$\mid \psi\rangle = a \mid \alpha\rangle + b \mid \beta\rangle \quad \text{with} \quad \mid a \mid^2 + \mid b \mid^2 = 1.$$

The probability that a state $\mid x\rangle$ corresponding to the index $x$ of an item which does not satisfy the query is $(N-M)/N$; similarly, the probability that a state $\mid x\rangle$ corresponding to the index $x$ of an item which satisfies the query is $(M)/N$. Thus,

$$a = \sqrt{\frac{N-M}{N}} \quad \text{and} \quad b = \sqrt{\frac{M}{N}}.$$

In this representation $\mathbf{O}_M$, the phase inversion of the $M$ solutions to the query carried out by the oracle, transforms the original state $\mid \psi\rangle = a \mid \alpha\rangle + b \mid \beta\rangle$ to $\mathbf{O}_M \mid \psi\rangle = a \mid \alpha\rangle - b \mid \beta\rangle$. The phase inversion is represented geometrically as a reflection of the state vector $\mid \psi\rangle$ about $\mid \alpha\rangle$; the projection on $\mid \beta\rangle$ changes its sign, as seen in Figure 37(b). But $\mid \psi\rangle$ is an equal superposition state, therefore, the inversion about the mean $\mathbf{D}_\psi$ is a reflection of the state $\mathbf{O}_M \mid \psi\rangle$ about $\mid \psi\rangle$. The product of the two reflections is a rotation. To determine the rotation angle we denote the angle between $\mid \psi\rangle$ and $\mid \alpha\rangle$ as $\theta/2$; then,

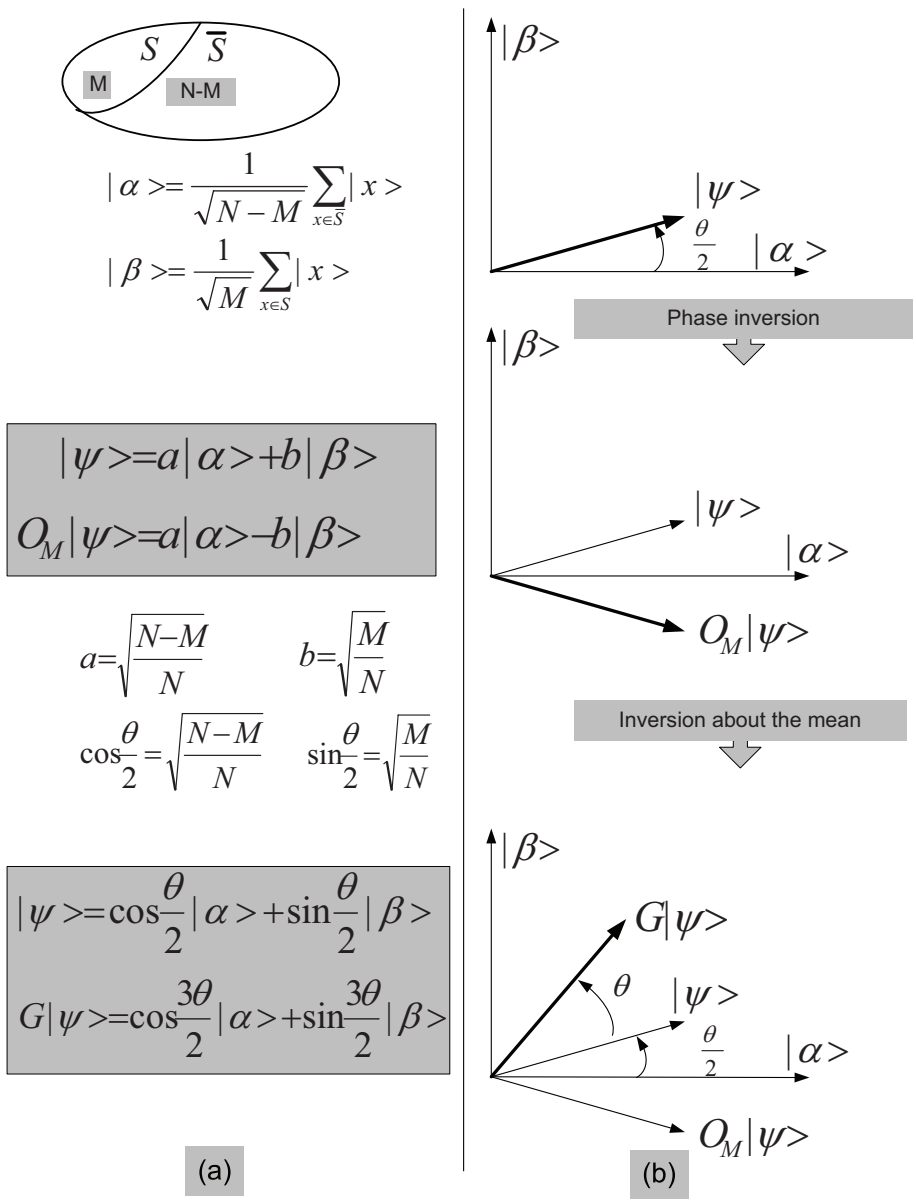



Figure 37: Geometric interpretation of one Grover iteration. (a) The $N$ items in the database $\mathcal{D}$ are partitioned in two disjoint subsets: $\bar{\mathcal{S}}$ includes items which do not satisfy the search criteria and $\mathcal{S}$ includes the ones which do. We define two normalized states: $\mid \alpha\rangle$, an equal superposition of the $N-M$ basis states $\mid x\rangle$ corresponding to indices $x \in \bar{\mathcal{S}}$ and $\mid \beta\rangle$, an equal superposition of the $M$ basis states $\mid x\rangle \in \mathcal{S}$. (b) Any superposition state can be represented as $\mid \psi\rangle = a \mid \alpha\rangle + b \mid \beta\rangle$. A phase inversion $\mathbf{O}_M$ produces the state $a \mid \alpha\rangle - b \mid \beta\rangle$ thus, it is a reflection of $\mid \psi\rangle$ about $\mid \alpha\rangle$. The inversion about the mean, $\mathbf{D}_\psi$, is a reflection of the state $\mathbf{O}_M \mid \psi\rangle$ about $\mid \psi\rangle$. The product of two reflections is a rotation; an iteration of Grover algorithm $G \mid \psi\rangle$ rotates the state $\mid \psi\rangle = \cos(\theta/2) \mid \alpha\rangle + \sin(\theta/2) \mid \beta\rangle$ by an angle $\theta$ towards $\mid \beta\rangle$, thus. towards the superposition of the states which satisfy the search criteria, thus $\mathbf{G} \mid \psi\rangle = \cos\frac{3\theta}{2} \mid \alpha\rangle + \sin\frac{3\theta}{2} \mid \beta\rangle$.

$$\cos\frac{\theta}{2} = \sqrt{\frac{N-M}{N}} \quad \text{and} \quad \sin\frac{\theta}{2} = \sqrt{\frac{M}{N}}.$$

We can express $\mid \psi\rangle$ as

$$\mid \psi\rangle = \cos\frac{\theta}{2} \mid \alpha\rangle + \sin\frac{\theta}{2} \mid \beta\rangle.$$

From Figure 37(b) we see that

$$\mathbf{G} \mid \psi\rangle = \mathbf{D}_\psi \mathbf{O}_M \mid \psi\rangle = \cos\frac{3\theta}{2} \mid \alpha\rangle + \sin\frac{3\theta}{2} \mid \beta\rangle.$$

We conclude that <u>one Grover iteration rotates the phase of the "solution" with $\theta$ radians.</u> After $k$ Grover iterations the state is

$$\mathbf{G}^k \mid \psi\rangle = \cos((2k+1)\theta/2) \mid \alpha\rangle + \sin((2k+1)\theta/2) \mid \beta\rangle.$$

Now we answer the question of how many iterations are needed when $M << N$; one iteration rotates the state vector with an angle $\theta$ thus, the number of iterations is

$$n_R \leq \lceil \frac{\pi/2}{\theta} \rceil.$$

When $M << N$, $\theta$ is a small angle and we approximate

$$\theta/2 \approx \sin(\theta/2) = \sqrt{\frac{M}{N}} \quad \Longrightarrow \quad \theta \approx 2\sqrt{\frac{M}{N}}.$$

It follows that

$$n_R \leq \lceil \frac{\pi/2}{2\sqrt{\frac{M}{N}}} \rceil = \lceil \pi/4\sqrt{\frac{N}{M}} \rceil.$$

We conclude that

$$n_R = \mathcal{O}\left(\sqrt{\frac{N}{M}}\right).$$

**Example.** Consider the case $n = 2$: then there are $N = 2^2 = 4$ items in the database with indices: $00, 01, 10$, and $11$; when $M = 1$, there is a unique solution. The quantum circuit in Figure 38 allows us to identify the solution in one iteration. Indeed, in this case the state after the first Walsh-Hadamard transform is

$$\mid \psi\rangle = \frac{\mid 00\rangle + \mid 01\rangle + \mid 10\rangle + \mid 11\rangle}{2}.$$

The angle $\theta/2$ of this state with $\mid \alpha\rangle$ is

$$\cos\theta/2 = \sqrt{\frac{N-M}{N}} = \frac{\sqrt{3}}{2} \quad \text{and} \quad \theta/2 = \frac{\pi}{6}.$$

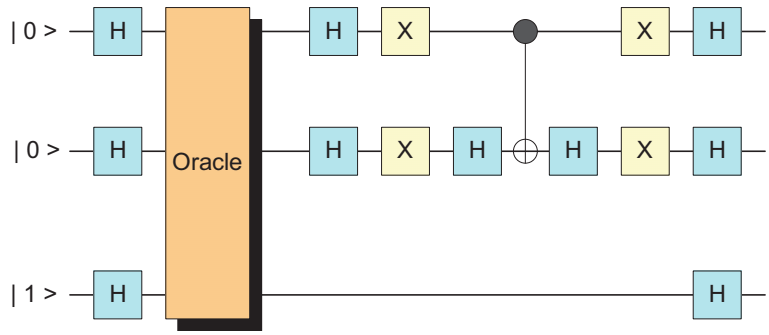



Figure 38: Two-qubit Grover circuit.

One Grover iteration rotates state $|\psi\rangle$ with an angle $\theta = \pi/3$. Thus, after one iteration the angle of the state vector is $\theta/2+\theta = \pi/2$, the state vector is along the solution $|\beta\rangle$. Figure 39 shows the four circuits of the oracle corresponding to solution $00, 01, 10$, and $11$, respectively.

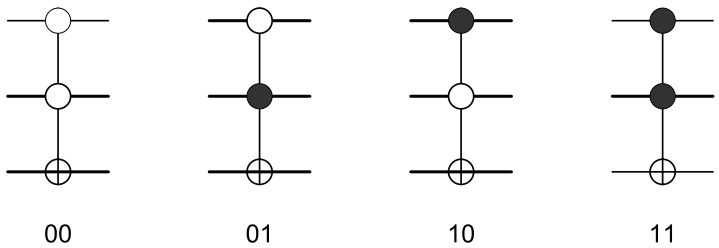



Figure 39: The circuits of the oracle for the Grover algorithm in Figure 38. The four circuits correspond to solutions $00, 01, 10$, and $11$, respectively.

**Stopping criteria.** The algorithm needs a precise stopping criterion, e.g., the precise number of target states. Without this stopping criterion the rotation of the current state could overshoot and drift away from the desired solution.

The sophisticated culinary analogy of this process attributed to Kristen Fuchs is described by Gilles Brassard [73]: "Quantum search is like cooking a souffle. You put the state obtained by quantum parallelism in a quantum oven and let the answer rise slowly. Success is almost guaranteed if you open the oven at the right time. But the souffle is likely to fall - the amplitude of the correct answer will go to zero - if you open the oven too early. Furthermore, the souffle could burn if you overcook it; strangely, the amplitude of the desired state shrinks after reaching its maximum."

Grover algorithm was generalized to the case in which the initial amplitudes are either real or complex and follow any arbitrary distribution [59, 60]. Recently, Grover showed that by replacing selective inversions with selective phase shifts of $\pi/3$ the algorithm preferentially converges to the target state irrespective of the step size or the number of iterations [190]. This strategy and the amplitude amplification are discussed in the next section.

## 1.24 Amplitude Amplification and Fixed-Point Quantum Search

We first emphasize the amplitude amplifications aspect of the quantum search, then discuss the general formulation of the amplitude amplification problem and, finally, we present the core ideas of the fixed point quantum search.

**Quantum search and amplitude amplification.** Grover quantum search algorithm can be summarized as follows: consider a quantum mechanical system with $N$ distinguishable states; all $N$ basis states have initially the same amplitude. We have $N$ objects, each identified by a unique label $x$. We map each label $x$ to one of the basis states of the system and iteratively amplify the amplitude of the basis state $\mid \tau\rangle$ corresponding to the solution $x = \tau$. The amplitude amplification is achieved by repeated state transformations as prescribed by the Walsh-Hadamard transform. Finally, we perform a measurement which collapses the system's state to a basis state. With high probability we end up observing the value $\tau$, the label of the object we are searching for, because the probability amplitude of the corresponding basis state, $\mid \tau\rangle$ has been amplified.

The centerpiece of the quantum search algorithm is a binary function $f(x)$ with the property that $f(x) = 1$ if and only if $x$ is the label of the object we are searching for, $f(\tau) = 1$. Recall that we can construct a quantum circuit that will evaluate any function that can be evaluated classically, thus we can construct a circuit to evaluate $f(x)$. If $b$ is an ancilla qubit we can always construct a circuit to carry out the following state transformation:

$$\mid x, b\rangle \mapsto \mid x, f(x)\ XOR\ b\rangle$$

If the ancilla, $b$, is initially in state $\mid b\rangle = 1/\sqrt{2}\,(\mid 0\rangle - \mid 1\rangle)$, then the circuit will invert the amplitude of all states when $f(x) = 1$.

Soon after Grover discovered the search algorithm, he realized that the Walsh-Hadamard transform can be replaced by almost any other transformation [186]. Let us follow Grover's arguments and notations. Assume that initially the system is in state $\mid \gamma\rangle$ and we wish to force it to state $\mid \tau\rangle$ through a sequence of unitary transformations $\mathbf{U}$. Call $U_{\tau\gamma}$ the amplitude of reaching the state $\mid \tau\rangle$ starting from the state $\mid \gamma\rangle$ after one application of the unitary transformation $\mathbf{U}$

$$U_{\tau\gamma} = \langle \tau \mid \mathbf{U} \mid \gamma\rangle.$$

The corresponding probability is $p_{\tau\gamma} = \mid U_{\tau\gamma} \mid^2$. In average we would need $\mathcal{O}(1/ \mid U_{\tau\gamma} \mid^2)$ trials (one trial is one application of $U$) for a success. Grover shows that in fact one can reach state $\mid \tau\rangle$ only in $\mathcal{O}(1/ \mid U_{\tau\gamma} \mid)$ trials. When $\mid U_{\tau\gamma} \mid << 1$ this results in a significant speedup.

Call $\mathbf{I}_x$ the transformation that inverts the amplitude of the basis state $\mid x\rangle$. The corresponding matrix, $I_x = [u_{ij}], 1 \leq i, j \leq N$, is a diagonal one with all diagonal terms equal to $+1$, except the element corresponding to $x$:

$$u_{ij} = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \neq x \\ -1 & \text{if} \quad i = j = x. \end{cases}$$

When $x$ corresponds to the basis vector $\mid x\rangle$ then the projector $\mathbf{P}_x = \mid x\rangle\langle x \mid$ is an $N \times N$ matrix whose only non-zero element is equal $+1$ and it is the diagonal element on row $x$ and

column $x$. We see that

$$I_x = I - 2 \mid x\rangle\langle x \mid$$

with $I$ the $N \times N$ identity matrix. For example for, $N = 4$ and $x = 3$:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mid 3\rangle\langle 3 \mid = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - 2 \times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad I_3 = I - 2 \mid 3\rangle\langle 3 \mid .$$

We define transformation $\mathbf{Q}$ as

$$\mathbf{Q} = -\mathbf{I}_\gamma \mathbf{U}^{-1} \mathbf{I}_\tau \mathbf{U}.$$

Consider the state obtained by applying the inverse of $\mathbf{U}$ to the target state

$$\mid \delta\rangle = \mathbf{U}^{-1} \mid \tau\rangle.$$

Now we apply the new transformation $\mathbf{Q}$ to the initial state $\mid \gamma\rangle$ and to the state $\mid \delta\rangle$. It is easy to show that:

$$\mathbf{Q} \mid \gamma\rangle = (1 - 4 \mid U_{\tau\gamma} \mid^2) \mid \gamma\rangle + 2U_{\tau\gamma} \mid \delta\rangle$$

Indeed,

$$\begin{aligned}
\mathbf{Q} \mid \gamma\rangle &= (-\mathbf{I}_\gamma \mathbf{U}^{-1} \mathbf{I}_\tau \mathbf{U}) \mid \gamma\rangle \\
&= [-(\mathbf{I} - 2 \mid \gamma\rangle\langle\gamma \mid)\mathbf{U}^{-1}(\mathbf{I} - \mid 2\tau\rangle\langle\tau \mid)\mathbf{U}] \mid \gamma\rangle \\
&= -\mathbf{I}\mathbf{U}\mathbf{U}^{-1} \mid \gamma\rangle - \mathbf{I}\mathbf{U}^{-1} 2 \mid \tau\rangle\langle\tau \mid \mathbf{U} \mid \gamma\rangle + 2 \mid \gamma\rangle\langle\gamma \mid \mathbf{U}^{-1}\mathbf{I}\mathbf{U} \mid \gamma\rangle \\
&\quad -4 \mid \gamma\rangle\langle\gamma \mid \mathbf{U}^{-1} \mid \tau\rangle\langle\tau \mid \mathbf{U} \mid \gamma\rangle \\
&= - \mid \gamma\rangle - 2\mathbf{I}\mathbf{U}^{-1} \mid \tau\rangle U_{\tau\gamma} + 2 \mid \gamma\rangle - 4 \mid \gamma\rangle U^*_{\tau\gamma} U_{\tau\gamma} \\
&= + \mid \gamma\rangle - 4 \mid U^*_{\tau\gamma} \mid^2 \mid \gamma\rangle + 2U_{\tau\gamma}(\mathbf{U}^{-1} \mid \tau\rangle) \\
&= (1 - 4 \mid U^*_{\tau\gamma} \mid^2) \mid \gamma\rangle + 2U_{\tau\gamma} \mid \delta\rangle.
\end{aligned}$$

It is left as an exercise for the reader to prove that

$$\mathbf{Q}(\mid \delta\rangle) = -2U^*_{\tau\gamma} \mid \gamma\rangle + \mid \delta\rangle.$$

The two equations can be combined as

$$\mathbf{Q} \begin{bmatrix} \mid \gamma\rangle \\ \mid \delta\rangle \end{bmatrix} = \begin{bmatrix} (1 - 4 \mid U_{\tau\gamma} \mid^2) & 2U_{\tau\gamma} \\ -2U^*_{\tau\gamma} & 1 \end{bmatrix} + \begin{bmatrix} \mid \gamma\rangle \\ \mid \delta\rangle \end{bmatrix}$$

Thus, *the transformation* $\mathbf{Q}$ *preserves the two-dimensional vector space spanned by* $\mid \gamma\rangle$ *and* $\mid \delta\rangle$, and each application of it rotates a vector in this space by approximately $2 \mid U_{\tau\gamma} \mid$ radians. The case of interest is when $\mid \gamma\rangle$ and $\mid \delta\rangle$ are almost orthogonal. Each application of $\mathbf{Q}$ rotates the system state by an angle $2 \mid U_{\tau\gamma} \mid$ thus the initial state $\mid \gamma\rangle$ is transformed to the final state $\mid \tau\rangle$ after a number of applications of $\mathbf{Q}$ equal to

$$n_Q = \frac{\frac{\pi}{2}}{2 \mid U_{\tau\gamma} \mid} = \frac{\pi}{4 \mid U_{\tau\gamma} \mid}.$$

Grover shows that the result extends to the case when the amplitudes in the states $\mid \gamma\rangle$ and $\mid \tau\rangle$ are rotated by arbitrary phases instead of being inverted, in other words when $\mathbf{I}_\gamma$ and $\mathbf{I}_\tau$ are replaced by arbitrary rotations.

This result shows that quantum search is a robust algorithm and that the Walsh-Hadamard transform on $n$ qubits, $\mathbf{W} = \mathbf{H}^{\otimes n}$, can in fact be replaced by almost any other unitary transformation.

If we wish to perform an exhaustive search starting from the zero state $\mid 0\rangle$ and use the Walsh-Hadamard transform on $n$ qubits as the unitary transformation, $\mathbf{U} = \mathbf{W} = \mathbf{W}^{-1}$ then $\mathbf{Q}$ becomes

$$\mathbf{Q} = -\mathbf{U}_0 \mathbf{W} \mathbf{I}_\tau \mathbf{W}.$$

Note that $-\mathbf{W}\mathbf{I}_0\mathbf{W}$ is in fact the inversion-about-zero operation described in Section 1.23. Since $\mathbf{I}_0 = \mathbf{I} - 2 \mid 0\rangle\langle 0 \mid$ it follows that:

$$-\mathbf{W}\mathbf{I}_0\mathbf{W} \mid x\rangle = -\mathbf{W}(\mathbf{I} - 2 \mid 0\rangle\langle 0 \mid)\mathbf{W} \mid x\rangle = - \mid x\rangle + 2\mathbf{W} \mid 0\rangle\langle 0 \mid \mathbf{W} \mid x\rangle.$$

If $\bar{X} = 1/N \sum_{i=1}^{N} x_i$ then we see that the $i$-th component of $-\mathbf{W}\mathbf{I}_0\mathbf{W}$ can be expressed as

$$-x_i + 2\bar{X} = \bar{X} + (\bar{X} - x_i).$$

**Amplitude amplification.** In 2000, Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Trapp generalized the amplitude amplification idea [215]. The new process, allows us to find a "Good" element $x \in X$ after an expected number of applications of the unitary transformation $\mathbf{U}$ and of its inverse, $\mathbf{U}^{-1}$; the number of iterations is proportional to $1/\sqrt{a}$ with $a$ the probability of getting a "Good" element $x$ if $\mathbf{U} \mid 0\rangle$ is measured.

Let $\mathbf{U}$ be a unitary operator in a Hilbert space, $\mathcal{H}_N$ with an orthonormal basis

$$X = \{\mid 0\rangle, \mid 1\rangle, \ldots, \mid N-1\rangle\}.$$

The only condition imposed on $\mathbf{U}$ is to be invertible, thus $\mathbf{U}$ must not involve any measurements.

If $\chi : X \mapsto \{0, 1\}$ is a Boolean function we say that the basis state $\mid x\rangle$ is a "Good" state if $\chi(x) = 1$ and $\mid x\rangle$ is a "Bad" state" if $\chi(x) = 0$. The Boolean function $\chi$ partitions the Hilbert space $\mathcal{H}_N$ into two sub-subspaces

$$\mathcal{H}_N = \mathcal{H}_{Good} \cup \mathcal{H}_{Bad}, \quad \mathcal{H}_{Good} \cap \mathcal{H}_{Bad} = \emptyset$$

such that

$$\begin{array}{lll} \mathcal{H}_{Good} \mapsto & \text{subspace spanned by the set of "Good" basis states} \mid x\rangle \text{ such that} & \chi(x) = 1 \\ \mathcal{H}_{Bad} \mapsto & \text{subspace spanned by the set of "Bad" basis states} \mid x\rangle \text{ such that} & \chi(x) = 0. \end{array}$$

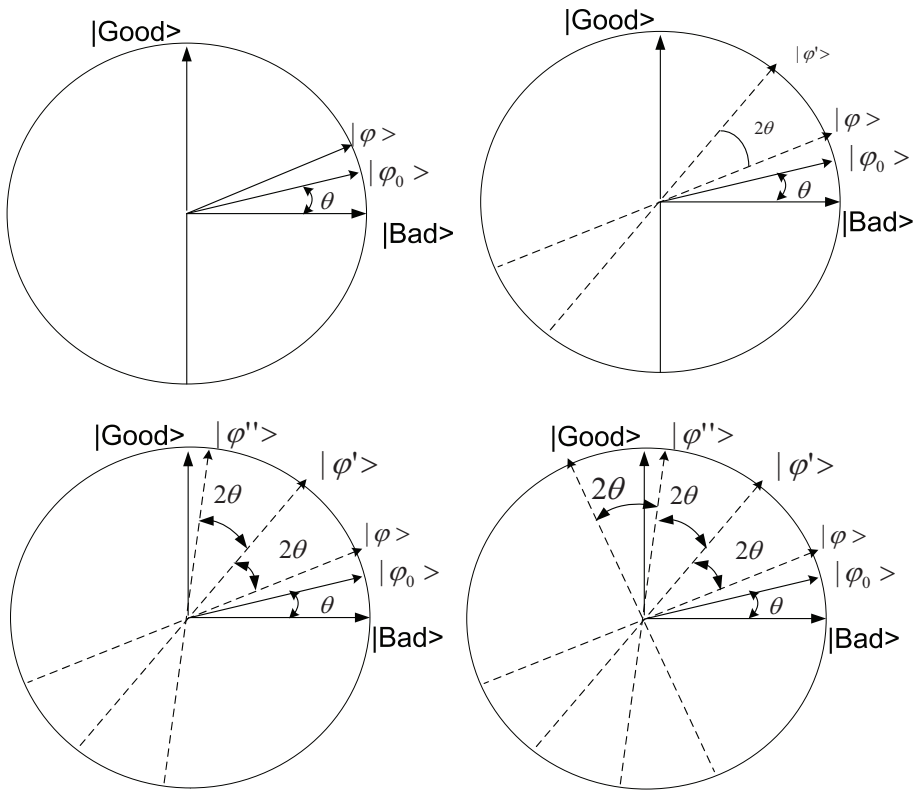



Figure 40: The effect of amplitude amplification; the original state $\mid \varphi\rangle$ is rotated towards the $\mid$ Good$\rangle$ state and may overshoot it.

Every pure state $\mid \gamma\rangle \in \mathcal{H}_N$ has a unique decomposition as a projection onto $\mathcal{H}_{Good}$ and $\mathcal{H}_{Bad}$

$$\mid \gamma\rangle = \mid \gamma_{Good}\rangle + \mid \gamma_{Bad}\rangle.$$

The amplitude amplification is based on an operator $\mathbf{Q}$ defined as

$$\mathbf{Q} = \mathbf{Q}(U, \chi, \phi, \varphi) = -\mathbf{U}\mathbf{S}_0(\phi)\mathbf{U}^{-1}\mathbf{S}_\chi(\varphi).$$

with $\phi$ and $\varphi$ two angles such that $0 \leq \phi, \varphi \leq \pi$ and $\mathbf{S}_\chi$ an operator which conditionally changes the amplitudes of "Good" states

$$\mathbf{S}_\chi(\varphi) \mid x\rangle \mapsto \begin{cases} e^{i\varphi} \mid x\rangle & \text{if} \quad \chi(x) = 1 \\ \mid x\rangle & \text{if} \quad \chi(x) = 0. \end{cases}$$

Similarly, $\mathbf{S}_0$ amplifies the amplitude by a factor $e^{i\phi}$ if the state is $\mid 0\rangle$

$$\mathbf{S}_0(\phi) \mid 0\rangle \mapsto \begin{cases} e^{i\phi} \mid 0\rangle & \text{if} \quad \chi(0) = 1 \\ \mid 0\rangle & \text{if} \quad \chi(0) = 0. \end{cases}$$

Let us apply the transformation $\mathbf{U}$ to a system in state $\mid 0\rangle$

$$\mid \Phi\rangle = \mathbf{U} \mid 0\rangle = \mid \Phi_{Good}\rangle + \mid \Phi_{Bad}\rangle$$

and consider the subspace $H_\Phi$ spanned by $\mid \Phi_{Good}\rangle$ and $\mid \Phi_{Bad}\rangle$. As before, $a$ denotes the probability of success

$$a = \langle \Phi_{Good} \mid \Phi_{Good}\rangle.$$

We also define the angle $\theta \leq \frac{\pi}{2}$ such that

$$\sin^2(\theta) = a.$$

**Proposition.** *The Hilbert subspace $\mathcal{H}_\Phi$ is stable under the transformation $\mathbf{Q}$. If $0 \leq a \leq 1$ then $\mathcal{H}_\Phi$ has dimension 2, otherwise it has dimension 1.*

This proposition is similar with the one we have discussed earlier and it implies the following equations

$$\begin{array}{rcrcr} \mathbf{Q} \mid \Phi_{Good}\rangle & = & (1-2a) \mid \Phi_{Good}\rangle & - & 2a \mid \Phi_{Bad}\rangle \\ \mathbf{Q} \mid \Phi_{Bad}\rangle & = & 2(1-a) \mid \Phi_{Good}\rangle & + & (1-2a) \mid \Phi_{Bad}\rangle \end{array}$$

$\mathbf{Q}$ is a unitary operator thus when $0 \leq a \leq 1$ then $\mathcal{H}_\Phi$ has an orthonormal basis consisting of the two eigenvectors of $\mathbf{Q}$, $\mid \Phi_\pm\rangle$

$$\mid \Phi_\pm\rangle = \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{a}} \mid \Phi_{Good}\rangle \pm \frac{i}{\sqrt{1-a}} \mid \Phi_{Bad}\rangle \right)$$

with the corresponding eigenvalues

$$\lambda_\pm = \exp(\pm i 2\theta).$$

We can express $\mid \Phi_{Good}\rangle$ and $\mid \Phi_{Bad}\rangle$ in terms of the eigenvectors

$$\mid \Phi_{Good}\rangle = \sqrt{\frac{a}{2}} \left( \mid \Phi_+\rangle + \mid \Phi_-\rangle \right)$$

and

$$\mid \Phi_{Bad}\rangle = -i\sqrt{\frac{1-a}{2}} \left( \mid \Phi_+\rangle - \mid \Phi_-\rangle \right).$$

Thus

$$\mathbf{U} \mid 0\rangle = \mid \Phi\rangle = \mid \Phi_{Good}\rangle + \mid \Phi_{Bad}\rangle = \frac{-i}{\sqrt{2}} \left[ \exp(i\theta) \mid \Phi_+\rangle - \exp(-i\theta) \mid \Phi_-\rangle \right].$$

It follows that after the $j$-th application of the transformation $\mathbf{Q}$ the state of the system is

$$\mathbf{Q}^{(j)} \mid \Phi\rangle = \frac{-i}{\sqrt{2}} \left(\exp[i\theta(2j+1)] \mid \Phi_+\rangle - \exp[-i\theta(2j+1)] \mid \Phi_-\rangle\right),$$

or

$$\mathbf{Q}^{(j)} \mid \Phi\rangle = \frac{1}{\sqrt{a}} \sin[(2j+1)\theta] \mid \Phi_{Good}\rangle + \frac{1}{\sqrt{1-a}} \cos[(2j+1)\theta] \mid \Phi_{Bad}\rangle.$$

This shows that after we apply $m$ times, with $m \geq 0$, the transformation $\mathbf{Q}$, a measurement will produce a "Good" state with probability $p_{Good} = \sin[(2m+1)\theta]^2$. The following proposition follows immediately.

**Proposition.** *Let* $\mathbf{U}$ *be any unitary transformation which does not perform any measurements and has an initial probability of success equal to* $0 \leq a \leq 1$ *and let* $\sin^2(\theta) = a$, $0 \leq \theta \leq \frac{\pi}{2}$. *Let* $\chi : \mathbb{Z} \mapsto \{0,1\}$ *be any Boolean function. Set* $m = \lceil \frac{\pi}{4\theta} \rceil$. *After* $m$ *applications of* $\mathbf{Q}$, *the resulting state,* $\mathbf{Q}^{(m)}\mathbf{U} \mid 0\rangle$, *is a "Good" state with probability at least* $\max(1-a,\ a)$.

We note that Grover algorithm is a particular instance of amplitude amplification when:

- the oracle implements the Boolean function $f = \chi$,
- the transformation $\mathbf{U}$ is the Walsh-Hadamard transform $\mathbf{W} = \mathbf{H}^{\otimes n}$ on $n$ qubits, and
- the Grover iteration $-\mathbf{W}\mathbf{S}_0\mathbf{W}\mathbf{S}_f$ corresponds to the operator $\mathbf{Q} = -\mathbf{U}\mathbf{S}_0\mathbf{U}^{-1}\mathbf{S}_\chi$.

This iteration carried out by transformation $\mathbf{Q}$ can be regarded as a *rotation* in the two-dimensional space spanned by the starting vector $\mid \psi_0\rangle = \mathbf{U} \mid 0\rangle = \sum_{x \in X} \alpha_x \mid x\rangle$ and the state consisting of a uniform superposition of solutions to the search problem. The initial state may be expressed as:

$$\mid \psi_0\rangle = \frac{1}{\sqrt{2}} \left(\sqrt{a} \mid Good\rangle + \sqrt{1-a} \mid Bad\rangle\right).$$

**Fixed-point quantum search.** The "classical" quantum search algorithm does not guarantee that the state vector does not over-rotate beyond the solution to the search problem, or, in terms of the culinary analogy discussed in Section 1.23, that the souffle is not overcooked. In his 2005 paper [190] Lov Grover shows that by replacing the selective amplitude inversion with selective phase shifts of $\theta = \frac{\pi}{3}$, the algorithm converges to the desired target state irrespective of the amount of rotation of the state vector at each iteration and the number of iterations. The new algorithm is thus extremely robust. Figure 41 inspired by [190] summarizes the differences between fixed point quantum search and amplitude amplification.

The system is initially in state $\mid \gamma\rangle$ and we wish it to end in state $\mid \tau\rangle$; when we apply the unitary operator $\mathbf{U}$ the state vector always moves closer to the target state.

We define the transformation $\mathbf{Q}$ as

$$\mathbf{Q}_{\tau\gamma} = \mathbf{U}\mathbf{R}_\gamma\mathbf{U}^\dagger\mathbf{R}_\tau U.$$

In this expression $\mathbf{R}_\gamma$ and $\mathbf{R}_\tau$ are selective phase shifts by $\theta = \frac{\pi}{3}$ of the states $\mid \gamma\rangle$ and $\mid \tau\rangle$, respectively

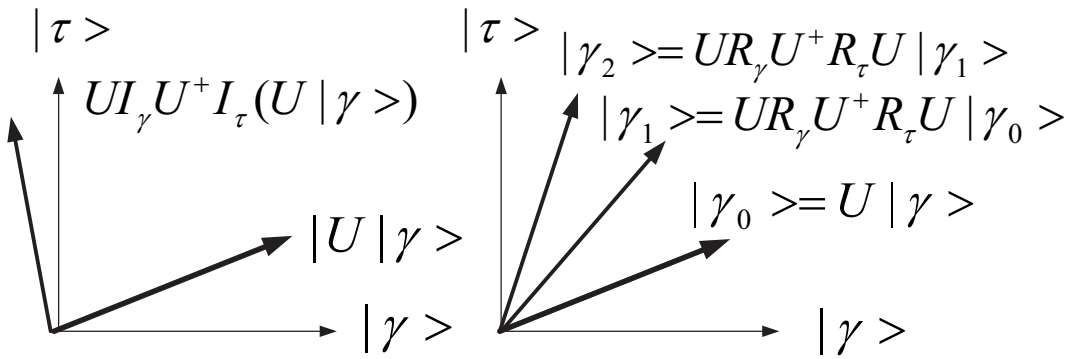



Figure 41: (Left) In case of amplitude amplification one application of the operator $\mathbf{U}\mathbf{I}_\gamma\mathbf{U}^\dagger\mathbf{I}_\tau\mathbf{U}$ could overshoot the desired target state $\mid \tau\rangle$. (Right) The fixed point quantum search guarantees that $\mathbf{U}\mathbf{R}_\gamma\mathbf{U}^\dagger\mathbf{R}_\tau\mathbf{U} \mid \gamma\rangle$ always moves towards the target state $\mid \tau\rangle$.

$$\begin{aligned}\mathbf{R}_\gamma &= \mathbf{I} - [1-\exp(i\theta)] \mid \gamma\rangle\langle\gamma \mid, \\ \mathbf{R}_\tau &= \mathbf{I} - [1-\exp(i\theta)] \mid \tau\rangle\langle\tau \mid .\end{aligned}$$

**Proposition.** *If when we apply the unitary transformation* $\mathbf{U}$ *the probability of reaching the state* $\mid \tau\rangle$ *from* $\mid \gamma\rangle$ *is* $p^{(U)}_{\tau\gamma} = \mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2 = (1-\epsilon)$ *then, when we apply the transformation* $\mathbf{Q}_{\tau\gamma}$ *to the system in state* $\mid \gamma\rangle$, $\mathbf{Q}_{\tau\gamma} \mid \gamma\rangle = \mathbf{U}\mathbf{R}_\gamma\mathbf{U}^\dagger\mathbf{R}_\tau\mathbf{U} \mid \gamma\rangle$ *the probability of reaching the state* $\mid \tau\rangle$ *becomes* $(1-\epsilon^3)$ *thus,*

$$\mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2 = (1-\epsilon) \quad \Longrightarrow \quad \mid\mid \mathbf{Q}_{\tau\gamma} \mid\mid^2 = \mid\mid \mathbf{U}\mathbf{R}_\gamma\mathbf{U}^\dagger\mathbf{R}_\tau\mathbf{U} \mid\mid^2 = (1-\epsilon^3).$$

The proof of this proposition follows from a one-by-one application of the successive transformations implied by $\mathbf{Q}_{\tau\gamma}$. First, we compute $\mathbf{U} \mid \gamma\rangle$, then we apply to the resulting state the transformation $\mathbf{R}_\tau = I - (1-\exp(i\theta)) \mid \tau\rangle\langle\tau \mid$, and so on. We use the fact that $\mathbf{U}_{\tau\gamma} = \langle\tau \mid \mathbf{U} \mid \gamma\rangle$, $\mathbf{U}^*_{\tau\gamma} = \langle\gamma \mid U^{-1} \mid \tau\rangle$, and $\mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2 = \mathbf{U}_{\tau\gamma}\mathbf{U}^*_{\tau\gamma}$. It follows that

$$\begin{aligned}\mathbf{Q}_{\tau\gamma} \mid \gamma\rangle &= \mathbf{U}\mathbf{R}_\gamma\mathbf{U}^\dagger\mathbf{R}_\tau\mathbf{U} \mid \gamma\rangle \\ &= [\exp(i\theta) + \mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2 (\exp(i\theta)-1)^2]\mathbf{U} \mid \gamma\rangle + \mathbf{U}_{\tau\gamma}(\exp(i\theta)-1) \mid \tau\rangle.\end{aligned}$$

Since $\mathbf{U} \mid \gamma\rangle$ deviates from the desired target state with probability $\epsilon = 1 - \mid\mid U_{\tau\gamma} \mid\mid^2$ it follows that the probability that this superposition deviates from the desired target state $\mid \tau\rangle$ is

$$\begin{aligned}q^{(Q)}_{\tau\gamma} &= (1 - \mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2) \mid\mid [\exp(i\theta) + \mid\mid \mathbf{U}_{\tau\gamma} \mid\mid^2 [\exp(i\theta)-1]^2] \mid\mid^2 \\ &= \epsilon \mid\mid \exp(i\theta) + (1-\epsilon)[\exp(i\theta)-1]^2] \mid\mid^2 \\ &= \epsilon^3.\end{aligned}$$

Thus, the probability of reaching the state $\mid \tau\rangle$ after one application of $\mathbf{Q}_{\tau\gamma}$ is

$$p^{(Q)}_{\tau\gamma} = 1 - \epsilon^3.$$

Figure 41 (b) shows that the successive applications of the fixed-point quantum search operator moves the state closer to the target state. The fixed-point quantum search transformation can be defined recursively:

$$
\begin{aligned}
\mathbf{Q}_{\tau\gamma}^{(0)} &= \mathbf{U} \\
\mathbf{Q}_{\tau\gamma}^{(1)} &= \mathbf{Q}_{\tau\gamma}^{(0)}\mathbf{R}_{\gamma}(\mathbf{Q}_{\tau\gamma}^{(0)})^{\dagger}\mathbf{R}_{\tau}\mathbf{Q}_{\tau\gamma}^{(0)} = \mathbf{U}\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U} \\
\mathbf{Q}_{\tau\gamma}^{(2)} &= \mathbf{Q}_{\tau\gamma}^{(1)}\mathbf{R}_{\gamma}(\mathbf{Q}_{\tau\gamma}^{(1)})^{\dagger}\mathbf{R}_{\tau}\mathbf{Q}_{\tau\gamma}^{(1)} \\
&= \mathbf{U}(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U}) \\
\mathbf{Q}_{\tau\gamma}^{(3)} &= \mathbf{Q}_{\tau\gamma}^{(2)}\mathbf{R}_{\gamma}(\mathbf{Q}_{\tau\gamma}^{(2)})^{\dagger}\mathbf{R}_{\tau}\mathbf{Q}_{\tau\gamma}^{(2)} \\
&= \mathbf{U}(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U}) \\
&\quad (\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U}) \\
&\quad (\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U}) \\
&\quad (\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U})(\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}^{\dagger}\mathbf{U})(\mathbf{R}_{\gamma}^{\dagger}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U}) \\
&\quad (\mathbf{R}_{\gamma}\mathbf{U}^{\dagger}\mathbf{R}_{\tau}\mathbf{U}) \\
\mathbf{Q}_{\tau\gamma}^{(4)} &= \ldots \\
\vdots &
\end{aligned}
$$

As we can see the recursion relation is not a simple one, we cannot derive a closed form expression for the $m$-th iteration.

We end this introductory chapter with a brief discussion of error models and quantum algorithms.

## 1.25 Error Models and Quantum Algorithms

Quantum errors affect the processing, storage, and transmission of quantum information. In this section we discuss models for errors affecting quantum computations described by quantum algorithms. The primary goal of a quantum algorithm is to produce a desired quantum state starting from a classical input state such as $\mid 0\rangle$. To obtain the result of a computation we perform a measurement of some or of all the qubits affecting the output state.

A quantum algorithms can be described using quantum networks. A *quantum network* is a space-time diagram of the operations applied by the algorithm to each qubit. The computational state of an algorithm is maintained with arbitrary precision if and only if the error of each quantum gate of the quantum network implementing the algorithm is below a certain threshold. Only if this condition is satisfied one could rely on the results of a quantum computation.

Without loss of generality we assume that errors occur at *given quantum network locations* and we consider two classes of errors:

1. *Operational errors.* They may occur during each gate operation and may affect all the qubits the gate is operating on.

2. *Memory errors.* A quantum network is partitioned into qubit time units determined by the maximum execution time of a gate. A memory error occurs if there is a loss of information during a qubit time unit. There is an interesting relationship between parallelism and memory errors. If $t$ is the longest interval of time a qubit can be

in storage without any significant loss of information and $q$ is the number of qubits actively involved in a computation thus, the number of gates executing in parallel, then the minimum number of operations per unit of time must be considerably larger than $q/t$.

To construct computational quantum error models we consider three aspects: (i) the type of error operators that may occur at a given location; (ii) the nature of the mixture of error operators at a given location; and (iii) wether the errors at different locations are independent or not. We make several simplifying assumptions:

- The operators used to transform a qubit are affected by a small number of errors; the error operators lead to either, no error, bit-flip, phase-flip, or bit and phase-flip.
- Each error occurs at a given location. The actual behavior of a quantum network can be represented as a sum of networks with linear error operators at each network location. This *error expansion* is not unique; it gives the correct input-output behavior of the network, yet is not capable of accurately representing the intermediate states of a qubit. The final state of the computation is obtained by adding the final state of each of the networks.
- *Leakage errors*, manifestations of loss of amplitude in the two-dimensional Hilbert space, are non-existent. This last assumption is unrealistic for some quantum systems, e.g., in an ion trap the amplitude of a level may be lost to other levels which in turn are storing information for different operations [244, 247].

The error analysis is simpler if one places stochastically and independently a standard error at an error location. The stochastic assumption means the state of the environment associated with each element in the error expansion be orthogonal. We present several quantum computational error models:

*Independent stochastic errors.* At each error location one makes an independent random choice of either the identity operator or a stochastic combination of linear operators satisfying the unitarity assumption. The probability of an error at that location is equal to the probability of the stochastic combination.

*Independent errors.* The error expansion is obtained by assigning a quantum operation to each error location. The quantum operation is expressed as a set of linear operators labelled by the states of the environment

$$\mid e_0\rangle I + \sum_i \mid e_i\rangle A_i$$

with $\mathrm{tr}\,[A_i] = 0,\ \forall i \geq 1$. The strength of the error associated with this quantum operation is

$$\mid \sum_i \mid e_i\rangle A_i \mid = \sup \mid \sum_i \mid e_i\rangle A_i \mid \psi\rangle \mid .$$

We cannot talk about the probability of error only about the error strength of the quantum operation.

*Quasi-independent stochastic errors.* In this model an error occurs with probability $p$ and the error expansion is obtained as a stochastic sum such that the probability of all terms of the sum which have a non-identity quantum operation at a given set of $k$ error locations is at most $p^k$.

*Quasi-independent errors.* Each summand of an error expansion is associated with a set of failed error locations and all other error locations are associated with an identity operator.

The following proposition [244] relates quantum algorithms using perfect operations to algorithm with imperfect operations.

**Proposition.** *There exists a constant $\delta$ such that for every $\epsilon > 0$ a quantum algorithm using perfect operations can be converted to an equivalent quantum algorithm with imperfect operations, each with an error at most $\delta$, such that the final error is at most $\epsilon$. The overhead of the converted algorithm is polylogarithmic in $\epsilon^{-1}$ and the number of computational steps of the algorithm.*

## 1.26 History Notes

Our knowledge about the physical world continually evolves simulated by new discoveries. In the seventeenth century Johannes Kepler formulated the laws of planetary motion which were empirically adequate but did not explain why the planets move the way they do. Sir Isaac Newton provided an adequate justification for Kepler's laws based on the three laws of physics he had formulated by 1666 and his law of gravitational forces; yet, Newton's mechanics could not answer why the actions at a distance obey an inverse square law. This question could be answered only after Albert Einstein and David Hilbert developed the general relativity theory, during the second decade of the twentieth century[11].

The history of science is ripe with similar examples: the Dutch mathematician Hendrik Antoon Lorentz showed that Maxwell's equations for the electromagnetic field are invariant to the *Lorentz transformations* he had proposed in 1904, but are not invariant under Galilean transformation of Newtonian mechanics; Lorentz's transformations are rather counterintuitive, predict length contraction and time dilation, while Galilean transformation are more intuitive and easier to comprehend. Only after Einstein postulated that the laws of physics are the same in all inertial frames of reference and that the speed of light is constant independent of the source did it become clear that Lorentz transformations are anchored in physical reality.

This brings up fundamental questions about our knowledge: How do we develop theories and models which are better than others? How do we distinguish true knowledge from inadequate knowledge? Is Quantum Mechanics an accurate description of the physical world? Is it the ultimate description of the physical world?

The discipline of philosophy which studies knowledge is called *epistemology*. If we examine the history of epistemology we notice a shift from the theories which stressed the absolute, the

[11] On 25 November 1915 Einstein submitted his paper *The Field Equations of Gravitation* which give the correct field equations for general relativity. Five days earlier Hilbert had submitted a paper *The Foundations of Physics* which contained the correct field equations for gravitation. Hilbert's paper contains some important contributions to relativity not found in Einstein's work.

permanent character of knowledge to theories which emphasize the relativity, the situation-dependence, the evolution of knowledge and its active interference with the physical world.

The Greek philosopher Plato believed that knowledge is absolute and reflects a set of universal "Ideas" or "Forms." Aristotle, his disciple, emphasized the role of logic, gathering knowledge through rational reflection, and empiricism which seeks knowledge through sensory perception. Aristotle's ideas were greatly admired during and immediately after the Renaissance, when empiricism and rationalism were fashionable. During the second half of the nineteenth century, Immanuel Kant introduced the human mind as an active originator of experience, rather than just a passive recipient of perception. Kant believed that knowledge results from organization of perceptual data on the basis of inborn "Categories" such as space, time, and causality. These a priori categories are static, even though the basic concepts are subjective. Kant's philosophy attempted to provide philosophical grounds for Newton's mechanics by showing that classical Physics is entirely compatible with transcendental conditions for objective knowledge. Niels Bohr was strongly influenced by Kant's philosophy.

The twentieth century was dominated by different flavors of *pragmatism*, e.g., logical positivism, and the "Copenhagen Interpretation" of Quantum Mechanics. The common theme of pragmatism is that knowledge consists of "models" which allow us to approach "problem-solving" in the simplest possible fashion. It is implicitly assumed that models abstract properties of the entities they describe and cannot, and should not capture all properties of these entities. Thus, multiple parallel models, possibly contradictory, may exist at any instant of time, and we should always choose the model which helps us solve the problem at hand. The ultimate truth, the reality behind these models, the so called "Ding an Sich[12]," is not only unattainable, but also meaningless.

While matter and energy preoccupied the minds of philosophers starting with Leucippus and Democritus several hundred years before our era and later preoccupied the minds of many generations of natural scientists, information per se became a subject of serious investigation only after the significant technological developments in communication in the late 1940s.

There is little wonder that information is not a central concept in Quantum Mechanics, or that information theory, as developed by Claude Shannon, is not concerned with the behavior of quantum systems capable of carrying information. The milestones that mark the inception of the information age happened in the second half of the twentieth century: the transistor was invented by William Shockley, John Bardeen and Walter Brattain, just before Christmas in 1947; the first commercial computer, UNIVAC I became operational in 1951; the DNA double helix structure was discovered by Sir Francis Harry Compton Crick and James Dewey Watson in 1953; the first microprocessor, 4004, was produced by Intel in 1971.

It is fascinating to follow the evolution of the ideas leading to today's quest for processing quantum information using quantum mechanical devices and to see the tremendous progress in quantum information processing we have witnessed since the early 1980's. The last two decades mirror the exciting decades from the beginning of the twentieth century when the basic concepts of Quantum Mechanics were developed.

The stage was set by the introduction of the quantum theory in 1900 by Max Planck which triggered a chain of scientific discoveries (Figure 42): in 1905, Albert Einstein developed the theory of the photoelectric effect; in 1911, Sir Ernest Rutherford presented the planetary model of the atom; in 1913, Niels Bohr introduced the quantum model of the hydrogen atom;

[12] "Ding an Sich" is a German expression; its ad litteram translation is "the thing in itself."

in 1923, Prince Louis de Broglie related the momentum $p$ of a particle with the wavelength $\lambda$ of the wave associated with the particle, $p = \lambda/h$, with Planck's constant $h = 6.626 \times 10^{-34}$ Joule.second

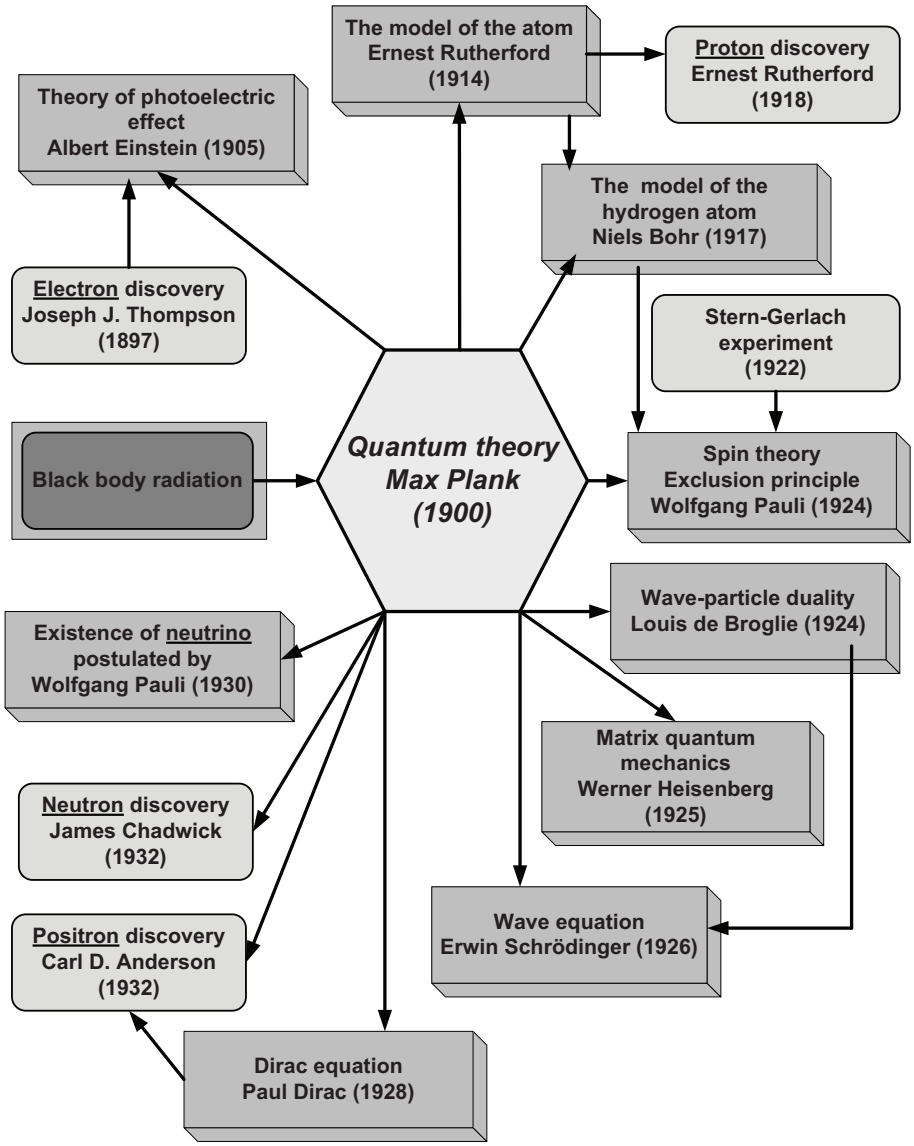



Figure 42: The chain of great discoveries in quantum physics during the *golden years* 1900 - 1932. Theoretical models were developed to explain the results of the experiments. In turn, the theoretical models predicted new properties and phenomena and suggested new experiments to establish the validity of their predictions.

Quantum Mechanics has its origins in the 1925 "matrix Quantum Mechanics" of Heisenberg [207]. Later that year, Max Born and Pascual Jordan used infinite matrices to represent basic physical quantities and developed a complete formalism for Quantum Mechanics. Inspired by de Broglie's wave-particle duality, Erwin Schrödinger introduced in 1926 the equation for the dynamics of the wave function [367]; the same year Schrödinger and Paul Dirac

showed the equivalence of Heisenberg's matrix formulation with Schrödinger's wave function.

The study of information and of the relation of information with Physics followed a sinuous path: in 1929, Leó Szilárd pioneered the study of the physics of information while analyzing the Maxwell Demon [415]; in 1961, Rolf Landauer showed that the erasure of information is a dissipative process and provided a quantitative characterization of the process, known as Landauer's principle [265].

The study of computability was a response to David Hilbert's Entscheidungsproblem formulated in 1930s; Hilbert asked if there was a mechanical procedure for separating mathematical truths from mathematical falsehoods. While studying this problem Alonzo Church and Stephen Kleene introduced the $\lambda$-definable functions in 1936. Church proved that "every effectively calculable function (effectively decidable predicate) is general recursive" [96]. Alan Turing introduced the concept of Turing Machine and proved that every function which could be regarded as computable is computable by a Turing Machine [429]. The Church-Turing principle is generally formulated as: "every function which can be regarded as computable can be computed by a universal computing machine." In layman's terms a universal computer is a single machine that can perform any physically possible computation.

The theoretical developments in mid 1930s were followed by the practical construction of mechanical devices for information processing and the formulation of basic principles for their construction. The first classical computers were designed and built in early 1940s; John von Neumann introduced the concept of "stored-program computer" and the "von Neumann computer architecture" in 1946 [80].

As faster and faster computers were built more attention was paid to energy dissipation. Theoretical physicists argued that if there is a minimum amount of energy required to perform an elementary logical operation, then the faster a computer becomes, the more power is needed for the operation of the computer, and the harder is to remove the heat generated during the computation. In 1973, Charles Bennett showed that any computation can be carried out using only reversible steps thus, in principle a computation can be carried out without dissipating any power [30]. A reversible Turing Machine performs only reversible computations as follows: it runs forward all the steps prescribed by the algorithm, outputs the result, and, finally, reverses all the steps to return to its initial state; at the end of this process no energy is dissipated.

The idea of using quantum effects for processing information followed a more direct path; in 1982, Richard Feynman envisioned the idea of a quantum computer, a physical device which takes advantage of the distinct properties of quantum systems to process information; Feynman conjectured that only a quantum computer will be able to carry out an "exact simulation" of a physical system [150]. In 1982, Paul Benioff recognized that the time evolution of an isolated quantum system described by the Hamiltonian is a reversible process and can mimic a reversible computation; he developed the idea of a quantum Turing Machine [28, 29].

In 1985, David Deutsch introduced the concept of "quantum parallelism," and gave a concrete example of a computation that showed the distinction between classical and quantum computers [115]. In 1994, Peter Shor developed an algorithm for factoring large numbers [380] and generated a wave of excitement for the newly founded discipline of quantum computing. Two years later, in 1996, Lov Grover discovered a quantum search algorithm for an unsorted database with $N$ elements.

Alexei Kitaev [241] introduced the topological quantum computing model in 1997 and in

2000 Michael Freedman, Alexei Kitaev, and Zhenghan Wang, showed that the topological quantum computing has a computing power similar to that of the quantum circuit model [159]. Preskill's class notes cover a rigorous presentation of topological quantum computing [345]; an elementary introduction to the subject can be found in [103].

In 1995, David DiVincenzo formulated the requirements for practical implementations of a quantum information processing system [127]. Deutsch, Barenco, and Ekert observed in 1995 that: "all computer programs my be regarded as symbolic representations of some of the laws of Physics, specialized to apply to specific processes, therefore the limits of computability coincide with the limits of science itself. If the laws of Physics did not support computational universality, they would be decreeing their own un-knowability" [118].

In the early 1990s, some 40 years after Shannon's development of information theory, Charles Bennett, Gilles Brassard, Richard Jozsa, Asher Peres, Benjamin Schumacher, Peter Shor, William Wootters and others introduced the basic concepts of quantum information theory [33, 34, 35, 37, 38, 39, 40, 42, 44, 46, 47, 48]. Quantum information theory had to reconcile two different views regarding nondeterminism, the one that is the basic tenet of Quantum Mechanics and the one at the foundation of information theory. The nondeterminism of Quantum Mechanics reflects our *inability to know precisely the state of atomic or subatomic particles,* a very provoking thought contested bitterly by some, including Albert Einstein. Information theory and classical Physics take a different view, *probabilities reflect lack of knowledge.* George Boole categorically declares that: "probability is expectation founded on partial knowledge. A perfect acquaintance with all the circumstances affecting the occurrence of an event would change expectation in certainty, and leave neither room nor demand for a theory of probabilities" [65].

## 1.27 Summary

A theme reverberating throughout the entire book is that *information is physical* and that the study of information and Physics are deeply intertwined. Classical and quantum information are anchored in physical reality, they are "engraved" onto some property of the physical system used to transmit, store, and transform information. Classical information we are familiar with is subject to the laws of classical Physics while quantum information obeys the laws of Quantum Mechanics.

The special properties of quantum information such as superposition and entanglement, the inability to clone quantum states, the fact that a measurement projects the quantum state pose considerable challenges to quantum information processing and, at the same time, open intriguing possibilities for communication, storage, and computing with quantum information. For example, we can detect with high probability the presence of an intruder on a quantum communication channel, but at the same time, we have to measure the state of a quantum register to detect errors without altering its state.

Quantum information is transformed using quantum gates, the building blocks for quantum circuits which, in turn, can be assembled to build quantum computing and communication devices. A quantum bit, a *qubit*, is a quantum system with two distinguishable states; the state of a qubit is represented by a vector in a two-dimensional Hilbert space, $\mathcal{H}_2$. A quantum computer is a physical device designed to transform quantum information embodied by the state of a quantum system.

The physical processes required to transform the quantum state in a controlled manner are different for ion traps, solid-state, optical, NMR, or other possible physical implementations. The requirements for the physical implementation of quantum devices come naturally to mind, they have an immediate correspondent for classical computers: we cannot conceive a state of the art computer built with circuits whose state cannot be controlled or be initialed to a desired state; it would be impractical to build a computer unless we have a finite set of building blocks; and it seems obvious that we should have access to the results of a computation.

An abstract computation could be viewed as the creation of symbols, the "output," which encode information according to some preexisting conventions in a systematic manner and have abstract properties specified by other symbols, the "input" of the process. Communication is a process whose output symbols are affected not only by the properties specified by the input but also by factors that are not entirely under our control. The "symbols" are physical objects thus, subject to the laws of physics.

An oracle is an abstraction for a black box that can follow a very elaborate procedure to answer a very complex question with a "yes" or "no" answer. Quantum computing has a natural affinity for oracles as the result of a quantum computation is probabilistic and it is a superposition of all possible results; increasingly more complex oracles are the ones proposed Deutsch, Deutsch-Jozsa, Bernstein-Vazirani, and Simon.

Quantum algorithms, start from an initial state and then cause a set of state transformations of the quantum computer which eventually lead to the desired result. Indeed, the first step for any quantum mechanical computation is to initialize the system to a state that we can easily prepare; then we carry out a sequence of unitary transformations that cause the system to evolve towards a state which provides the answer to the computational problem. A quantum operation is a rotation of the state $\mid \psi\rangle$ in N-dimensional Hilbert space. Thus, the ultimate challenge is to build up powerful N-dimensional rotations as sequences of one and two dimensional rotations. For any quantum algorithm there are multiple paths leading from the initial to the final state and there is a degree of interference among these paths. The amplitude of the final state thus, the probability of reaching the desired final state, depends on the interference among these paths. This justifies the common belief that quantum algorithms are very sensitive to perturbations and one has to be extremely careful when choosing the transformations the quantum mechanical system is subjected to.

Grover introduced a quantum algorithm for searching an unsorted database containing $N$ items in a time of order $\sqrt{N}$ while on a classical computer the search requires a time of order $N$. The speedup of Grover algorithm is achieved by exploiting both quantum parallelism and the fact that, according to quantum theory, a probability is the square of an amplitude. No classical or quantum algorithm can solve this problem faster than time of order $\sqrt{N}$. Grover search algorithm can be applied directly to a wide range of problems.

## 1.28 Exercises and Problems

**Problem 1.** Let $A$ and $B$ be two linear operators. Prove that:

$$\mathrm{tr}\left(AA^{\dagger}\right)\mathrm{tr}\left(BB^{\dagger}\right) \geq \mid \mathrm{tr}\left(AB^{\dagger}\right) \mid^{2} .$$

This is the Schwarz inequality for the operator inner product.

*Hint:* Assume $A = [a_{ij}]$ and $B = [b_{ij}]$, $1 \leq i \leq n$, $1 \leq j \leq m$. Then:

$$\text{tr}\left(AA^\dagger\right) = (\mid a_{11} \mid^2 + \ldots + \mid a_{1m} \mid^2) + (\mid a_{21} \mid^2 + \ldots + \mid a_{2m} \mid^2) + \ldots + (\mid a_{n1} \mid^2 + \ldots + \mid a_{nm} \mid^2)$$

and

$$\text{tr}\left(AB^\dagger\right) = (a_{11}b_{11}^* + \ldots + a_{1m}b_{1m}^*) + (a_{21}b_{21}^* + \ldots + a_{2m}b_{2m}^*) + \ldots + (a_{n1}b_{n1}^* + \ldots + a_{nm}b_{nm}^*).$$

Use the Cauchy-Schwartz inequality: if $(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{C}$ then:

$$\left(\mid \alpha_1 \mid^2 + \mid \alpha_2 \mid^2 + \ldots \mid \alpha_n \mid^2\right)\left(\mid \beta_1 \mid^2 + \mid \beta_2 \mid^2 + \ldots + \mid \beta_n \mid^2\right) \geq \mid \alpha_1\beta_1^* + \alpha_2\beta_2^* + \ldots + \alpha_n\beta_n^* \mid^2 .$$

**Problem 2.** Show that $\mathbb{C}^{m\times n}$, the set of all matrices $A = [a_{ij}]$ with elements $a_{ij} \in \mathbb{C}$ is a vector space. The addition of two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ is defined as $A + B = [a_{ij} + b_{ij}]$, the inverse of $A = [a_{ij}]$ is $-A = [-a_{ij}]$ and the identity element is $E = [0]$.

**Problem 3.** Show that the set of $n \times n$ square matrices with elements from $\mathbb{C}$ form an inner product vector space. The inner product of two matrices is defined as

$$\langle A, B\rangle = \text{tr}(A^*B).$$

with $A^*$ the complex conjugate of the matrix $A$.

**Problem 4.** Show that opposite points on the Bloch sphere correspond to orthogonal qubit states.

**Problem 5.** Consider a point $A$ on the Bloch sphere with coordinates $x_A, y_A, z_A$ and the vector connecting the origin of the sphere with $A$. Show that a rotation with an angle $\theta$ around this vector is described by:

$$\mathcal{R}_A(\theta) = \cos\left[\frac{\theta}{2}\right]\sigma_I + i\sin\left[\frac{\theta}{2}\right](x_A\sigma_x + y_A\sigma_y + z_A\sigma_z)$$

**Problem 6.** Show that $\mid 0\rangle, \mid 1\rangle, \mid 2\rangle, \mid 3\rangle$ form an orthonormal basis in $\mathcal{H}_4$. Construct the operator $\Pi_4$ that permutes circularly the basis in $\mathcal{H}_4$ as follows $\mid 0\rangle \rightarrow \mid 1\rangle$, $\mid 1\rangle \rightarrow \mid 2\rangle$, $\mid 2\rangle \rightarrow \mid 3\rangle$ and $\mid 3\rangle \rightarrow \mid 0\rangle$. Construct its matrix representation. Calculate $\Pi_4 \mid \psi\rangle$ and $\langle\psi \mid \Pi_4$ with $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle + \alpha_2 \mid 2\rangle + \alpha_3 \mid 3\rangle$. *Hint:* Express the canonical basis states as $\mid 00\rangle, \mid 01\rangle, \mid 10\rangle, \mid 11\rangle$.

**Problem 7.** The controlled-phase, `CPHASE` gate in Figure 9 transforms the two input qubits in state $\mid \varphi_1\rangle$ and $\mid \varphi_2\rangle$ as follows:

$$\mid \varphi_1\varphi_2\rangle \mapsto (-1)^{\varphi_1\varphi_2} \mid \varphi_1\varphi_2\rangle.$$

Construct the matrix describing this transformation. Show that this gate is in fact a controlled-Z gate which flips the phase of the target qubit when the control qubit is set; show that the role of the target and control qubits could be reversed.

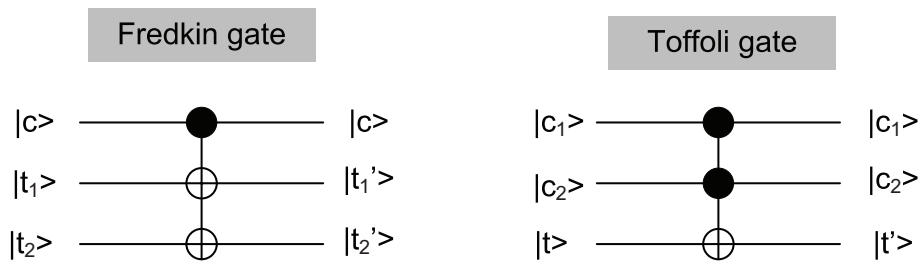



Figure 43: The Fredkin gate operates as follows: the control qubit $\mid c\rangle$ is transferred to the output unchanged; when $\mid c\rangle = \mid 0\rangle$ then the target qubits $\mid t_1\rangle$ and $\mid t_2\rangle$ are transferred to the output unchanged; when $\mid c\rangle = \mid 1\rangle$ then the two target qubits $\mid t_1\rangle$ and $\mid t_2\rangle$ are swapped. The Toffoli gate operates as follows: the control qubits $\mid c_1\rangle$ and $\mid c_2\rangle$ are transferred to the output unchanged; when $\mid c_1\rangle = \mid c_2\rangle = \mid 0\rangle$ then the target qubit is transferred to the output unchanged; when $\mid c_1\rangle = \mid c_2\rangle = \mid 1\rangle$ then the target qubit is flipped.

**Problem 8.** The Fredkin and Toffoli gates in Figure 43 operate on three qubits.

Show that the unitary transformations carried out by the Fredkin and Toffoli gates are respectively:

$$U_{Fredkin} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad U_{Toffoli} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

*Hint:* The Fredkin gate transforms the basis vectors in $\mathcal{H}_8$ as follows:

$\mid 000\rangle \mapsto \mid 000\rangle$, $\mid 001\rangle \mapsto \mid 001\rangle$, $\mid 010\rangle \mapsto \mid 010\rangle$, $\mid 011\rangle \mapsto \mid 011\rangle$,
$\mid 100\rangle \mapsto \mid 100\rangle$, $\mid 101\rangle \mapsto \mid 110\rangle$, $\mid 110\rangle \mapsto \mid 101\rangle$, $\mid 111\rangle \mapsto \mid 111\rangle$.

The Toffoli gate transforms the basis vectors in $\mathcal{H}_8$ as follows:

$\mid 000\rangle \mapsto \mid 000\rangle$, $\mid 001\rangle \mapsto \mid 001\rangle$, $\mid 010\rangle \mapsto \mid 010\rangle$, $\mid 011\rangle \mapsto \mid 011\rangle$,
$\mid 100\rangle \mapsto \mid 100\rangle$, $\mid 101\rangle \mapsto \mid 101\rangle$, $\mid 110\rangle \mapsto \mid 111\rangle$, $\mid 111\rangle \mapsto \mid 110\rangle$.

**Problem 9.** Show that classical Fredkin and Toffoli gates allow <u>fanout</u>, in other words can replicate one of their inputs at two different outputs, while the quantum Fredkin and Toffoli gates do not violate the <u>no cloning theorem</u>. *Hint:* Consider the case when the input of the two quantum gates are in the basis states $\mid 0\rangle$ and $\mid 1\rangle$ as well as the case when the input is in a superposition state, e.g., $1/\sqrt{2}(\mid 0\rangle + \mid 1\rangle)$.

**Problem 10.** Provide an intuitive justification of the reason why a quantum circuit must have the same number of input and output qubits and relate your answer to the reversibility of quantum transformations. *Hint:* The transformation $\mathbf{U}$ carried out by the circuit is unitary, $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$.

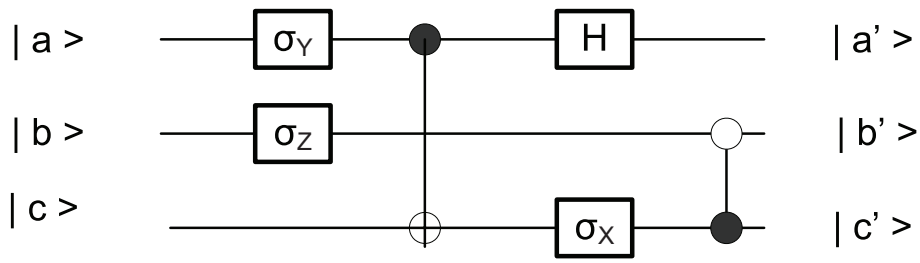



Figure 44: $\sigma_x, \sigma_y, \sigma_z$, are transformation given by Pauli matrices, $H$ is the Hadamard transform.

**Problem 11.** Construct the transfer matrix of the quantum circuit in Figure 44:

**Problem 12.** If $\sigma_I, \sigma_x, \sigma_y$, and $\sigma_z$ are Pauli matrices and $\rho$ is the density matrix of a qubit show that:

$$\frac{1}{2}\sigma_I = \frac{1}{4}\left(\rho + \sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z\right).$$

**Problem 13.** Construct a quantum circuit able to solve the Deutsch-Jozsa problem.

**Problem 14.** By replacing selective inversions by selective phase shifts of $\pi/3$, Grover algorithm preferentially converges to the target state irrespective of the step size, or the number of iterations [190]. Discuss the new algorithm and explain its novel features.

**Problem 15.** Show that if $H(n)$ is a Hadamard matrix then so is

$$H(2n) = \begin{pmatrix} H(n) & H(n) \\ H(n) & -H(n) \end{pmatrix}.$$

*Hint:* show that

$$H(2n)H(2n)^\dagger = 2nI$$

with $I$ the $n \times n$ identity matrix.

**Problem 16.**

Consider the transformation $\mathbf{Q}$ defined in Section 1.24 as

$$\mathbf{Q} = -\mathbf{I}_\gamma \mathbf{U}^{-1}\mathbf{I}_\tau \mathbf{U}.$$

Initially the system is in state $\mid \gamma\rangle$ and we wish to force it to state $\mid \tau\rangle$ through a sequence of unitary transformations $\mathbf{U}$. We apply the new transformation $\mathbf{Q}$ to the initial state $\mid \gamma\rangle$ and to the state $\mid \delta\rangle = \mathbf{U}^{-1} \mid \tau\rangle$, obtained by applying the inverse of $\mathbf{U}$ to the target state.

Show that

$$\mathbf{Q}(\mid \gamma\rangle) = \mid \gamma\rangle(1 - 4 \mid U_{\tau\gamma} \mid^2) + 2U_{\tau\gamma}(\mathbf{U}^{-1} \mid \tau\rangle)$$

and

$$\mathbf{Q}(\mathbf{U}^{-1} \mid \tau\rangle) = \mathbf{U}^{-1} \mid \tau\rangle - 2U^*_{\tau\gamma} \mid \gamma\rangle.$$

Here $U_{\tau\gamma}$ is the amplitude of reaching the state $\tau$ starting from the state $\gamma$ after one application of the unitary transformation $U$:

$$U_{\tau\gamma} = \langle \tau \mid \mathbf{U} \mid \gamma \rangle.$$

*Hint:* Use the fact that $\mathbf{U}$ is unitary $\mathbf{U}\mathbf{U}^{-1} = \mathbf{I}$. Recall also that

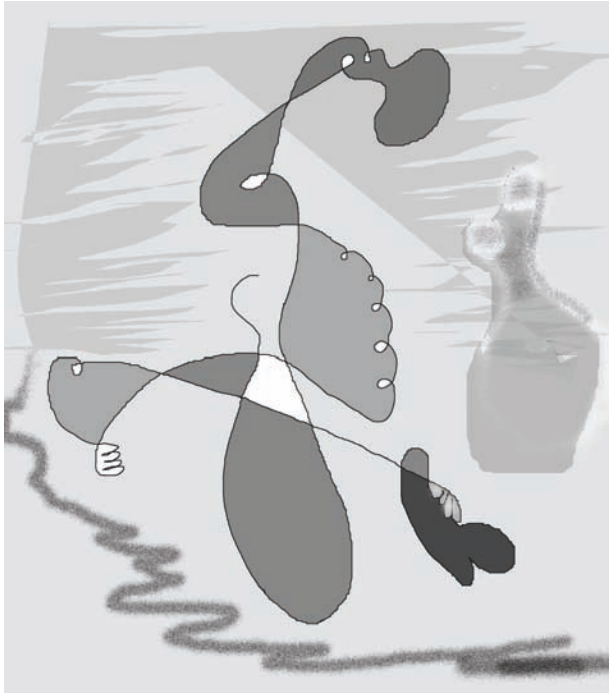
$$\langle \gamma \mid \gamma \rangle = 1 \quad \text{and} \quad \langle \tau \mid \tau \rangle = 1,$$

and that

$$\mathbf{U}^*_{\tau\gamma} = (\langle \tau \mid \mathbf{U} \mid \gamma \rangle)^\dagger = \langle \gamma \mid \mathbf{U}^{-1} \mid \tau \rangle.$$

We have also shown that $\mathbf{I}_\gamma = \mathbf{I} - \mid \gamma \rangle \langle \gamma \mid$ and $\mathbf{I}_\tau = I - \mid \tau \rangle \langle \tau \mid$.

"Not everything that can be counted counts, not everything that counts can be counted."
Albert Einstein.



# 2 Measurements and Quantum Information

Quantum information processing requires a deeper understanding of the measurement process and its effects on a quantum system; measurements allow us to observe the outcome of transformations carried out by quantum circuits and quantum communication channels. This justifies the important role measurements play in quantum information theory and the extended coverage of measurements of quantum systems in this chapter.

Recall that a quantum measurement is an irreversible process which transforms quantum information to classical information and that there is a strong coupling between the system being measured and the measuring system. The measurements postulate tells us that the outcome of a measurement of an observable $A$ of a quantum system is an eigenvalue of the self-adjoint operator $\mathbf{A}$ associated with the measurement of the observable. Operator $\mathbf{A}$ is a *projector* for observable $A$ of a quantum system. The quantum state immediately after the measurement is an eigenvector of the operator $\mathbf{A}$. We say that a quantum measurement *collapses the state* of a quantum system.

We have seen in Chapter 1 that *we cannot simultaneously measure observables associated with non-commuting operators;* if two operators **A** and **B** do not commute, then measuring observable $A$ using projector **A** will interfere with the eigenvectors of projector **B** and vice versa. Indeed, only two commuting matrices can be simultaneously diagonalized thus, we can measure the eigenvalues of one of them without affecting the eigenvectors of the other.

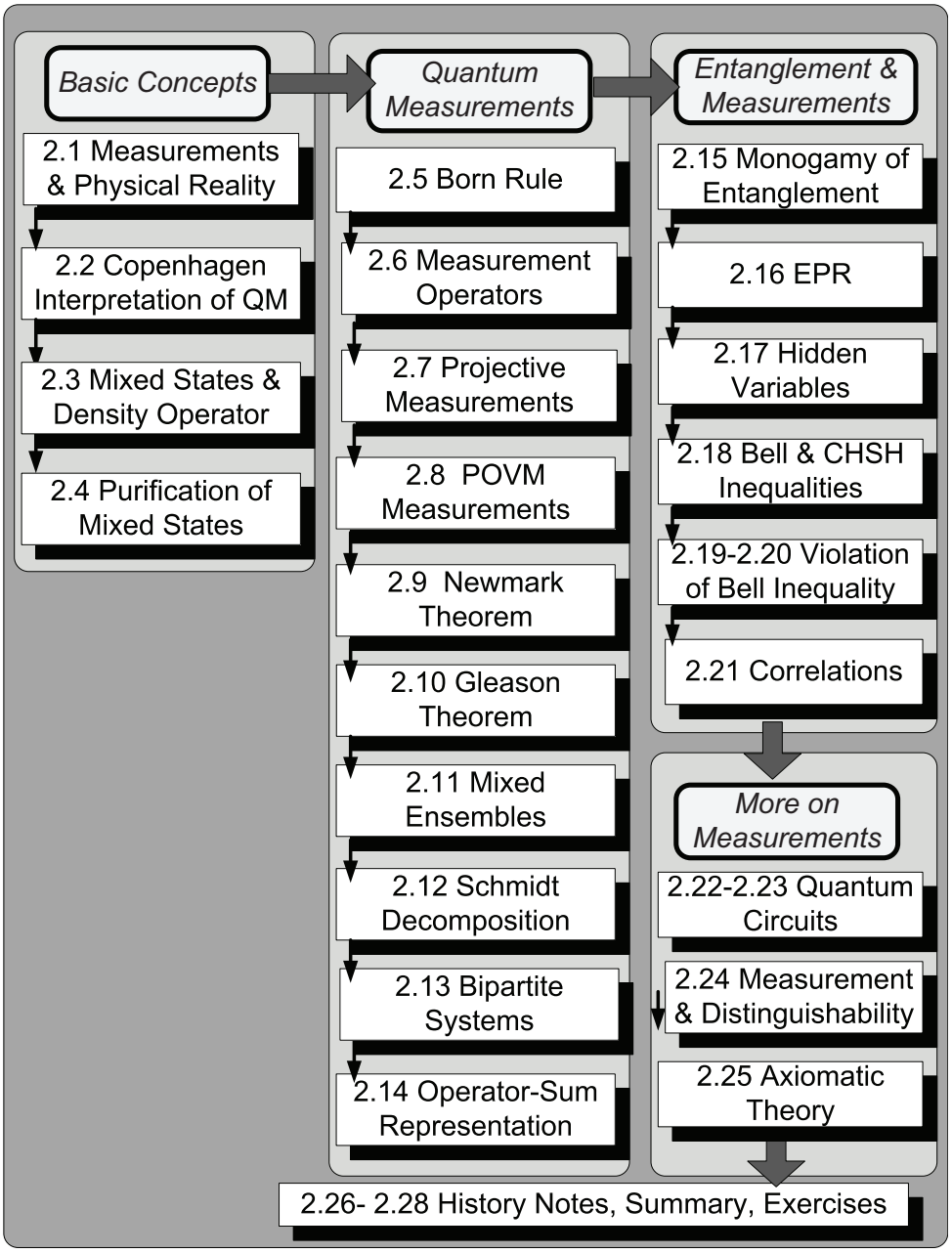



Figure 45: The organization of Chapter 2 at a glance.

Figure 45 summarizes the organization of the chapter; after the introduction of basic concepts we analyze quantum measurements, then discuss the entanglement and applications.

## 2.1 Measurements and Physical Reality

A measurement is the process of acquiring information about our environment, about physical objects, phenomena, or any entities using a measuring apparatus and our senses; this information eventually evolves into knowledge and allows us to construct models of physical systems. Measurements are an important step in the process of scientific discovery; any experiment to assess the correctness of a theoretical model of the physical world requires measurements.

The relationship between our perceptions and the "physical reality" surrounding us is central to the theory of knowledge and requires answers to questions such: Is there a clear separation between the measuring instrument and the system being measured? Is a measurement objective? Are we able to make "ideal measurements" and learn the "true state" of a system? Classical and Quantum Physics do not give similar answers to these questions.

In Classical Physics there is an unambiguous distinction between a system and the measuring instrument. Also, a description of the system developed as a result of measurements does not depend on a measurement of the system in the same state at a later point in time; thus, we have reasons to believe that the description of a classical system is objective.

From the inception of Quantum Mechanics the concept of "measurement" and the relationship between the outcome of the measurement and our knowledge of the state of the quantum system prior, during, and after the measurement proved to be intensely controversial. In 1934, Niels Bohr made a stunning observation regarding the measurement of quantum systems: "a subsequent measurement to a certain degree deprives the information given by a previous measurement of its significance for predicting the future course of phenomena. Obviously, these facts not only set a limit to the extent of the information obtainable by measurement, but they also set a limit to the meaning which we may attribute to such information. We meet here in a new light the old truth that in our description of nature the purpose is not to disclose the real essence of the phenomena, but only to track down so far as possible, relations between the manifold aspects of our experience."

The question if a numerical value of a physical quantity has any meaning until an observation has been performed was answered differently by the representatives of the metaphysical and realist schools of thought. The Copenhagen Interpretation of Quantum Mechanics of Niels Bohr and Werner Heisenberg was contested by many physicists including Albert Einstein. John von Neumann provided an axiomatic definition of a quantum measurement process; he treats the measurement apparatus as a quantum object and argues that a measurement consists of a linear stage characterized by the joint state of the quantum system and the measuring instrument followed by a non-linear stage, when this joint state collapses.

In 1958 Werner Heisenberg gave a crisp description of the relationship between our knowledge and the physical reality surrounding us: "...The laws of nature which we formulate mathematically in quantum theory deal no longer with the particles themselves but with our knowledge of the elementary particles. ... The conception of objective reality ... evaporated into the ... mathematics represents no longer the behavior of elementary particles but rather our knowledge of this behavior...."

Several views regarding the relationship between the measurement process and the information obtained as a result of a measurement are possible. The common wisdom is that we create new information as a result of a measurement. Another view is that we only transform information from one form to another and that *information is conserved* during the mea-

surement process. For example, when we use an instrument to measure the intensity of the electric current, we in fact transform microscopic information provided by the state of the individual electrons and the settings of the electronic apparatus, used to measure the electric current, to another type of information, a real number and a unit, e.g., 3.5 Ampere[13].

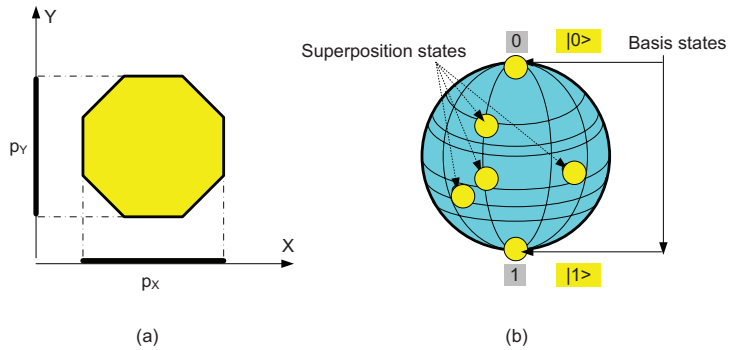



Figure 46: (a) Information is lost when we project an object; we cannot recognize the shape of a 2-D object from a 1-D projection. (b) A quantum measurement projects quantum information on a subspace of the original Hilbert space and replaces quantum with classical information; the measurement of a qubit in a superposition state $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ projects it on one of the basis states, $\mid 0\rangle$ or $\mid 1\rangle$ and reveals classical information, 0 or 1.

Projecting an object leads to a loss of information, as we can see in Figure 46(a)[14]; the measurement postulate tells us that the state of a quantum system is altered as a result of a measurement, Figure 46(b). A naive solution to the problem of measuring a quantum system without altering the state, would be to clone a quantum state and then measure the copy and leave the original unperturbed. Unfortunately, state replication, one of the favorite methods to increase the reliability of classical systems, is not possible for quantum systems. The "no-cloning" theorem, a consequence of the linearity of Quantum Mechanics, states that a transformation of an arbitrary state $\mid \psi\rangle$ of the type

$$\mid \psi\rangle \quad \mapsto \quad \mid \psi\rangle \otimes \mid \psi\rangle.$$

is not possible. If cloning of quantum states were possible, then we would be able to clone two arbitrary states $\mid \psi\rangle$ and $\mid \varphi\rangle$ as

$$\mid \psi\rangle \quad \mapsto \quad \mid \psi\rangle \otimes \mid \psi\rangle \qquad \text{and} \qquad \mid \varphi\rangle \quad \mapsto \quad \mid \varphi\rangle \otimes \mid \varphi\rangle.$$

By linearity it follows that

$$\alpha \mid \psi\rangle + \beta \mid \varphi\rangle \quad \mapsto \quad \alpha \mid \psi\rangle \otimes \alpha \mid \psi\rangle + \beta \mid \varphi\rangle \otimes \beta \mid \varphi\rangle.$$

On the other hand, a quantum copy machine would then be able to clone the superposition state $\alpha \mid \psi\rangle + \beta \mid \varphi\rangle$ as

[13] 1 Ampere is the constant current which, when flowing through two parallel conductors of infinite length and negligible cross-section placed in vacuum 1 m apart, produce between them a force of $2x10^{-7}$ Newton/m).

[14] We can reconstruct an $n$-dimensional object from multiple $(n-1)$-dimensional projections using Radon Transforms, or its complex analog, the Penrose Transform; CAT scan is an application of this process.

$$\alpha \mid \psi\rangle + \beta \mid \varphi\rangle \quad \mapsto \quad (\alpha \mid \psi\rangle + \beta \mid \varphi\rangle) \otimes (\alpha \mid \psi\rangle + \beta \mid \varphi\rangle),$$

as shown in Figure 47. But we see that

$$\alpha \mid \psi\rangle \otimes \alpha \mid \psi\rangle + \beta \mid \varphi\rangle \otimes \beta \mid \varphi\rangle \neq (\alpha \mid \psi\rangle + \beta \mid \varphi\rangle) \otimes (\alpha \mid \psi\rangle + \beta \mid \varphi\rangle).$$

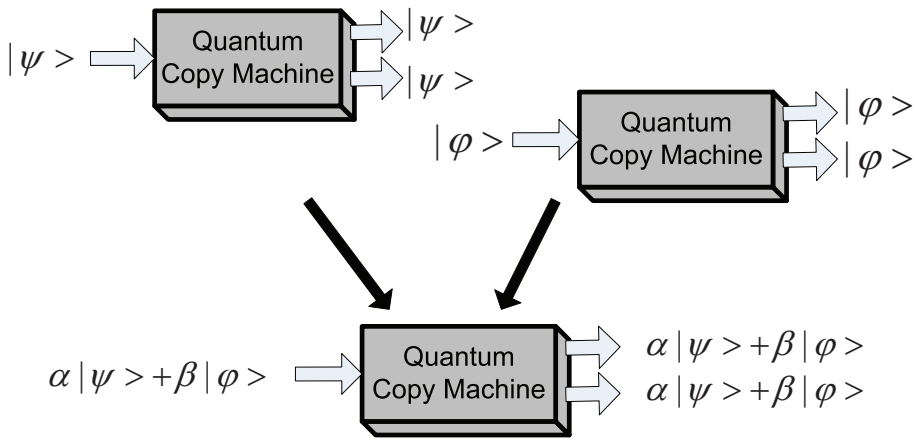



Figure 47: A quantum copy machine able to clone two arbitrary states $\mid \psi\rangle$ and $\mid \varphi\rangle$ it could also clone the superposition state $(\alpha \mid \psi\rangle + \beta \mid \varphi\rangle)$. Then the output of the quantum copy machine would be the tensor product of the original superposition state and its clone $(\alpha \mid \psi\rangle + \beta \mid \varphi\rangle) \otimes (\alpha \mid \psi\rangle + \beta \mid \varphi\rangle)$ different than the value one would expect due to the linearity of quantum state transformations, $(\alpha \mid \psi\rangle \otimes \alpha \mid \psi\rangle + \beta \mid \varphi\rangle \otimes \beta \mid \varphi\rangle)$.

We conclude that a *quantum copy machine able to clone two arbitrary quantum states* $\mid \psi\rangle$ *and* $\mid \phi\rangle$ *cannot clone a superposition of the two states,* $\mid \xi\rangle = \alpha \mid \psi\rangle + \beta \mid \phi\rangle$.

Surprisingly, the fact that a measurement alters the state of a quantum system could be very beneficial; one cannot eavesdrop on a quantum communication channel without being detected and this has important applications in computer security. Quantum key distribution protocols exploit this property of quantum information.

The rejection of non-determinism, a basic tenet of Quantum Mechanics, has motivated the development of several alterative models such as the hidden variable theory, discussed at length in Section 2.17. Bell theorem provides the theoretical foundation for the study of correlations of distant events and enables us to show the nonexistence of "local hidden variables." Bell inequality is derived using a very simple model of a physical system that makes only two common sense assumptions: physical properties are independent of observations, the *realism principle*, and the measurement of different physical properties of different objects carried out by different observers at distinct locations cannot influence each other, *the locality principle*. John Bell proved that independently of any specific physical theory *there is an upper limit to the correlation of distant events* if the principle of locality is valid; any hidden variable account of the quantum theory applied to quantum systems consisting of distant subsystems must have the separability property thus, no such hidden variable theory can be correct. Bell showed that a deterministic model based on hidden variables generates results

whose averages are identical with those predicted by Quantum Mechanics, but the hidden variables do not satisfy the separability requirement discussed in Section 2.20.

While no experiment can prove a theory to be right, a single experiment could prove a theory to be wrong; no experiment carried out from mid 1920s until this day has produced results contradicting the prediction of Quantum Mechanics. Yet, quantum systems violate Bell inequality and the only logical conclusion is that one of the two assumptions of the Bell model, or possibly both must be false. It is hard to accept that physical properties depend on our observations and discount the realism; alternatively, can we discount the locality?

## 2.2 Copenhagen Interpretation of Quantum Mechanics

Quantum Mechanics, perhaps one of the most successful scientific theory in the history of mankind, is in some respects counterintuitive, challenges our imagination, seems to violate fundamental principles of classical Physics, and raises many questions that continue to preoccupy the minds of physicists and philosophers since the early 1930s. Among them: Is there an objective reality? What is the role of the experiment? How accurate can our knowledge about the physical objects be? What is the relationship between Quantum Mechanics and classical Physics?

Classical Physics is deterministic; we know how the system will evolve in time if we determine the initial state of a physical system by observing its properties at the time we choose to be the initial moment, as well as, all external forces acting on the system. Moreover, the act of observing the system does not affect its later behavior. In classical physics there is a clear and unambiguous distinction between the system and the measuring instrument; the description of a physical system is objective, it does not depend on measurements carried out at a later time.

The *Copenhagen interpretation of Quantum Mechanics* is the name given to a set of principles formulated by Niels Bohr, Werner Heisenberg, Max Born, and several other physicists who spent time at the Institute for Theoretical Physics of the University of Copenhagen[15] in their attempt to answer the questions raised by Quantum Mechanics. Niels Bohr had a very practical approach to these questions; he believed that Quantum Mechanics is a natural generalization of classical physics, that classical concepts did not change their meaning, but that their application was restricted. Early on Bohr observed that his quantum theory of the hydrogen atom leads to results that coincide approximately with the ones provided by classical electrodynamics. He then postulated the so called correspondence rule, a heuristic principle which states that the behavior of systems described by the Quantum Mechanics reproduces classical Physics in the limit of large quantum numbers. The correspondence principle stipulates that a new scientific theory should be able to explain the phenomena under the condition when the earlier theory is valid. For example, Einstein special relativity satisfies the correspondence principle, because it reduces to classical mechanics in the limit of velocities small compared to the speed of light; general relativity reduces to Newtonian gravity in the limit of weak gravitational fields; Laplace theory of celestial mechanics reduces to Kepler's when interplanetary interactions are ignored, and Kepler's reproduces Ptolemy's equation in

[15]The Institute for Theoretical Physics of the University of Copenhagen was named "Niels Bohr Institute" on October 7, 1965 on Niels Bohr's 80th birthday.

a coordinate system where the Earth is stationary; statistical mechanics reproduces thermodynamics when the number of particles is large. Paul Feyerabend, another philosopher of science, believes that *succeeding theories like classical and quantum mechanics are incommensurable, and that there is partial lack of rationality in the choice between incommensurable theories.*

Complementarity is a principle formulated by Niels Bohr in 1928. Complementarity reflects de Broglie duality principle; photons, electrons, and other atomic or sub-atomic objects sometimes exhibit wave-like behavior, and at other times particle-like behavior. Sometimes it is possible to observe simultaneously both types of behavior, e.g., in case of double-slit experiments. Bohr, expressed the view that the complementarity of kinematic and dynamic properties of quantum particles can only be observed in mutually exclusive experiments.

The Copenhagen interpretation refers to concepts such as Bohr complementarity and correspondence principle, Born statistical interpretation of the wave function, and nondeterminism. Almost synonymous with nondeterminism, the Copenhagen interpretation aims to account for the violation of the basic principles classical physics rests on [477]:

- *Causality:* every event has a cause.
- *Determinism:* every later state of a system is uniquely determined by any earlier state.
- *Continuity:* when a process evolves from an initial to a final state it has to go through all intervening states.
- *Conservation of energy:* the energy of a closed system can be transformed into various forms but is never gained, lost, or destroyed.
- *The principle of space and time:* physical objects exist separately in space and time in such a way that they are localizable and countable. Processes that describe the evolution of physical objects take place in space and time.

Some of the tenets of the philosophy of Bohr are:

(i) Theory must rely on experimental practice.

(ii) Experimental practice assumes a pre-scientific practice of (1) description which establishes the requirements for a measurement apparatus and defines what scientific experience means, and (2) understanding the role of the environment, the identification of physical objects. This pre-scientific experience defines primitive categories such as: duration and change of duration, position of a physical object and change of position, and cause and effect relations. These categories are necessary for objective knowledge and must be part of any description of Nature.

(iii) Classical physics is based on exact concepts derived from these categories. The concepts of classical physics allow us to understand and to communicate results of our experiments.

(iv) The quantization of the energy of a harmonic oscillator (implicitly, of the energy radiated by a blackbody) discovered by Plank requires a revision of the foundation for the use of classical concepts. The use of classical concepts is well defined only when effects of energy quantization are negligible, when Plank constant can be neglected. Classical concepts can only

be applied to phenomena which are macroscopic manifestations of a measurement performed on the object. They reflect the uncontrollable interaction between a measurement apparatus and the physical object.

(v) The quantum mechanical description of an object requires that a separation line be drawn between the description of the object and of the measurement apparatus. But, this line of separation is not between the macroscopic instrument and the microscopic object. Parts of the measurement apparatus may sometimes be included in the quantum mechanical description of the physical object.

(vi) Schrödinger wave equation has a symbolic character, it cannot provide a visual description of physical reality, but it can be used to predict the outcome of a measurement.

We conclude this section with the words of Richard Feynman in *Lecture Notes on Physics* [149]: "The principle of science, the definition, almost, is the following: The test of all knowledge is the experiment. Experiment is the sole judge of the scientific *truth*. But what is the source of knowledge? Where do the laws to be tested come from? Experiment, itself helps to produce these laws, in the sense that it gives us hints. But also needed is the imagination to create from these hints the great generalizations - to guess at the wonderful, simple, but very strange patterns beneath them all, and then to experiment to check again whether we have made the right guess."

The next topics we address cover the measurements of quantum systems and require a closer examination of the density operator.

## 2.3 Mixed States and the Density Operator

The density operator allows us to distinguish pure states from mixed states of a quantum system. Pure states are characterized by maximum knowledge, they can be expressed as superpositions of basis vectors of an orthonormal basis. Mixed states require a statistical characterization and describe:

(i) Ensembles, or statistical mixtures of pure states; the system can be in any of the pure states $\mid \psi_1\rangle, \mid \psi_2\rangle, \mid \psi_3\rangle, \ldots$ with probabilities $p_1, p_2, p_3, \ldots$.

(ii) Composite systems; for example, two systems, $\mathcal{A}$ and $\mathcal{B}$, components of a larger system, $\mathcal{AB}$, in an entangled pure state.

The density operator of a system in a pure state $\mid \psi\rangle$ is $\rho = \mid \psi\rangle\langle\psi \mid$, while the density operator of the ensemble of pure states $\mid \psi_1\rangle, \mid \psi_2\rangle, \mid \psi_3\rangle \ldots$ with probabilities $p_1, p_2, p_3, \ldots$ is

$$\rho = \sum_{(i)} p_i \mid \psi_i\rangle\langle\psi_i \mid .$$

Given the bipartite system $\mathcal{AB}$ in the entangled pure state $\mid \psi_{AB}\rangle$ the density operator of $\mathcal{A}$ captures only the information available to an observer who has infinitely many opportunities to examine subsystem $\mathcal{A}$; the reduced density operator of subsystem $\mathcal{A}$ is defined as

$$\rho_A = \mathrm{tr}_B \mid \psi_{AB}\rangle\langle\psi_{AB} \mid .$$

**Example.** *A quantum system in a mixed state.* To relate the concept of a mixed state with physical systems we discuss briefly the Stern-Gerlach experiment. Recall that the spin of an atomic or subatomic particle is the intrinsic angular momentum of the particle. The spin is quantified; the spin quantum number $s$ takes discrete values. There are two types of particles: *fermions*, or 1/2-spin particles, with $s = \pm 1/2$ and *bosons*, or integer-spin particles with $s = 0, \pm 1$; for example, the electrons are fermions, while the photons are bosons.

The spin interacts with a magnetic field and the Stern-Gerlach experiment shows that a beam of hydrogen atoms consists of a mix of atoms with two different spin values. The experimental setup in Figure 48 uses a non-uniform magnetic field whose $z$ axis component is normal to the planar cap; the components of the magnetic field along $x$ and $y$ axis are negligible. The atomic beam entering the apparatus consists of a statistical mixture of hydrogen atoms in spin states $|\uparrow\rangle$ and $|\downarrow\rangle$ deflected upwards and, respectively, downwards. If the atomic beam consists of $N$ particles and $N_\uparrow$ is the number of atoms deflected upwards then, for very large $N$, the probability of an atom to have its spin up approaches the ratio $N_\uparrow/N$; similarly, if $N_\downarrow$ is the number of atoms deflected downwards the probability of an atom to have its spin down approaches the ratio $N_\downarrow/N$.

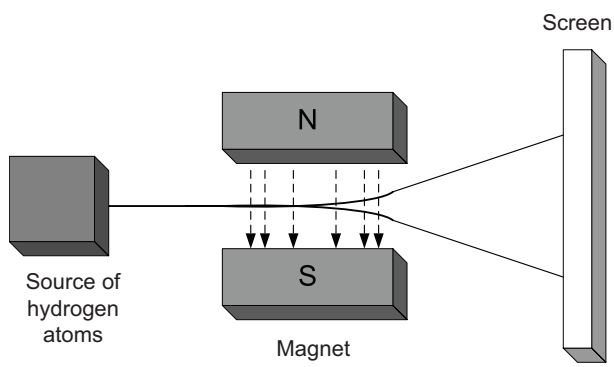



Figure 48: The hydrogen atomic beam entering the apparatus used in the Stern-Gerlach experiment consists of a statistical mixture of atoms in spin states $|\uparrow\rangle$ and $|\downarrow\rangle$ deflected upwards and, respectively, downwards.

In the Stern-Gerlach experiment a *pure spin state* corresponds to a *completely polarized beam*, while a statistical mixture corresponds to either a *partially polarized beam* when the probabilities of possible states are unequal, or to an *unpolarized beam* if the probabilities of the states are equal. The value of an observable cannot be predicted with certainty. Yet, we have to make a distinction between the probabilities associated with a pure state and the probabilities associated with impure states or statistical mixtures. For example, in case of 1/2-spin particles discussed above there are two possible pure spin states, $|\uparrow\rangle$ and $|\downarrow\rangle$ each with density matrix $\rho_\uparrow = |\uparrow\rangle\langle\uparrow|$ and $\rho_\downarrow = |\downarrow\rangle\langle\downarrow|$, respectively. If $N = N_\uparrow + N_\downarrow$ then $p_\uparrow = N_\uparrow/N$, $p_\downarrow = N_\downarrow/N$, and $p_\uparrow + p_\downarrow = 1$. The density matrix of the mixed state is $\rho = p_\uparrow\rho_\uparrow + p_\downarrow\rho_\downarrow$.

After this qualitative discussion of pure and mixed states we switch our focus to properties of the density operator; recall from Section 1.11 that the density is a self-adjoint operator. It

is easy to show that any self-adjoint $2 \times 2$ matrix $\rho$ can be represented as

$$\rho = \frac{1}{2}(\sigma_I + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z) = \frac{1}{2}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1-r_z \end{pmatrix}$$

with $r_x, r_y, r_z \in \mathbb{C}$ and $\sigma_I, \sigma_x, \sigma_y, \sigma_z$ the Pauli matrices

$$\sigma_I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then $r_x, r_y$, and $r_z$ are equal to

$$r_x = \rho_{12} + \rho_{21}, \quad ir_y = \rho_{21} - \rho_{12}, \quad r_z = \rho_{11} - \rho_{22} \quad \text{if} \quad \rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}.$$

The trace of the density operator has the following properties

$$\text{tr}(\rho) = 1, \qquad \text{tr}(\rho^2) = \frac{1}{2}(1+r^2), \qquad \text{where} \qquad r^2 = r_x^2 + r_y^2 + r_z^2 \leq 1.$$

*Proof:* We use the previous expression of the density operator and construct the trace of the density operator as the sum of the diagonal elements:

$$\text{tr}(\rho) = \frac{1}{2}((1+r_z) + (1-r_z)) = 1.$$

It is easy to see that

$$\rho^2 = \frac{1}{2}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1-r_z \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1+r_z & r_x - ir_y \\ r_x + ir_y & 1-r_z \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1+r^2+2r_z & 2(r_x - ir_y) \\ 2(r_x + ir_y) & 1+r^2-2r_z \end{pmatrix}$$

Thus,

$$\text{tr}(\rho^2) = \frac{1}{2}(1+r^2).$$

The condition $\det(\rho) \geq 0$ becomes

$$\det(\rho) = \frac{1}{4}[(1-r_z^2) - (r_x^2 + r_y^2)] = \frac{1}{4}(1-r^2) \geq 0$$

thus, $r^2 \leq 1$. This expression implies that there is a one-to-one correspondence between the possible density matrices of a single qubit and points either on the surface of the Bloch sphere, $\mid r \mid = 1$, or within the Bloch ball $0 \leq \mid r \mid < 1$, Figure 49.

**Properties of the density operator.** The density operator has several properties:

I. If a qubit is in a pure basis then it can be represented as a point on the Bloch sphere, $\det(\rho) = 0$, and its density matrix could only be

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{or} \qquad \rho = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

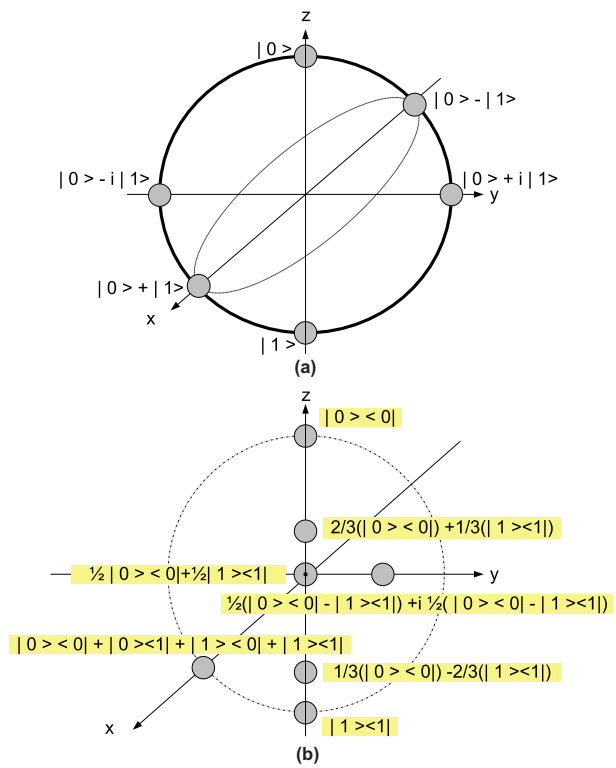



Figure 49: Bloch sphere and Bloch ball. (a) Pure states are points on the surface of a Bloch sphere. (b) Pure and mixed states are points on the surface of or inside a Bloch ball, respectively.

II. If the density operator of a single qubit has only one non-zero eigenvalue then the state is a pure state. If the density operator has more than one non-zero eigenvalue, then the state is a mixed state.

III. Infinitely many ensembles of pure states may be characterized by the same density operator.

IV. The density operator of an equal mixture of two orthogonal single qubit pure states is

$$\rho_{eq} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \rho_{ne}.$$

with $\rho_{ne}$ the density operator of a mixture of states with unequal probabilities

*Proof:* For pure states $r^2 = 1$ thus $\det(\rho) = 1 - r^2 = 0$; the determinant of the density operator of states represented as points on the surface of the Bloch (solid) sphere is equal to zero. Any $2 \times 2$ density operator $\rho$ can be diagonalized

$$\rho = \begin{pmatrix} \alpha_{11} & 0 \\ 0 & \alpha_{22} \end{pmatrix}.$$

Then $\det(\rho) = \alpha_{11}\alpha_{22} = 0$; the trace of the density operator must be one, $\mathrm{tr}(\rho) = \alpha_{11}+\alpha_{22} = 1$.

The two equations have two sets of solutions

$$\alpha_{11} = 1,\ \alpha_{22} = 0 \quad \text{or} \quad \alpha_{11} = 0,\ \alpha_{22} = 1.$$

We conclude that the density operator of a pure basis state could only be

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{or} \quad \rho = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

We notice that the first expression of the density operator corresponds to $\rho =\mid 0\rangle\langle 0 \mid$ and the second to $\rho =\mid 1\rangle\langle 1 \mid$. The density operator of a pure state can be represented as the sum of outer products of basis states.

**Example.** *The density matrix of the state* $\mid \varphi\rangle = 1/\sqrt{2}(\mid 0\rangle + \mid 1\rangle)$ *is*

$$\rho_\varphi =\mid \varphi\rangle\langle\varphi \mid = \frac{\mid 0\rangle\langle 0 \mid + \mid 0\rangle\langle 1 \mid + \mid 1\rangle\langle 0 \mid + \mid 1\rangle\langle 1 \mid}{2} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

The diagonalization of the density operator may have caused a rotation of the basis vectors. If we call the new basis vectors $\mid \hat{0}\rangle$ and $\mid \hat{1}\rangle$ resulting from the rotation of $\mid 0\rangle$ and $\mid 1\rangle$, then the density operator of a pure state can only be $\rho =\mid \hat{0}\rangle\langle\hat{0} \mid$ or $\rho =\mid \hat{1}\rangle\langle\hat{1} \mid$.

**Example.** *An equal mixture* ($p_0 = p_1 = 1/2$) *of the pure states of a single qubit* $\mid 0\rangle$ *and* $\mid 1\rangle$ *and an equal mixture of the pure states* $\mid +\rangle = 1/\sqrt{2}\,(\mid 0\rangle + \mid 1\rangle)$ *and* $\mid -\rangle = 1/\sqrt{2}\,(\mid 0\rangle - \mid 1\rangle)$ *have the same density operator.*

The density operator of the first mixture is

$$\rho = \sum_i p_i\rho_i = \frac{1}{2}\rho_0 + \frac{1}{2}\rho_1 = \frac{1}{2}\mid 0\rangle\langle 0 \mid + \frac{1}{2}\mid 1\rangle\langle 1 \mid = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\mathrm{tr}\,[\rho_0] = (1+0) = 1 \quad \text{and} \quad \mathrm{tr}\,[\rho_1] = (0+1) = 1.$$

The density operator of the equal mixture of the states $\mid +\rangle = 1/\sqrt{2}\,(\mid 0\rangle + \mid 1\rangle)$ and $\mid -\rangle = 1/\sqrt{2}\,(\mid 0\rangle - \mid 1\rangle)$ is

$$\rho_+ =\mid +\rangle\langle + \mid = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \mathrm{tr}\,[\rho_+] = \frac{1}{2}(1+1) = 1$$

$$\rho_- =\mid -\rangle\langle - \mid = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad \mathrm{tr}\,[\rho_-] = \frac{1}{2}(1+1) = 1$$

and

$$\rho = \sum_i p_i\rho_i = \frac{1}{2}\rho_+ + \frac{1}{2}\rho_- = \frac{1}{2}\times\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2}\times\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathrm{tr}(\rho^2) = \mathrm{tr}\left[\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\right] = \mathrm{tr}\left[\begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}\right] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} < 1.$$

The trace of the density operator satisfies the condition $\mathrm{tr}(\rho^2) = 1$ for pure states and $\mathrm{tr}(\rho^2) < 1$ for mixed states.

**Example.** *Consider two mixtures of states with unequal probabilities $p_0 = 1/4$ and $p_1 = 3/4$.*

First, we consider a mixture of states $\mid 0\rangle$ and $\mid 1\rangle$; the density operator is

$$\rho'_{ne} = \sum_i p_i \rho_i = \frac{1}{4}\rho_0 + \frac{3}{4}\rho_1 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{3}{4}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Now we consider a mixture of the states $\mid +\rangle = 1/\sqrt{2}\,(\mid 0\rangle + \mid 1\rangle)$ and $\mid -\rangle = 1/\sqrt{2}\,(\mid 0\rangle - \mid 1\rangle)$; the density operator is

$$\rho''_{ne} = \sum_i p_i \rho_i = \frac{1}{4}\rho_+ + \frac{3}{4}\rho_- = \frac{1}{4} \times \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{3}{4} \times \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}$$

or

$$\rho''_{ne} = \frac{1}{4}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

We see that $\rho'_{ne} \neq \rho_{eq}$ and $\rho''_{ne} \neq \rho_{eq}$.

We now discuss the defining properties of a density operator; an operator $\rho$ in the Hilbert $\mathcal{H}_n$ can be considered a density operator if and only if it has three properties:

(i)t is self-adjoint, $\rho = \rho^\dagger$;

(ii) it is non-negative, $\rho \geq 0$; and

(iii) its trace is equal to 1, $\mathrm{tr}\,[\rho] = 1$.

The following proposition shows that operators with these properties form a convex set[16].

**Proposition.** *The set of density operators in a Hilbert space $\mathcal{H}_n$ is a convex linear subset of $n \times n$ self-adjoint matrices.*

*Proof:* Assume that $\rho_a$ and $\rho_b$ are density operators, thus satisfy the three properties. Given a real $\lambda$ such that $0 \leq \lambda \leq 1$ we construct the following operator:

$$\rho(\lambda) = \lambda\rho_a + (1-\lambda)\rho_b.$$

We see immediately that $\rho(\lambda)$ is self-adjoint because $\rho_a$ and $\rho_b$ are self-adjoint:

$$\rho^\dagger(\lambda) = [\lambda\rho_a + (1-\lambda)\rho_b]^\dagger = \lambda\rho_a^\dagger + (1-\lambda)\rho_b^\dagger = \lambda\rho_a + (1-\lambda)\rho_b = \rho(\lambda).$$

[16] A convex subspace of a vector space is a subset that contains all the points of a straight line connecting any two points in the subset.

Given any state $\mid \varphi\rangle \in \mathcal{H}_n$ we see that

$$\langle \varphi \mid \rho(\lambda) \mid \varphi\rangle = \lambda\langle \varphi \mid \rho_a \mid \varphi\rangle + (1-\lambda)\langle \varphi \mid \rho_b \mid \varphi\rangle \geq 0.$$

Finally,

$$\mathrm{tr}\left[\rho(\lambda)\right] = \lambda \mathrm{tr}\left[\rho_a\right] + (1-\lambda)\mathrm{tr}\left[\rho_b\right] = \lambda + (1-\lambda) = 1.$$

The property of density matrices to form a convex set has important consequences, as we shall see in the next chapters. There are infinitely many ways to express $\rho(\lambda)$ as a convex combination of other states thus, infinitely many ways to prepare a mixed state; if we have an observable $M$ associated with the self-adjoint operator $\mathbf{M}$ and wish to determine the expected value for different preparations of the state $\rho_a$ and $\rho_b$ these values are indistinguishable. The expected value of the observable $\mathbf{M}$ for the system with density matrix $\rho(\lambda)$ is defined in Section 2.10

$$\langle \mathbf{M}\rangle = \mathrm{tr}\left[\mathbf{M}\rho(\lambda)\right].$$

As before, $\rho(\lambda) = \lambda\rho_a + (1-\lambda)\rho_b$. Call $\langle \mathbf{M}_a\rangle = \mathrm{tr}\left[\mathbf{M}\rho_a\right]$ and $\langle \mathbf{M}_b\rangle = \mathrm{tr}\left[\mathbf{M}\rho_b\right]$ the average values of the measurements for the systems with density matrices $\rho_a$ and $\rho_b$. Then

$$\langle \mathbf{M}\rangle = \lambda\langle \mathbf{M}_a\rangle + (1-\lambda)\langle \mathbf{M}_b\rangle.$$

Thus, regardless on how we have prepared a mixed state with the density operator $\rho$, the average value of the outcome of a measurement of any observable of the system will be the same. On the other hand, a pure state can be prepared in one and only one way; a pure state is an eigenvector of some observable and the outcome of the measurement of that observable will always be equal to 1.

The density operator, $\rho$, of an ensemble captures only the information available to an observer who has the opportunities to examine infinitely many states of the ensemble. The entropy of a mixed state of non-orthogonal pure states, which are not mutually distinguishable, is given by von Neumann entropy:

$$S(\rho) = -\mathrm{tr}\left(\rho \log \rho\right).$$

When the pure states of the ensemble are orthogonal, they can be treated as classical states and then the entropy is given by the known Shannon expression:

$$H = -\sum_{(i)} p_i \times \log p_i.$$

In summary, the density operator is a positive semi-definite, self-adjoint operator with trace equal to unity. The eigenvalues of a positive semi-definite operator are real numbers greater than or equal to zero, thus, the operator must have a positive trace and a non-negative determinant. The trace of the density operator satisfies the condition $\mathrm{tr}(\rho^2) = 1$ for pure states and $\mathrm{tr}(\rho^2) < 1$ for mixed states.

## 2.4 Purification of Mixed States

Pure states have a particular appeal for Quantum Information Theory thus, a legitimate question is if given a quantum system $\mathcal{A}$ in a mixed state we can identify another system $\mathcal{B}$ such that the bipartite system $\mathcal{C} = \mathcal{AB}$ is in a pure state. System $\mathcal{B}$, called a *reference system*, is only a mathematical construct without a physical support.

The process of finding a reference system given a quantum system $\mathcal{A}$ in a mixed state $\rho_{\mathcal{A}}$ is called purification. If $\mid \psi_{\mathcal{C}}\rangle$ is the pure state of the bipartite system $\mathcal{C} = \mathcal{AB}$ we wish that

$$\rho_{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}} \left( \mid \psi_{\mathcal{C}}\rangle\langle\psi_{\mathcal{C}} \mid \right).$$

Let $\mathcal{H}_n$ be the Hilbert space for system $\mathcal{A}$ with an orthonormal basis $\{e_i\}$; then the decomposition of the density operator of $\mathcal{A}$ is

$$\rho_{\mathcal{A}} = \sum_i p_i \mid e_i^A\rangle\langle e_i^A \mid .$$

The procedure of purification consists of the following steps:
(i) Introduce $\mathcal{B} \in \mathcal{H}_n$, which shares the same space with $\mathcal{A}$. Let $\{\mid e_j^B\rangle\} \in \mathcal{H}_n$ be an orthonormal basis for $\mathcal{H}_n$.
(ii) Define a pure state of the bipartite system

$$\mid \psi_{\mathcal{C}}\rangle = \sum_i \sqrt{p_i} \mid e_i^A\rangle \otimes \mid e_i^B\rangle.$$

We now compute the reduced density operator for $\mathcal{A}$ for this state as

$$\begin{aligned} \mathrm{tr}_B \left( \mid \psi_{\mathcal{C}}\rangle\langle\psi_{\mathcal{C}} \mid \right) &= \textstyle\sum_i \sum_j \sqrt{p_i p_j} \mid e_i^A\rangle\langle e_i^A \mid \mathrm{tr}\left( \mid e_i^B\rangle\langle e_j^B \mid \right) = \\ &= \textstyle\sum_i \sum_j \sqrt{p_i p_j} \mid e_i^A\rangle\langle e_i^A \mid \delta_{ij} = \sum_i p_i \mid e_i^A\rangle\langle e_i^A \mid = \rho_{\mathcal{A}}. \end{aligned}$$

A formal definition of purification requires a more general formalism for positive semidefinite operators [445]. First, we review several linear algebra concepts: the kernel of a linear transformation $\mathcal{L} : \mathcal{H}_a \mapsto \mathcal{H}_b$, denoted as $\mathrm{Ker}(\mathcal{L})$, is the set of all vectors $\mid \zeta\rangle \in \mathcal{H}_a$ such that $\mathcal{L}(\mid \zeta\rangle) = 0$; the rank of a linear transformation is the dimension of its range, in our example $\mathrm{rank}(\mathcal{H}_a) = \dim(\mathcal{H}_b)$; if $\mathcal{L}(\mathcal{H}_a)$ denotes a linear transformation of the vector space $\mathcal{H}_a$ then $\mathrm{supp}\,[\rho]$, the support of a normal operator $\rho \in \mathcal{L}(\mathcal{H}_a)$ is defined as the subspace of $\mathcal{H}_a$ spanned by the eigenvectors of $\rho$ having a nonzero eigenvalue

$$\mathrm{supp}\,[\rho] = \mathrm{span}\{\mid \xi_i\rangle : 1 \leq i \leq n, \lambda_i \neq 0\}$$

where $\rho = \sum_{i=1}^n \lambda_i \mid \xi_i\rangle\langle\xi_i \mid$ is a spectral decomposition of $\rho$. Obviously

$$\dim(\mathrm{supp}\,[\rho]) = \mathrm{rank}(\rho).$$

**Formal definition of purification.** Let $\mathcal{H}_a$ and $\mathcal{H}_b$ be Hilbert spaces and $\rho \in \mathrm{Pos}(\mathcal{H}_a)$ be any positive semidefinite operator. Then a purification of $\rho$ in $\mathcal{H}_a \otimes \mathcal{H}_b$ is any vector $\mid \varphi\rangle$ such that

$$\rho = \mathrm{tr}_{\mathcal{H}_b} [| \varphi\rangle\langle\varphi |] .$$

Let $\Pi_{\mathcal{H}_c}$ be the projection of $\mathcal{H}_a$ onto a subspace $\mathcal{H}_c$. Then

$$\Pi_{\mathrm{supp}(\rho)} = \sum_{i \in S} | \xi_i\rangle\langle\xi_i |$$

with $\{| \xi_i\rangle : i \in R\}$, an orthonormal basis in $\mathcal{H}_a$. If $\dim [\mathcal{H}_b] \geq \mathrm{rank}\,[\rho]$ then there exists a mapping $A \in \mathcal{L}(\mathcal{H}_a, \mathcal{H}_b)$ such that:

$$A^\dagger A = \Pi_{\mathrm{supp}(\rho)}.$$

Indeed, $\{| \xi_i\rangle : i \in R\}$, an orthonormal basis in $\mathcal{H}_a$, exists because $\dim [\mathcal{H}_b] \geq \mathrm{rank}\,[\rho] =| R |$. Thus,

$$A = \sum_{i \in S} \gamma_i \langle\xi_i |$$

satisfies the condition $A^\dagger A = \Pi_{\mathrm{supp}(\rho)}$.

**The existence of purification.** Let $\mathcal{H}_a$ and $\mathcal{H}_b$ be Hilbert spaces and $\rho \in \mathrm{Pos}(\mathcal{H}_a)$ be any positive semidefinite operator in $\mathcal{H}_a$. Then a purification of $\rho$ exists in $\mathcal{H}_a \otimes \mathcal{H}_b$ if and only if $\dim [\mathcal{H}_b] \geq \mathrm{rank}\,[\rho]$. Moreover, when $\dim [\mathcal{H}_b] \geq \dim [\mathcal{H}_a]$ a purification of $\rho$ always exists.

*Proof:* the proof of the two clauses is rather short; when $\dim [\mathcal{H}_b] \geq \mathrm{rank}\,[\rho]$ then there exist the mapping $A$ such that $A^\dagger A = \Pi_{\mathrm{supp}(\rho)}$. Given a matrix $A$ we denote by $\mathrm{vec}(A)$ the column vector constructed as follows: we take each row vector of $A$, transpose it and stack the vectors in order, on top on one another. Define

$$| \psi\rangle = \mathrm{vec}(\sqrt{\rho} A^\dagger) \in \mathcal{H}_a \otimes \mathcal{H}_b.$$

Then $| \psi\rangle$ is a purification of $\rho$ because

$$\mathrm{tr}_{\mathcal{H}_b} [| \psi\rangle\langle\psi |] = \sqrt{\rho} A^\dagger A \sqrt{\rho} = \rho.$$

Conversely, when $| \psi\rangle \in \mathcal{H}_a \otimes \mathcal{H}_b$ satisfies the condition $\mathrm{tr}_{\mathcal{H}_b} [| \psi\rangle\langle\psi |] = \rho$, consider $\Psi = \mathrm{vec}^{-1}(| \psi\rangle)$. Then,

$$\rho = \mathrm{tr}_{\mathcal{H}_b} [| \psi\rangle\langle\psi |] = \Psi\Psi^\dagger,$$

therefore $\mathrm{rank}(\psi) = \mathrm{rank}(\rho)$; as the rank of $\psi$ cannot exceed the dimension of $\mathcal{H}_b$ it follows that

$$\dim [\mathcal{H}_b] \geq \mathrm{rank}\,[\rho] = \mathrm{rank}(\rho).$$

This property will be used in Chapter 3 when we discuss information transmission through a noisy quantum channel.

Next we discuss quantum measurements and start our presentation with a review of a postulate of Quantum Mechanics, the Born rule.

## 2.5 Born Rule

Quantum measurements are subject to Born rule; this fundamental postulate[17] of Quantum Mechanics is typically formulated as:

*Consider a system in state* $\mid \psi\rangle \in \mathcal{H}_n$ *and an observable* $A$*. Let* $\mathbf{P}_i, 1 \leq i \leq n$ *be a set of projectors on* $\{\mid a_1\rangle, \mid a_1\rangle, \ldots \mid a_n\rangle\}$*, in an orthonormal basis of* $\mathcal{H}_n$ *where we measure the observable* $A$*; the projector* $\mathbf{P}_i$ *has* $\mid a_i\rangle$ *as an eigenvector and* $\lambda_i$ *as the corresponding eigenvalue. Then the probability for a measurement of the observable* $A$ *of yielding a particular value* $\lambda_i$ *is*

$$p(\lambda_i) = \mid \langle a_i \mid \psi\rangle \mid^2 .$$

In Section 1.7 we have already shown that the probability of an outcome $\lambda_i$ is

$$p(\lambda_i) = \langle \psi \mid \mathbf{P}_i \mid \psi\rangle$$

with $\mathbf{P}_i = \mid a_i\rangle\langle a_i \mid$; this expression is equivalent to Born rule

$$\langle \psi \mid \mathbf{P}_i \mid \psi\rangle = \langle \psi \mid a_i\rangle\langle a_i \mid \psi\rangle = (\langle a_i \mid \psi\rangle)^\dagger(\langle a_i \mid \psi\rangle) = \mid \langle a_i \mid \psi\rangle \mid^2 .$$

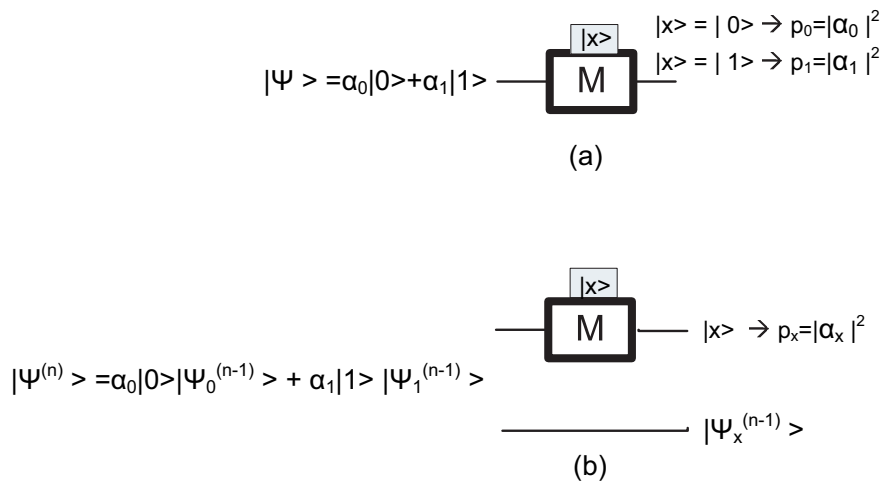



Figure 50: (a) The measurement of a single qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ using a one-qubit measurement gate. (b) We use a one-qubit measurement gate to measure only the first qubit of a register of $n$ qubits in state $\mid \psi^{(n)}\rangle = \alpha_0 \mid 0\rangle \otimes \mid \psi_0^{(n-1)}\rangle + \alpha_1 \mid 1\rangle \otimes \mid \psi_1^{(n-1)}\rangle$.

The measurement of a single qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ using a one-qubit measurement gate is illustrated in Figure 50(a). The measurement is performed in the orthonormal basis $(\mid 0\rangle, \mid 1\rangle)$. The instrument displays the result 0 corresponding to the state

[17] Even though no violation of the Born rule has been observed experimentally it would be desirable to replace this postulate by a derivation based on a set of fundamental assumptions. A. M. Gleason has presented a mathematical motivation for the form of Born rule [172] and more recently W. H. Zurek used a mechanism termed environment-assisted invariance, or *envariance*, for a novel derivation of Born rule [473].

$\mid 0\rangle$ with probability $p_0 = \mid \alpha_0 \mid^2$ and the result 1 corresponding to the state $\mid 1\rangle$ with probability $p_1 = \mid \alpha_1 \mid^2$.

The generalized Born rule for a register of $n$ qubits in state $\mid \psi^n\rangle \in \mathcal{H}_{2^n}$ when we carry out a measurement of only the first qubit using a one-qubit measurement gate is presented in Figure 50(b). In this case we consider two subspaces $\mathcal{H}_2$ and $\mathcal{H}_{2^{n-1}}$ of $\mathcal{H}_{2^n}$ and express the state of the $n$-qubit register as

$$\mid \psi^n\rangle = \alpha_0(\mid 0\rangle \otimes \mid \psi_0^{(n-1)}\rangle) + \alpha_1(\mid 1\rangle \otimes \mid \psi_1^{(n-1)}\rangle).$$

Here $\mid \psi_0^{(n-1)}\rangle$ and $\mid \psi_1^{(n-1)}\rangle$ are eigenvectors of two projectors in $\mathcal{H}_{2^{n-1}}$. If the result displayed by the instrument is 0 (when the single qubit measured is projected on $\mid 0\rangle$), then the remaining $n-1$ qubits are in state $\mid \psi_0^{(n-1)}\rangle$; when the result displayed by the instrument is 1 (when the single qubit measured is projected on $\mid 1\rangle$), then the remaining $n-1$ qubits are in state $\mid \psi_1^{(n-1)}\rangle$.

We can also apply the generalized Born rule to measuring the first $k$ qubits of an $n$-qubit register originally in state

$$\mid \psi^n\rangle = \sum_{i=0}^{2^k-1} \alpha_i \mid i^{(k)}\rangle \otimes \mid \psi_i^{(n-k)}\rangle.$$

When the result of the measurement performed on the first $k$ qubits is $x = r$ then we know that the remaining $(n-k)$ qubits are in state $\mid \psi_r^{(n-k)}\rangle$.

## 2.6 Measurement Operators

The measurement of an observable of a quantum system in state $\mid \psi\rangle \in \mathcal{H}_n$ is characterized by a set of operators, $\{\mathbf{M}_i\}$. Call $\{\lambda_i\}$ the set of possible outcomes of the measurement of the quantum system and recall that these outcomes are eigenvalues of the measurement operators and must be real numbers to allow a physical interpretation of the measurement process. Only self-adjoint operators have real eigenvalues thus, *the measurement operators* $\mathbf{M}_i$ *must be self-adjoint*

$$\mathbf{M}_i^\dagger \mathbf{M}_i = \mathbf{M}_i^2 = \mathbf{M}_i.$$

According to the measurement postulate of Quantum Mechanics, the probability that the outcome with index $i$, $\lambda_i$, occurs as a result of the measurement $\mathbf{M}_i$ is

$$p(i) = \langle \psi \mid \mathbf{M}_i \mid \psi\rangle.$$

The sum of the probabilities of all possible outcomes of the measurement must be 1

$$\sum_i p(i) = \sum_i \langle \psi \mid \mathbf{M}_i \mid \psi\rangle = 1.$$

This implies that the measurement operators must satisfy the <u>completeness equation</u>

$$\sum_i \mathbf{M}_i = \mathbf{I}.$$

The measurement causes the system to change its state; if the state of the system immediately prior to the measurement is $\mid \psi\rangle$ then, after the measurement $\mathbf{M}_i$, the state of the system is $\mid \varphi_i\rangle = \mathbf{M}_i \mid \psi\rangle$. The *normalized state* of the system is

$$\mid \varphi_i'\rangle = \frac{\mathbf{M}_i \mid \psi\rangle}{\sqrt{\langle\psi \mid \mathbf{M}_i \mid \psi\rangle}}.$$

**Example.** *Consider a single qubit in state* $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$. There are only two possible outcomes of a measurement in the canonical basis $\{\mid 0\rangle, \mid 1\rangle\}$; we can only observe the basis states $\mid 0\rangle$ or $\mid 1\rangle$. The corresponding projectors are

$$\mathbf{P}_0 = \mid 0\rangle\langle 0 \mid = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{P}_1 = \mid 1\rangle\langle 1 \mid = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

These two operators are self-adjoint and $\mathbf{P}_0^\dagger \mathbf{P}_0 = \mathbf{P}_0$ and $\mathbf{P}_1^\dagger \mathbf{P}_1 = \mathbf{P}_1$. The probability of the outcome 0 corresponding to basis state $\mid 0\rangle$ is

$$p(0) = \langle\psi \mid \mathbf{P}_0 \mid \psi\rangle = \mid \alpha_0 \mid^2$$

Indeed,

$$\mathbf{P}_0 \mid \psi\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix}$$

and

$$p(0) = \langle\psi \mid (\mathbf{P}_0 \mid \psi\rangle) = (\ \alpha_0^* \ \ \alpha_1^*) \begin{pmatrix} \alpha_0 \\ 0 \end{pmatrix} = \mid \alpha_0 \mid^2.$$

The normalized state of the qubit after applying the projector $\mathbf{P}_0$, up to a phase, is given by

$$\mid \varphi_0'\rangle = \frac{\mathbf{P}_0 \mid \psi\rangle}{\sqrt{\langle\psi \mid \mathbf{P}_0 \mid \psi\rangle}} = \frac{\alpha_0 \mid 0\rangle}{\sqrt{\mid \alpha_0 \mid^2}} = \mid 0\rangle.$$

Similarly, the probability of the outcome 1 corresponding to the basis state $\mid 1\rangle$ is

$$p(1) = \langle\psi \mid \mathbf{P}_1 \mid \psi\rangle = \mid \alpha_1 \mid^2 .$$

The normalized state of the qubit after applying the projector $\mathbf{P}_1$, up to a phase, is given by

$$\mid \varphi_1'\rangle = \frac{\mathbf{P} \mid \psi\rangle}{\sqrt{\langle\psi \mid \mathbf{P}_1 \mid \psi\rangle}} = \frac{\alpha_1 \mid 1\rangle}{\sqrt{\mid \alpha_1 \mid^2}} = \mid 1\rangle.$$

The completeness condition is satisfied: $p(0) + p(1) = \mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 = 1$ .

Next we discuss projective measurements and then POVM measurements.

## 2.7 Projective Measurements

In this section we discuss projection measurements whose outcomes are deterministic and introduce von Neumann axiomatic definition of the measurement process,

When the projectors $\mathbf{M}_k \in \mathcal{H}_n$ form an orthogonal set, $\mathbf{M}_j\mathbf{M}_k = 0$ when $j \neq k$, and if the set is complete, $\sum_k \mathbf{M}_k = \mathbf{I}$, then any observable $\mathbf{M}$ of the quantum system has a spectral decomposition:

$$\mathbf{M} = \sum_k k\mathbf{M}_k$$

with $\mathbf{M}_k$ the projector onto the eigenspace of $\mathbf{M}$ with eigenvalue $k$. The measurement operator corresponding to a basis vector $\mid k\rangle \in \mathcal{H}_n$ is $\mathbf{M}_k = \mid k\rangle\langle k \mid$; the operator $\mathbf{M}_k$ is self-adjoint and idempotent:

$$\mathbf{M}_k = \mathbf{M}_k^\dagger \quad \text{and} \quad \mathbf{M}_k^2 = \mid k\rangle\langle k \mid k\rangle\langle k \mid = \mid k\rangle\langle k \mid = \mathbf{M}_k.$$

To answer the question when is the outcome of a measurement deterministic, we use the concept of *projective measurement*.

Projective measurements. Let $\{\mathbf{M}_k\}$ be a set of self-adjoint measurement operators. The set is complete and consists of orthogonal operators and

$$\mathbf{M}_i\mathbf{M}_j = \delta_{ij}\mathbf{M}_i,$$

with $\delta_{ij}$ the Kronecker delta function; the operators allow us to conduct projective measurements.

When measuring a quantum system in state $\mid \psi\rangle$ the probability of observing the outcome with index $k$ is

$$p(k) = \langle \psi \mid \mathbf{M}_k \mid \psi\rangle.$$

*Projective measurements are repeatable; the outcome observed as a result of projective measurement is deterministic.* Indeed, consider a quantum system in state $\mid \psi\rangle$ immediately prior to a measurement carried out using the projective measurement operator $\mathbf{M}_k$; the state immediately after the measurement, $\mid \psi'\rangle$, is

$$\mid \psi'\rangle = \frac{\mathbf{M}_k \mid \psi\rangle}{\sqrt{p(k)}}.$$

If we apply the operator $\mathbf{M}_k$ again, this time to the state $\mid \psi'\rangle$, there is no state change. Indeed, the state after applying $\mathbf{M}_k$ for the second time is

$$\mid \psi''\rangle = \frac{\mathbf{M}_k \mid \psi'\rangle}{\sqrt{p'(k)}} = \frac{\mathbf{M}_k}{\sqrt{p'(k)}}\frac{\mathbf{M}_k \mid \psi\rangle}{\sqrt{p(k)}}$$

with

$$p(k) = \langle \psi \mid \mathbf{M}_k \mid \psi\rangle \quad \text{and} \quad p'(k) = \langle \psi' \mid \mathbf{M}_k \mid \psi'\rangle.$$

We can now derive the expression for $p'(k)$

$$p'(k) = \langle \psi' \mid \mathbf{M}_k \mid \psi' \rangle = \left[ \langle \psi \mid \frac{\mathbf{M}_k^\dagger}{\sqrt{p(k)}} \right] \mathbf{M}_k \left[ \frac{\mathbf{M}_k}{\sqrt{p(k)}} \mid \psi \rangle \right] = \frac{\langle \psi \mid \mathbf{M}_k \mid \psi \rangle}{p(k)}.$$

Then,

$$\sqrt{p'(k)p(k)} = \sqrt{\langle \psi \mid \mathbf{M}_k \mid \psi \rangle} = \sqrt{p(k)}.$$

Finally, we see that the state has not been affected by the second measurement

$$\mid \psi'' \rangle = \frac{\mathbf{M}_k}{\sqrt{p'(k)}} \frac{\mathbf{M}_k \mid \psi \rangle}{\sqrt{p(k)}} = \frac{\mathbf{M}_k \mid \psi \rangle}{\sqrt{p(k)}} = \mid \psi' \rangle.$$

The fact that projective measurements are repeatable has straightforward mathematical and physical interpretations: projecting a quantum state on the basis vectors of the Hilbert space transforms quantum information into classical one thus, applying again the same operator does not lead to a state change. The outcome observed as a result of such a measurement is deterministic: $p'(k) = \langle \psi' \mid \mathbf{M}_k \mid \psi' \rangle = 1$.

The <u>expected value of the projective measurement $\mathbf{M}$, $\mathbf{E}(\mathbf{M})$, denoted as $\langle \mathbf{M} \rangle$</u>, is defined as

$$\langle \mathbf{M} \rangle = \mathbf{E}(\mathbf{M}) = \sum_k kp(k) = \sum_k k \langle \psi \mid \mathbf{M}_k \mid \psi \rangle = \langle \psi \mid \left[ \sum_k k\mathbf{M}_k \right] \mid \psi \rangle = \langle \psi \mid \mathbf{M} \mid \psi \rangle.$$

The <u>standard deviation of a projective measurement $\mathbf{M}$</u> is defined as

$$\mathbf{E}\left[\mathbf{M} - \langle \mathbf{M} \rangle\right]^2 = \mathbf{E}\left[\mathbf{M}^2 - 2(\mathbf{M})(\langle \mathbf{M} \rangle) + (\langle \mathbf{M} \rangle)^2\right] = \langle \mathbf{M}^2 \rangle - (\langle \mathbf{M} \rangle)^2.$$

**Example.** *Assume that an observable of a qubit in state* $\mid \psi \rangle = (\mid 0 \rangle - \mid 1 \rangle)/\sqrt{2}$ *has two eigenvectors* $\mid 0 \rangle$ *and* $\mid 1 \rangle$ *with the eigenvalues* $+1$ *and* $-1$*, respectively. Then the probabilities of the two outcomes are* $p(+1) = p(-1) = 1/2$.

The projector operators corresponding to the two eigenvectors, $\mathbf{M}_0$ and $\mathbf{M}_1$ are:

$$\mathbf{M}_0 = \mid 0 \rangle \langle 0 \mid = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{M}_1 = \mid 1 \rangle \langle 1 \mid = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

The projective measurement of the state $\mid \psi \rangle$ produces the eigenvalue $(+1)$ corresponding to the eigenvector $\mid 0 \rangle$ with probability

$$p(+1) = \langle \psi \mid \mathbf{M}_0 \mid \psi \rangle = \langle \psi \mid 0 \rangle \langle 0 \mid \psi \rangle,$$

and the eigenvalue $(-1)$ corresponding to the eigenvector $\mid 1 \rangle$ with probability

$$p(-1) = \langle \psi \mid \mathbf{M}_1 \mid \psi \rangle = \langle \psi \mid 1 \rangle \langle 1 \mid \psi \rangle.$$

It is easy to see that

$$\mathbf{M}_0 \mid \psi\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$p(+1) = \langle \psi \mid (\mathbf{M}_0 \mid \psi\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}.$$

Similarly, $p(-1) = \langle \psi \mid (\mathbf{M}_1 \mid \psi\rangle) = 1/2$.

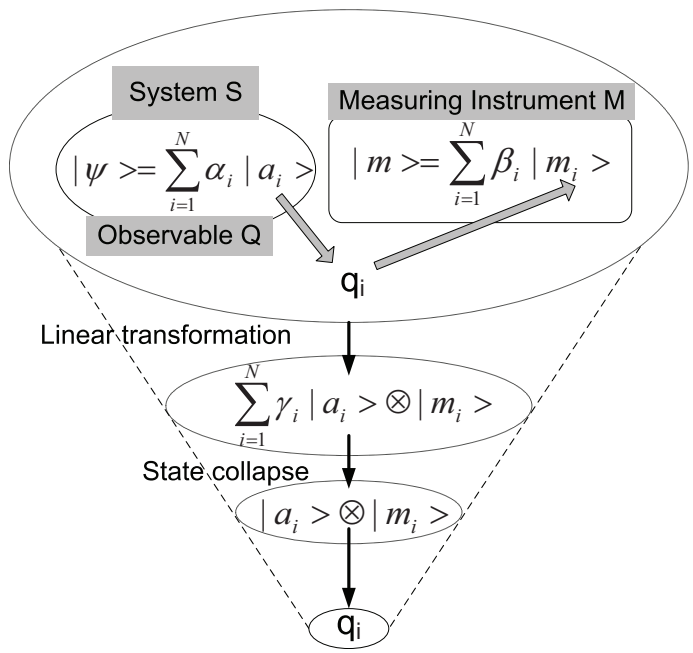



Figure 51: Two stage von Neumann-type measurement. We measure the observable $Q$ of the quantum system $S$ using a macroscopic apparatus $M$. At time $t$ the quantum system is in a superposition state $\mid \psi\rangle = \sum_i \alpha_i \mid a_i\rangle$ and the apparatus is in state $\mid m\rangle$. $q_i$ is the eigenvalue corresponding to the eigenvector $\mid a_i\rangle$ of observable $Q$. At the end of the first phase the composite system $S + M$ is in state $\sum_i \gamma_i \mid a_i\rangle \otimes \mid m_i\rangle$. Then, as a result of the measurement, the system collapses in the state $\mid a_i\rangle \otimes \mid m_i\rangle$ for some $i$.

John von Neumann provided an axiomatic definition of a quantum measurement process; his measurement scheme is an ancestor of the quantum decoherence theory and treats the measurement apparatus as a quantum object. According to this axiomatic definition the measurement consists of two stages, Figure 51:

1. The quantum system $S$ interacts with a macroscopic measuring apparatus $M$ to measure the quantity $Q$. This interaction is linear; if at time $t$ the quantum system is in a superposition state $\mid \psi\rangle = \sum_i \alpha_i \mid a_i\rangle$ and the apparatus is in state $\mid m\rangle$ the ensemble $S + M$ is in state $\sum_i \alpha_i \mid a_i\rangle \otimes \mid m\rangle$. In this expression $\mid a_i\rangle$ are the eigenvectors of $Q$ corresponding to the possible outcomes of the measurement $q_i$ (eigenvalues of $Q$). At the end of this stage the system is in state $\sum_i \gamma_i \mid a_i\rangle \otimes \mid m_i\rangle$ with $\mid m_i\rangle$ the state in which $M$ registers the outcome $q_i$; the states of the apparatus, $\mid m_i\rangle$, are orthonormal.

2. As a result of a nondeterministic process the ensemble $S+M$ "collapses" from the state $\sum_i \gamma_i \mid a_i\rangle \otimes \mid m_i\rangle$ to the state $\mid a_i\rangle \otimes \mid m_i\rangle$ for some $i$.

In summary, in a projective measurement the measurement operators are self-adjoint and idempotent. The number of such operators is equal to the dimension of the Hilbert space. Since orthogonal measurement operators commute, they correspond to simultaneous observables. A von Neumann measurement is a conditional expectation onto a maximal Abelian subalgebra of the algebra of all bounded operators acting on the Hilbert space [330].

## 2.8 Positive Operator Valued Measures (POVM)

Projective measurements are restrictive and not always possible; sometimes we destroy a quantum particle in the process of measurement thus, the repeatability of a projective measurement is violated. For example, a photon may be absorbed by a polarization filter and no longer available for measurements. In such cases where the system is measured only once, the post measurement state of the system is no longer of interest and the probabilities of the measurement outcomes are the ones that count.

A generalized type of measurement, the *Positive Operator-Valued Measure* or POVM, is concerned only with the statistics of the measurement. The POVM operators are not necessarily orthogonal or commutative and *allow the possibility of measurement outcomes associated with non-orthogonal states.* Recall that measurement operators corresponding to non-orthogonal states do not commute and are therefore not simultaneously observable. The number of POVM operators may differ from the dimension of the Hilbert space, while the number of projective operators is precisely equal to the dimension of the Hilbert space.

Given the set of measurement operators $\{\mathbf{M}_i\}$ describing a measurement performed on a quantum system in state $\mid \psi\rangle$, a <u>POVM</u> has the elements $\mathbf{A}_i$ defined by

$$\mathbf{A}_i = \mathbf{M}_i^\dagger \mathbf{M}_i \quad \text{and} \quad \sum_i \mathbf{A}_i = \mathbf{I}.$$

The operators $\mathbf{A}_i$ have several properties:

(i) Are positively defined; recall that a positively defined operator has real and positive eigenvalues. This property follows immediately from the definition of $\mathbf{A}_i$.

(ii) $p(i) = \langle \psi \mid \mathbf{A}_i \mid \psi\rangle$. This follows immediately from the fact that $p(i) = \langle \psi \mid \mathbf{M}_i^\dagger \mathbf{M}_i \mid \psi\rangle$. The set of operators $\{\mathbf{A}_i\}$ can be used, instead of $\{\mathbf{M}_i\}$, to estimate the probabilities of various measurement outcomes.

Not all the elements of a POVM are measurement operators. Only in the case of projective measurements, all the POVM elements $\mathbf{A}_i$, where

$$\mathbf{A}_i \equiv \mathbf{P}_i^\dagger \mathbf{P}_i = \mathbf{P}_i$$

are the same as the measurement operators.

POVM measurements prevent us from making misidentification errors but sometimes they may not provide any information at all. POVM allows us to distinguish the cases when the

outcome of a measurement identifies with certainty the original state from the ones when the identification is not possible. POVM measurements can be optimal, a topic discussed in more depth in Section 3.12; optimality implies that the average information gained as a result of the measurement is maximized.

**Example.** *Consider a qubit in one of two possible states:*

$$| \psi_1\rangle =| 0\rangle \;\Rightarrow\; | \psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad | \psi_2\rangle = \frac{(| 0\rangle - | 1\rangle)}{\sqrt{2}} \;\Rightarrow\; | \psi_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We define a POVM with three operators (elements). The first two are

$$\mathbf{A}_1 = \frac{\sqrt{2}}{1+\sqrt{2}} | 1\rangle\langle 1 | \quad \Rightarrow \quad \mathbf{A}_1 = \frac{\sqrt{2}}{1+\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \frac{\sqrt{2}}{1+\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

and

$$\mathbf{A}_2 = \frac{\sqrt{2}}{1+\sqrt{2}} \frac{(| 0\rangle + | 1\rangle)(\langle 0 | + \langle 1 |)}{2} \quad \Rightarrow \quad \mathbf{A}_2 = \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix}$$

or

$$\mathbf{A}_2 = \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

To ensure completeness we define a third operator

$$\mathbf{A}_3 = \mathbf{I} - \mathbf{A}_1 - \mathbf{A}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{\sqrt{2}}{1+\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} - \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Thus,

$$\mathbf{A}_3 = \frac{1}{2(1+\sqrt{2})} \begin{pmatrix} 2+\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 2-\sqrt{2} \end{pmatrix}.$$

Now we show that the three possible outcomes of the measurement associated with $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$, lead to the following conclusions regarding the original state of the system:

$\mathbf{A}_1 \;\Rightarrow\;$ the original state was $| \psi_2\rangle$

$\mathbf{A}_2 \;\Rightarrow\;$ the original state was $| \psi_1\rangle$

$\mathbf{A}_3 \;\Rightarrow\;$ do not know if the original state was $| \psi_1\rangle$ or $| \psi_2\rangle$

First, we show that when there is an outcome associated with $\mathbf{A}_1$, then, the probability that the original state was $| \psi_1\rangle$ is $p_{\mathbf{A}_1}(| \psi_1\rangle) = 0$ and $p_{\mathbf{A}_1}(| \psi_2\rangle) > 0$ thus, the system could only have been in state $| \psi_2\rangle$. Indeed,

$$\mathbf{A}_1 | \psi_1\rangle = \frac{\sqrt{2}}{1+\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and

$$p_{\mathbf{A}_1}(\mid \psi_1\rangle) = \langle\psi_1 \mid [\mathbf{A}_1 \mid \psi_1\rangle] = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Also,

$$\mathbf{A}_1 \mid \psi_2\rangle = \frac{\sqrt{2}}{1+\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{1+\sqrt{2}} \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

and

$$p_{\mathbf{A}_1}(\mid \psi_2\rangle) = \langle\psi_2 \mid [A_1 \mid \psi_2\rangle] = \frac{1}{\sqrt{2}(1+\sqrt{2})} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}(1+\sqrt{2})}.$$

When there is an outcome associated with $\mathbf{A}_2$, then the probability that the original state was $\mid \psi_2\rangle$ is equal to zero, $p_{\mathbf{A}_2}(\mid \psi_2\rangle) = 0$ and $p_{\mathbf{A}_2}(\mid \psi_1\rangle) > 0$, therefore the system could only have been in state $\mid \psi_1\rangle$. Indeed,

$$\mathbf{A}_2 \mid \psi_2\rangle = \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{2(1+\sqrt{2})} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and

$$p_{\mathbf{A}_2}(\mid \psi_2\rangle) = \langle\psi_2 \mid [\mathbf{A}_2 \mid \psi_2\rangle] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \frac{1}{2(1+\sqrt{2})} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Also,

$$\mathbf{A}_2 \mid \psi_1\rangle = \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$p_{\mathbf{A}_2}(\mid \psi_1\rangle) = \langle\psi_1 \mid [\mathbf{A}_2 \mid \psi_1\rangle] = \begin{pmatrix} 1 & 0 \end{pmatrix} \frac{\sqrt{2}}{2(1+\sqrt{2})} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}(1+\sqrt{2})}.$$

Finally, when there is an outcome associated with $\mathbf{A}_3$, then there is a non-zero probability that the original state was either $\mid \psi_1\rangle$, or $\mid \psi_2\rangle$; thus $p_{\mathbf{A}_3}(\mid \psi_1\rangle) > 0$ and $p_{\mathbf{A}_3}(\mid \psi_2\rangle) > 0$. Indeed,

$$\mathbf{A}_3 \mid \psi_1\rangle = \frac{1}{2(1+\sqrt{2})} \begin{pmatrix} 2+\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 2-\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}(1+\sqrt{2})} \begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix}$$

$$p_{\mathbf{A}_3}(\mid \psi_1\rangle) = \langle\psi_1 \mid [\mathbf{A}_3 \mid \psi_1\rangle] = \frac{1}{\sqrt{2}(1+\sqrt{2})} \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

Similarly,

$$\mathbf{A}_3 \mid \psi_2\rangle = \frac{1}{2(1+\sqrt{2})} \begin{pmatrix} 2+\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 2-\sqrt{2} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}(1+\sqrt{2})} \begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix}$$

$$p_{\mathbf{A}_3}(|\ \psi_2\rangle) = \langle\psi_2 \mid [\mathbf{A}_3 \mid \psi_2\rangle] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}(1+\sqrt{2})} \begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

Clearly, $0 \leq (p_{\mathbf{A}_3}(|\ \psi_1\rangle), p_{\mathbf{A}_3}(|\ \psi_2\rangle)) \leq 1$ and, in this case,

$$p_{\mathbf{A}_3}(|\ \psi_1\rangle) = p_{\mathbf{A}_3}(|\ \psi_2\rangle) = \frac{1}{\sqrt{2}}.$$

There is never a mistake in identifying the original state when we apply the POVM measurement operators $\mathbf{A}_1$ and $\mathbf{A}_2$. This example shows when we use the POVM element $\mathbf{A}_3$ there is not enough information for a reliable identification.

A clear exposition of POVM can be found in Peres [326]; POVM has been applied to the analysis of quantum key distribution and the analysis of eavesdropping strategy by Ekert and his coworkers [137]. A review of POVM, followed by a discussion of an implementation of POVM for photonic qubits can be found in [71, 72].

We conclude the discussion of projective and POVM measurements with the observation that the main difference between them is the number of available preparations of the system and that of available outcomes may be different from each other and different from the dimension of the Hilbert space. The probability of an outcome $i$ is given by $p_i = \mathrm{tr}(\mathbf{A}_i\rho)$ in the case of POVM, while for a von Neumann measurement is $p_i = \mathrm{tr}(\mathbf{P}_i\rho)$.

## 2.9 Neumark Theorem

Neumark theorem addresses the problem of extending POVM operators from a smaller to a larger Hilbert space; a consequence of this theorem is that a realizable experimental procedure for a POVM measurement with operators $\mathbf{A}_i$ always exists.

To show that we introduce the concept of maximal quantum test: consider a quantum system and assume that the maximum number of different outcomes obtainable in a test of the system is $N$; any test with exactly $N$ different outcomes is a maximal quantum test. For example, when atoms with spin $s$ are used in a Stern-Gerlach experiment $2s+1$ different outcomes are observed, regardless of the orientation of the magnetic field. The "strong superposition principle" [325] states that any complete set of mutually orthogonal vectors has a physical realization as a maximal quantum test. Newmark theorem can be formulated as [6]:

**Theorem.** *The Hilbert space $\mathcal{H}_n$ in which the set $\mathbf{A}_i, 1 \leq i \leq n$, of POVM operators are defined can be extended to a larger Hilbert space, $\mathcal{H}_N$, with $N > n$ such that in $\mathcal{H}_N$ there exist a set of orthogonal projectors $\mathbf{P}_i$ with two properties:*

- $\sum_i \mathbf{P}_i = \mathbf{I}$ *and*
- $\mathbf{A}_i$ *is the result of projecting* $\mathbf{P}_i \in \mathcal{H}_N$ *into* $\mathcal{H}_n$.

*Proof:* we add $N-n$ extra dimensions to $\mathcal{H}_n$ by introducing the unit vectors $\mid v_k\rangle$ orthogonal to each other, $\langle v_j \mid v_k\rangle = \delta_{jk}, \quad (n+1) \leq (j,k) \leq N$, and also orthogonal to all $\mid a_i\rangle$ [325]; then the vectors

$$\mid w_i\rangle = \mid a_i\rangle + \sum_{l=n+1}^{N} \gamma_{il} \mid v_l\rangle \quad 1 \leq i \leq N$$

form a complete orthonormal basis in $\mathcal{H}_N$ if and only if

$$\langle w_i \mid w_j\rangle = \delta_{ij}.$$

This means that

$$\langle a_i \mid a_j\rangle + \sum_{l=n+1}^{N} \gamma_{il}^* \gamma_{jl} = \delta_{ij}.$$

The unknown coefficients $\gamma_{il}$ can be determined from this system of equations. We notice that we have more equations than unknowns; we also notice that $\mid a_i\rangle$ are not arbitrary, they satisfy the closure property $\sum \mid a_i\rangle\langle a_i \mid = \sum \mathbf{A}_i = \mathbf{I}$. We write this system as

$$\sum_{k=1}^{n} \alpha_{ik}^* \alpha_{jk} + \sum_{l=n+1}^{N} \gamma_{il}^* \gamma_{j,l} = \delta_{ij} \qquad (i,j = 1,\ldots,N).$$

We now consider the square matrix of order $N$

$$M = \begin{pmatrix} \alpha_{p1} & \ldots & \alpha_{pn} & \gamma_{p(n+1)} & \ldots \gamma_{pN} \\ \alpha_{q1} & \ldots & \alpha_{qn} & \gamma_{q(n+1)} & \ldots \gamma_{qN} \\ \vdots & & \vdots & \vdots & \vdots \\ \alpha_{N1} & \ldots & \alpha_{Nn} & \gamma_{N(n+1} & \ldots \gamma_{NN} \end{pmatrix}.$$

The first $n$ columns of the matrix $M$ consist of the $\alpha_{ik}$ given to us, and the last $N-n$ columns are the unknown $\gamma_{il}$. This matrix is unitary, the first $n$ columns can be considered $n$ orthonormal vectors in an $N$-dimensional space and the remaining $(N-n)$ column vectors can be constructed in infinitely many ways. The $N$ column vectors allow us to construct the orthonormal vectors $\mid w_i\rangle$ and then the projections of $\mid w_i\rangle$ to $\mathcal{H}_n$ gives us the $\mid a_i\rangle$.

In the general case, the extension from $\mathcal{H}_n$ to $\mathcal{H}_N$ requires the introduction of ancilla qubits. Neumark theorem proves that an arbitrary POVM with a finite number of elements can, in principle, be converted to a maximal quantum test by using the ancilla, an auxiliary independently prepared quantum system.

## 2.10 Gleason Theorem

A measurement provides additional information about a quantum system in a pure or in a mixed state and it affects the density operator of the system. The post-measurement state is characterized by new probabilities and it is different from the state of the quantum system prior to the measurement.

We now consider $N$ independent versions of a quantum system, each in a different state $\mid \psi_i\rangle$, $1 \leq i \leq N$. Gleason theorem allows us to compute the ensemble average of the operator $\mathbf{A}$ after we measure the same observable $A$ on all $N$ versions of the system [2].

**Theorem.** *The ensemble average of the observable $A$ is equal to the trace of the product of the measurement operator $\mathbf{A}$ and the density operator $\rho$*

$$\langle \mathbf{A} \rangle = \mathrm{tr}(\mathbf{A}\rho).$$

*Proof:* Let $\{\mid a_j\rangle\}$ denote the set of vectors of an orthonormal basis and assume that these orthogonal vectors are the eigenstates of a self-adjoint operator. The ensemble average, $\langle \mathbf{A} \rangle$, of the measurements performed on the observable $A$ is defined as

$$\langle \mathbf{A} \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \psi_i \mid \mathbf{A} \mid \psi_i \rangle.$$

If the system is initially in state $\mid \psi_i\rangle$, as a result of applying the measurement operator $\mathbf{A}$ it ends up in state $\mid \varphi_i\rangle = \mathbf{A} \mid \psi_i\rangle$. Thus, the ensemble average can be expressed as the average of inner products of the states before and after the measurement

$$\langle \mathbf{A} \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \psi_i \mid \varphi_i \rangle.$$

The vectors of the orthonormal basis $\{\mid a_j\rangle\}$ and the corresponding projectors satisfy the completeness condition $\sum_j \mid a_j\rangle\langle a_j \mid = \mathbf{I}$. Then we rewrite

$$\langle \psi_i \mid \mathbf{A} \mid \psi_i \rangle = \langle \psi_i \mid \mathbf{IA} \mid \psi_i \rangle = \langle \psi_i \left[ \sum_j \mid a_j\rangle\langle a_j \mid \mathbf{A} \right] \mid \psi_i \rangle = \sum_j \langle \psi_i \mid a_j\rangle\langle a_j \mid \mathbf{A} \mid \psi_i \rangle,$$

or

$$\langle \psi_i \mid \mathbf{A} \mid \psi_i \rangle = \sum_j \langle a_j \mid \mathbf{A} \mid \psi_i \rangle\langle \psi_i \mid a_j \rangle.$$

Note that in the previous expression $\langle \psi_i \mid a_j \rangle$ is a complex number, the projection of $\mid \psi_i\rangle$ on the basis vector $\mid a_j\rangle$. It follows that

$$\langle \mathbf{A} \rangle = \frac{1}{N} \sum_{i=1}^{N} \sum_j \langle a_j \mid \mathbf{A} \mid \psi_i \rangle\langle \psi_i \mid a_j \rangle = \sum_j \langle a_j \mid \mathbf{A} \frac{1}{N} \sum_{i=1}^{N} \mid \psi_i\rangle\langle \psi_i \mid a_j \rangle.$$

When $p_i = 1/N$ the density operator $\rho$ is

$$\rho = \frac{1}{N} \sum_{i=1}^{N} \mid \psi_i\rangle\langle \psi_i \mid .$$

Then the ensemble average $\langle \mathbf{A} \rangle$ can be written as

$$\langle \mathbf{A} \rangle = \sum_j \langle a_j \mid \mathbf{A}\rho \mid a_j \rangle = \mathrm{tr}(\mathbf{A}\rho).$$

To complete the proof we recall that $\mathrm{tr}(\mathbf{A}\rho) = \mathrm{tr}(\rho\mathbf{A})$.

**Example.** *Consider a mixed ensemble of quantum states in a two-dimensional Hilbert space with orthonormal basis vectors* $\mid 0\rangle$ *and* $\mid 1\rangle$.

The states and their respective occurrence probabilities in the mixed ensemble are:

| State | Probability |
|---|---|
| $\mid \psi_A\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ | $p$ |
| $\mid \psi_B\rangle = \beta_0 \mid 0\rangle + \beta_1 \mid 1\rangle$ | $1-p$ |

The density operator of this mixed ensemble of states is

$$\rho = p \mid \psi_A\rangle\langle\psi_A \mid + (1-p) \mid \psi_B\rangle\langle\psi_B \mid .$$

A measurement on this system is specified by the projectors

$$\mathbf{P}_0 = \mid 0\rangle\langle 0 \mid \quad \text{and} \quad \mathbf{P}_1 = \mid 1\rangle\langle 1 \mid .$$

The two possible outcomes of a measurement of the system in the mixed state are 0 and 1. In a series of such measurements we expect to obtain the outcome 0 with probability:

$$p(0) = \text{Prob}(\mid \psi_A\rangle)\text{Prob}(0, \mid \psi_A\rangle) + \text{Prob}(\mid \psi_B\rangle)\text{Prob}(0, \mid \psi_B\rangle).$$

In this expression $\text{Prob}(\mid \psi_A\rangle)$ is the probability of state $\mid \psi_A\rangle$ in the mix, $\text{Prob}(0, \mid \psi_A\rangle)$ is the conditional probability of observing the outcome 0 in state $\mid \psi_A\rangle$ and it is equal to $\langle\psi_A \mid \mathbf{P}_0 \mid \psi_A\rangle$; similarly, $\text{Prob}(0, \mid \psi_B\rangle)$ is the conditional probability of observing the outcome 0 in state $\mid \psi_B\rangle$ and it is equal to $\langle\psi_B \mid \mathbf{P}_0 \mid \psi_B\rangle$; but $\text{Prob}(\mid \psi_A\rangle) = p$ and $\text{Prob}(\mid \psi_B\rangle) = 1-p$ thus,

$$p(0) = p\langle\psi_A \mid \mathbf{P}_0 \mid \psi_A\rangle + (1-p)\langle\psi_B \mid \mathbf{P}_0 \mid \psi_B\rangle$$

with $\text{Prob}(0, \mid \psi_A\rangle) = \mid \alpha_0 \mid^2$ and $\text{Prob}(0, \mid \psi_B\rangle) = \mid \beta_0 \mid^2$. It follows that

$$p(0) = p \mid \alpha_0 \mid^2 + (1-p) \mid \beta_0 \mid^2 .$$

According to Gleason theorem ensemble-averaged quantities such as $p(0)$ above can be represented in terms of the density operator $\rho$ as

$$\begin{aligned} p(0) &= \text{tr}[\rho\mathbf{P}_0] \\ &= \text{tr}\,[\mathbf{P}_0\rho]. \end{aligned} .$$

Indeed,

$$\begin{aligned} \text{tr}\,[\rho\mathbf{P}_0] &= \langle 0 \mid \rho\mathbf{P}_0 \mid 0\rangle + \langle 1 \mid \rho\mathbf{P}_0 \mid 1\rangle \\ &= \langle 0 \mid \rho \mid 0\rangle \\ &= \langle 0 \mid [p \mid \psi_A\rangle\langle\psi_A \mid + (1-p) \mid \psi_B\rangle\langle\psi_B \mid ] \mid 0\rangle \\ &= p\,\langle 0 \mid \psi_A\rangle\langle\psi_A \mid 0\rangle + (1-p)\,\langle 0 \mid \psi_B\rangle\langle\psi_B \mid 0\rangle \\ &= p \mid \alpha_0 \mid^2 + (1-p) \mid \beta_0 \mid^2 \\ &= p(0). \end{aligned}$$

The density operator plays an important role in Quantum Information Theory, notably in the characterization of the noise inevitable in quantum communication. Consider quantum

states affected by noise and characterized by vectors that are not necessarily normalized in a Hilbert space, e.g., $\mid \widetilde{\psi}_i\rangle$ and $\mid \widetilde{\varphi}_i\rangle$. The relationship among quantum states with the same density operator is captured by the following proposition.

*Consider two sets of mixed states* $\mid \widetilde{\psi}_i\rangle$ *and* $\mid \widetilde{\varphi}_i\rangle$ *that may not be normalized to unit length; the two sets are extended to have the same number of component states by padding the one with fewer states with additional zero vectors. Let* $U = [u_{ij}]$ *be a unitary matrix of complex numbers.* $\mid \widetilde{\psi}_i\rangle$ *and* $\mid \widetilde{\varphi}_i\rangle$ *generate the same density operator if and only if:*

$$\mid \widetilde{\psi}_i\rangle = \sum_j u_{ij} \mid \widetilde{\varphi}_j\rangle.$$

*Proof:* We only prove one side of the proposition, namely that if $\mid \widetilde{\psi}_i\rangle = \sum_j u_{ij} \mid \widetilde{\varphi}_j\rangle$ then the two vectors characterize states with the same density operator.

$$\sum_i \mid \widetilde{\psi}_i\rangle\langle\widetilde{\psi}_i \mid = \sum_{i,j,k} u_{ij}u_{ik}^* \mid \widetilde{\varphi}_j\rangle\langle\widetilde{\varphi}_k \mid = \sum_{j,k}\left(\sum_i u_{ki}^* u_{ij}\right) \mid \widetilde{\varphi}_j\rangle\langle\widetilde{\varphi}_k \mid .$$

Thus,

$$\sum_i \mid \widetilde{\psi}_i\rangle\langle\widetilde{\psi}_i \mid = \sum_{j,k} \delta_{k,j} \mid \widetilde{\varphi}_j\rangle\langle\widetilde{\varphi}_k \mid = \sum_j \mid \widetilde{\varphi}_j\rangle\langle\widetilde{\varphi}_j \mid .$$

**Corollary.** *Given a Hilbert space* $\mathcal{H}_n$ *of dimension* $n \geq 3$*, the only possible measure of the probability of the state associated with a particular linear subspace* $\mathcal{H}_m$*, where* $m < n$*, of the Hilbert space* $\mathcal{H}_n$ *is the trace of the operator product of the projection operator* $\mathbf{E}$ *that projects* $\mathcal{H}_n$ *onto* $\mathcal{H}_m$ *and the density matrix* $\rho$ *of the system,* $tr[\rho\mathbf{E}]$.

The corollary makes the assumption that in Quantum Mechanics we should be able to measure the events and assign probabilities to all possible orthogonal projections in a Hilbert space $\mathcal{H}$. The projectors are unitary, $\mathbf{E}^2 = \mathbf{E}$ and self-adjoint $\mathbf{E} = \mathbf{E}^\dagger$. The state of a quantum system is a mapping of each projection to a real number between zero and one, the probability of a particular outcome of the measurement:

$$\mathbf{E} \mapsto p\,(\mathbf{E})\,; \quad 0 \leq p\,(\mathbf{E}) \leq 1.$$

This mapping must have three properties:

1. $\mathrm{p}(\mathbf{0}) = 0$.
2. $\mathrm{p}(\mathbf{1}) = \mathrm{I}$.
3. $\mathbf{E}_1\mathbf{E}_2 = 0 \implies p\,(\mathbf{E}_1 + \mathbf{E}_2) = p\,(\mathbf{E}_1) + p\,(\mathbf{E}_2)$

The first and second properties give the probabilities for the projectors on the null subspace and the entire space, $\mathbf{0}$ and $\mathbf{1}$ respectively; the last property reflects the fact that probabilities assigned to mutually orthogonal spaces must be additive; this makes sense because mutually orthogonal subspaces can be considered exclusive alternatives.

Gleason theorem is an important result in quantum logic; quantum logic treats quantum events as logical propositions, and studies the relationships and structures formed by these events. Formally, a quantum logic is a set of events that is closed under a countable disjunction of countably many mutually exclusive events. The representation theorem in quantum logic shows that these logics form a lattice which is isomorphic to the lattice of subspaces of a vector space with a scalar product[18]. The representation theorem allows us to treat quantum events as a lattice $\mathcal{L} = \mathcal{L}(\mathcal{H})$ of subspaces of a real or complex Hilbert space $\mathcal{H}$. Gleason theorem allows us to assign probabilities to these events. Gleason theorem [172] also shows that the Born rule for the probability of obtaining a specific outcome of a given measurement, follows naturally from the structure formed by the lattice of events in a real or complex Hilbert space.

## 2.11 Mixed Ensembles and their Time Evolution

We now discuss the effect of a measurement of a mixed ensemble on its density operator and show that the time evolution of mixed ensembles can be expressed through the time evolution of its density operator. The density operator $\rho$, or the quantum state vector $\mid \psi\rangle$, are *mathematical objects* that represent our knowledge about the preparation of a physical system. They are not 'physical' objects in the way hydrodynamic waves or electromagnetic wavepackets are, though we could be tempted to interpret them as such. In this mathematical framework we can think of pure states $\mid \psi\rangle$ as representing "minimal uncertainty" states, while a density operator $\rho$ represents the additional uncertainties, the probabilities, associated with a mixed ensemble.

We wish to express the probability of an outcome of the measurement knowing the density operator of the initial state of a system and we also wish to determine the density operator of the resulting state when one of the projectors is used for the measurement.

**Proposition.** *Consider a mixed ensemble of states $\mid \psi_i\rangle$ with the corresponding probabilities, $p_i$, and a density operator $\rho = \sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid$. Call $p(k)$ the probability of obtaining the outcome with index $k$ and $\rho_k$ the post-measurement density operator of the mixed ensemble in a measurement of an observable corresponding to the projector $\mathbf{P}_k$; then*

$$p(k) = \text{tr}\,[\rho\,\mathbf{P}_k] \quad \text{and} \quad \rho_k = \frac{\mathbf{P}_k\,\rho\,\mathbf{P}_k}{\text{tr}\,[\rho\,\mathbf{P}_k]}.$$

*Proof:* $p_i = \text{Prob}\,(\mid \psi_i\rangle)$ denotes the probability of state $\mid \psi_i\rangle$; call $\text{Prob}\,(k, \mid \psi_i\rangle)$ the probability of observing outcome $k$ of a measurement of state $\mid \psi_i\rangle$ and recall that this probability is $\langle\psi_i \mid \mathbf{P}_k \mid \psi_i\rangle$, thus

$$p(k) = \sum_i \text{Prob}\,(\mid \psi_i\rangle)\ \text{Prob}\,(k, \mid \psi_i\rangle) = \sum_i p_i\langle\psi_i \mid \mathbf{P}_k \mid \psi_i\rangle = \sum_i p_i\langle\psi_i \mid \mathbf{P}_k\mathbf{I} \mid \psi_i\rangle.$$

[18] A necessary result for Gleason theorem to be applicable is that the field K over which the vector space is defined, is either the real numbers, complex numbers, or the quaternions; the definition of the inner product of a non-zero vector with itself will satisfy the requirements to make the vector space a Hilbert space for real numbers, complex numbers, or the quaternions.

We use the fact that $\mathbf{I} = \sum_j \mid j\rangle\langle j \mid$ and obtain

$$p(k) = \sum_i p_i \langle \psi_i \mid \mathbf{P}_k \sum_j (\mid j\rangle\langle j \mid) \mid \psi_i\rangle = \sum_j \langle j \mid \rho \mathbf{P}_k \mid j\rangle = \text{tr } [\rho \ \mathbf{P}_k].$$

Now we concentrate on the second relationship and denote by $\mid \varphi'_{(i,k)}\rangle$ the normalized state resulting when the projector, or measurement operator, $\mathbf{P}_k$ is applied to state $\mid \psi_i\rangle$ of the mixed ensemble. Call $p(i \mid k)$ the conditional probability of applying $\mathbf{P}_k$ to the state $\mid \psi_i\rangle$ of the mixed ensemble; then the density operator $\rho_k$ of the mixed ensemble after the application of $\mathbf{P}_k$ is

$$\rho_k = \sum_i p(i \mid k) \mid \varphi'_{(i,k)}\rangle\langle \varphi'_{(i,k)} \mid .$$

The normalized state resulting when $\mathbf{P}_k$ is applied to state $\mid \psi_i\rangle$ of the mixed ensemble is

$$\mid \varphi'_{(i,k)}\rangle = \frac{\mathbf{P}_k \mid \psi_i\rangle}{\sqrt{\langle \psi_i \mid \mathbf{P}_k \mid \psi_i\rangle}}.$$

The relation between conditional probabilities allow us to express

$$p(i \mid k) = \frac{p(k \mid i) p_i}{p(k)}$$

with the probability of outcome with index $k$, $p(k) = \text{tr } [\rho \ \mathbf{P}_k]$ and the conditional probability of applying $\mathbf{P}_k$ to the state $\mid \psi_i\rangle$ of the mixed ensemble, $p(i \mid k) = \langle \psi_i \mid \mathbf{P}_k \mid \psi_i\rangle$. Here $p(k \mid i)$ is the conditional probability of applying $\mathbf{P}_i$ to the state $\mid \psi_k\rangle$ of the mixed ensemble.

The density operator of the mixed ensemble after the application of $\mathbf{P}_k$ is

$$\rho_k = \sum_i p_i \frac{\langle \psi_i \mid \mathbf{P}_k \mid \psi_i\rangle}{\text{tr } [\rho \ \mathbf{P}_k]} \times \frac{\mathbf{P}_k \mid \psi_i\rangle\langle \psi_i \mid \mathbf{P}_k}{\langle \psi_i \mid \mathbf{P}_k \mid \psi_i\rangle} = \frac{\mathbf{P}_k [\sum_i p_i \mid \psi_i\rangle\langle \psi_i \mid] \mathbf{P}_k}{\text{tr } [\rho \ \mathbf{P}_k]}.$$

Thus,

$$\rho_k = \frac{\mathbf{P}_k \rho \mathbf{P}_k}{\text{tr } [\rho \ \mathbf{P}_k]}.$$

Mixed ensembles, such as $\{p_i^A, \mid \psi_i^A\rangle\}$ with density operator $\rho_A$ and $\{p_i^B, \mid \psi_i^B\rangle\}$ with density operator $\rho_B$, can be combined together. The states $\mid \psi_A\rangle$ and $\mid \psi_B\rangle$ in the combined ensemble occur with probabilities $pp_i^A$ and $(1-p)p_i^B$, respectively. The density operator $\rho$ of the combined ensemble is

$$\rho \ = \ p \ \rho_A \ + \ (1-p) \ \rho_B.$$

Here, $p$ and $(1-p)$, with $0 \ \leq \ p \ \leq \ 1$ are the probability of occurrence in the combined ensemble of the individual mixed ensembles $A$ and $B$, respectively.

Recall that the Schrödinger equation describes the evolution in time of the mathematical objects describing the system preparation, given that we know the Hamiltonian of the system. When we perform a measurement on a quantum system, we gain new information about that system and our initial description of it needs to change instantaneously in order to make

possible the prediction of further measurements. According to the Schrödinger equation, a unitary time evolution operator, $\mathbf{T}$ can be used to express the evolution in time of a pure quantum state

$$| \psi(t)\rangle \;=\; \mathbf{T}(t,0) \mid \psi(0)\rangle \qquad \text{and} \qquad \langle\psi(t) \mid \;=\; \langle\psi(0) \mid \mathbf{T}(0,t).$$

According to its definition, the density operator at time $t$ can be written as

$$\begin{aligned}
\rho(t) &= \textstyle\sum_i p_i \;\; \mid \psi_i(t)\rangle\langle\psi_i(t) \mid \\
&= \textstyle\sum_i p_i \mathbf{T}(t,0) \mid \psi_i(0)\rangle\langle\psi_i(0) \mid \mathbf{T}(0,t) \\
&= \mathbf{T}(t,0) \textstyle\sum_i p_i \;\; \mid \psi_i(0)\rangle\langle\psi_i(0) \mid \mathbf{T}(0,t) \\
&= \mathbf{T}(t,0)\; \rho(0)\; \mathbf{T}(0,t).
\end{aligned}$$

Now, the trace of the density operator can be written as

$$\text{tr}\;[\rho(t)] = \text{tr}\;[\mathbf{T}(t,0)\rho(0)\mathbf{T}(0,t)] = \text{tr}\;[\mathbf{T}(0,t)\mathbf{T}(t,0)\rho(0)] = \text{tr}\;[\rho(0)],$$

where $\mathbf{T}(0,t)\mathbf{T}(t,0) = I$. We see that the trace of the density operator is preserved under Hamiltonian evolution; the Hamiltonian evolution preserves the eigenvalues of the density operator.

## 2.12 Bipartite Systems; Schmidt Decomposition

Bipartite systems are of interest in quantum information theory because quantum systems interact with one another and with the environment; such interactions affect the information encoded in the quantum state. For example, a quantum system $\mathcal{A}$ in a pure state may interact with the environment and as a result of this interaction the state of the system may become a mixed state; the system $\mathcal{A}$ may be looked at as part of a bipartite system, $\mathcal{A}$ + environment. The formalism to describe bipartite systems is based on Schmidt decomposition.

**Proposition.** *Let $| \psi_{\mathcal{C}}\rangle$ be the state of a bipartite system $\mathcal{C} = \mathcal{AB}$ with $\mathcal{A} \in \mathcal{H}_A$ and $\mathcal{B} \in \mathcal{H}_B$; there are two orthonormal bases, called Schmidt bases $\{| e_j^A\rangle\} \in \mathcal{H}_A$ and $\{| e_j^B\rangle\} \in \mathcal{H}_B$ such that*

$$| \psi_{\mathcal{C}}\rangle = \sum_j \lambda_j \mid e_j^A\rangle\otimes \mid e_j^B\rangle.$$

In this expression $\lambda_j$ are called Schmidt coefficients; $\lambda_j$ are real numbers and $\sum_j \lambda_j^2 = 1$.

*Proof:* When $\dim[\mathcal{H}_A] = \dim[\mathcal{H}_B] = n$ then we can express the state of $\mathcal{C} \in \mathcal{H}_A \otimes \mathcal{H}_B$ as a superposition

$$| \psi_{\mathcal{C}}\rangle = \sum_{i=1}^{n}\sum_{k=1}^{n} \gamma_{ik} \mid u_i\rangle\otimes \mid v_k\rangle$$

using the orthonormal bases $\{\mid u_j\rangle\} \in \mathcal{H}_A$ and $\{\mid v_j\rangle\} \in \mathcal{H}_B, 1 \leq j \leq n$.

The complex coefficients $\gamma_{ik},\ 1 \leq i,k \leq n$ can be viewed as elements of a matrix $\Gamma = [\gamma_{ik}]$. The singular value decomposition theorem states that an $n \times n$ matrix $\Gamma$ over a field $F$ of real or complex numbers, can be decomposed as a product of three $n \times n$ matrices over $F$: $\Gamma = A\Delta B^\dagger$ with: $A = [\alpha_{ij}]$ and $B = [\beta^*_{ij}]$ unitary matrices, and $\Delta = [\delta_{ij}]$ a diagonal matrix. This allows us to write the previous expression as

$$\mid \psi_{\mathcal{C}}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_{ij}\delta_{jj}\beta_{jk} \mid u_j\rangle \otimes \mid v_k\rangle.$$

We denote

$$\mid e_i^A\rangle = \sum_{j=1}^{n} \alpha_{ij} \mid u_j\rangle, \quad \mid e_j^B\rangle = \sum_{k=1}^{n} \beta_{jk} \mid v_k\rangle, \quad \text{and} \quad \lambda_j = \delta_{jj}.$$

Recall that $\{\mid u_j\rangle\}$ form an orthonormal basis in $\mathcal{H}_A$ and that $A$ is unitary. It is easy to see that the vectors $\{\mid e_i^A\rangle\}$ form an orthonormal basis in $\mathcal{H}_A$

$$\langle e_i^A \mid e_k^A\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha^*_{ji}\alpha_{ik}\langle u_j \mid u_k\rangle = \delta_{jk}$$

because $\sum_i \mid \alpha_{ii} \mid^2 = 1$. Similarly the vectors $\{\mid e_j^B\rangle\}$ form an orthonormal basis in $\mathcal{H}_B$. Thus,

$$\mid \psi_{\mathcal{C}}\rangle = \sum_j \lambda_j \mid e_j^A\rangle \otimes \mid e_j^B\rangle.$$

The proof for the case when the two subsystems have Hilbert spaces with different dimensions is more elaborate.

Consider a bipartite pure state $\mid \psi_{\mathcal{C}}\rangle$ of a composite system $\mathcal{C} = \mathcal{AB}$. The number of non-zero eigenvalues $\lambda_j$ of $\rho^{\mathcal{C}}$, thus the number of terms in the Schmidt decomposition of $\mid \psi_{\mathcal{C}}\rangle$ is called the Schmidt number of state $\mid \psi_{\mathcal{C}}\rangle$. The Schmidt number is a measure of the degree of entanglement of the two subsystems.

A separable bipartite pure state $\mid \varphi_{\mathcal{AB}}\rangle$ over the Hilbert space $\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ is a tensor product of two pure states $\mid \zeta_{\mathcal{A}}\rangle \in \mathcal{H}_{\mathcal{A}}$ and $\mid \xi_{\mathcal{A}}\rangle \in \mathcal{H}_{\mathcal{B}}$

$$\mid \varphi_{\mathcal{AB}}\rangle = \mid \zeta_{\mathcal{A}}\rangle \otimes \mid \xi_{\mathcal{B}}\rangle.$$

It follows that the reduced density matrices are pure

$$\rho^{\mathcal{A}} = \mid \zeta_{\mathcal{A}}\rangle\langle\zeta_{\mathcal{A}} \mid \quad \text{and} \quad \rho^{\mathcal{B}} = \mid \xi_{\mathcal{B}}\rangle\langle\xi_{\mathcal{B}} \mid .$$

The concepts "separability," "correlation," and "entanglement" should be correctly understood in the following sense [343]:

- Two subsystems $\mathcal{A}$ and $\mathcal{B}$ can be correlated even if they are separable; for example, two spins in the state

$$\mid \varphi_{\mathcal{AB}}\rangle = \mid\uparrow_{\mathcal{A}}\rangle \otimes \mid\uparrow_{\mathcal{B}}\rangle$$

are separable, but correlated, both spins are up. This state can be created by acting individually on the two subsystems; the two subsystems need not be at any instant of time together, we can only send a message to the two individuals preparing the two systems that each one of them should have the spin up.

- Entanglement is a special type of correlation, the two correlated systems are not separable. The entanglement cannot be created locally at the two sites where the subsystems $\mathcal{A}$ and $\mathcal{B}$ are located; the two subsystems $\mathcal{A}$ and $\mathcal{B}$ must interact directly with each other to become entangled; the state

$$\mid \varphi_{\mathcal{AB}}\rangle = \frac{1}{2}\left(|\uparrow_{\mathcal{A}}\rangle \otimes |\uparrow_{\mathcal{B}}\rangle + |\downarrow_{\mathcal{A}}\rangle \otimes |\downarrow_{\mathcal{B}}\rangle\right)$$

  requires a unitary transformation to be applied to the two qubits that have to be brought to some place and allowed to interact.

We conclude this discussion with the observation that the Schmidt number provides a clear distinction between bipartite pure states $\mid \psi_{\mathcal{AB}}\rangle$ that are separable and those that are entangled: *a bipartite pure state* $\mid \psi_{\mathcal{AB}}\rangle$ *is entangled if its Schmidt number is greater than one, otherwise it is separable.*

## 2.13 Measurements of Bipartite Systems

We continue our discussion of bipartite quantum systems and explore the question how to gather information about one of the subsystems, $\mathcal{A}$, from measurements performed on the bipartite system, $\mathcal{C} = \mathcal{AB}$ [44, 308]. The bipartite system may be in a product state when $\mid \psi_{\mathcal{C}}\rangle = \mid \psi_{\mathcal{A}}\rangle \otimes \mid \psi_{\mathcal{B}}\rangle$ with $\mid \psi_{\mathcal{A}}\rangle$ the state of subsystem $\mathcal{A}$ and $\mid \psi_{\mathcal{B}}\rangle$ the state of subsystem $\mathcal{B}$, or in a state with a some degree of entanglement between the two subsystems when $\mid \psi_{\mathcal{C}}\rangle \neq \mid \psi_{\mathcal{A}}\rangle \otimes \mid \psi_{\mathcal{B}}\rangle$, see Figure 52. The following proposition allows us to find the average value of an observable of one subsystem only.

**Proposition.** *Let* $\mid \psi_{\mathcal{C}}\rangle$ *be the current state of a bipartite system* $\mathcal{C}$ *consisting of two subsystems* $\mathcal{A}$ *and* $\mathcal{B}$*; we measure an observable* $M$ *of the subsystem* $\mathcal{A}$*. Then the average value of the corresponding measurement operator* $\mathbf{M}^{\mathcal{A}}$ *is*

$$\langle \mathbf{M}^{\mathcal{A}} \rangle = \langle \psi_{\mathcal{C}} \mid \mathbf{M}^{\mathcal{A}} \otimes \mathbf{I}^{\mathcal{B}} \mid \psi_{\mathcal{C}} \rangle.$$

This follows immediately if we think about the measurement as applying the operator $\mathbf{M}^{\mathcal{A}} \otimes \mathbf{I}^{\mathcal{B}}$ to the bipartite system; $\mathbf{I}^{\mathcal{B}}$ is the identity operator applied to subsystem $\mathcal{B}$.

**Example.** *The measurement of a bipartite system of* $1/2$*-spin particles.* If $\mathcal{C}$ is a system of two entangled particles and each subsystem can be described using the basis vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ then $\mid \psi_{\mathcal{C}}\rangle = \alpha_0 \, |\uparrow\uparrow\rangle + \alpha_1 \, |\downarrow\downarrow\rangle$. The expected value of the observable $O$ for one of the two particles, without specifying which one, is

$$\langle O \rangle = \langle \psi_{\mathcal{C}} \mid \mathbf{O} \otimes \mathbf{I} \mid \psi_{\mathcal{C}} \rangle = \mid \alpha_0 \mid^2 \left[\langle\uparrow\mid \mathbf{O} \mid\uparrow\rangle\right] + \mid \alpha_1 \mid^2 \left[\langle\downarrow\mid \mathbf{O} \mid\downarrow\rangle\right] = \mathrm{tr}(\rho\mathbf{O}).$$

where the density operator is given by $\rho = \mid \alpha_0 \mid^2 |\uparrow\rangle\langle\uparrow| + \mid \alpha_1 \mid^2 |\downarrow\rangle\langle\downarrow|$. The operators $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ are projectors on the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$. The result is similar to the case where any of the two particles is detected with probability $\mid \alpha_0 \mid^2$ or $\mid \alpha_1 \mid^2$ to be in state $|\uparrow\rangle$ or $|\downarrow\rangle$, respectively. If we repeat the experiment many times the result is also similar to that obtained for an ensemble of particles, some in state $|\uparrow\rangle$, some in state $|\downarrow\rangle$.

Let $\mathcal{C} = \mathcal{AB}$ be a bipartite system consisting of subsystems $\mathcal{A}$ and $\mathcal{B}$ and described by the density operator $\rho_{\mathcal{C}}$. The <u>partial trace of $\rho_{\mathcal{C}}$ over system $\mathcal{B}$</u> is

$$\mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{C}}\right] = \mathrm{tr}_{\mathcal{B}}\left[\ \mid a_1\rangle\langle a_2 \mid \otimes \mid b_1\rangle\langle b_2 \mid\ \right] = \mid a_1\rangle\langle a_2 \mid\ \mathrm{tr}_{\mathcal{B}}\left[\ \mid b_1\rangle\langle b_2 \mid\ \right] = \mid a_1\rangle\langle a_2 \mid \langle b_1 \mid b_2\rangle.$$

with $\mid a_1\rangle, \mid a_2\rangle$ any two vectors in the state space of $\mathcal{A}$ and $\mid b_1\rangle, \mid b_2\rangle$ any two vectors in the state space of $\mathcal{B}$. The <u>reduced density operator of subsystem $\mathcal{A}$</u> is defined as $\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{C}}\right]$.

**Proposition.** *Let $\mathcal{C} = \mathcal{AB}$ be a bipartite system in a product state with $\rho_{\mathcal{A}}$ the density operator of subsystem $\mathcal{A}$, $\rho_{\mathcal{B}}$ the density operator of subsystem $\mathcal{B}$, and $\rho_{\mathcal{C}} = \rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}$. Then the reduced density operator of each subsystem is equal to the density operator of the subsystem:*

$$\rho^{\mathcal{A}} = \rho_{\mathcal{A}} \quad \text{and} \quad \rho^{\mathcal{B}} = \rho_{\mathcal{B}}.$$

*Proof:* Recall that the trace of the density operator of a system is equal to unity:

$$\mathrm{tr}_{\mathcal{A}}\left[\rho_{\mathcal{A}}\right] = 1 \quad \text{and} \quad \mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{B}}\right] = 1.$$

According to the definition of the reduced density operator:

$$\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{C}}\right] = \mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}\right] = \rho_{\mathcal{A}}\mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{B}}\right] = \rho_{\mathcal{A}},$$

$$\rho^{\mathcal{B}} = \mathrm{tr}_{\mathcal{A}}\left[\rho_{\mathcal{C}}\right] = \mathrm{tr}_{\mathcal{A}}\left[\rho_{\mathcal{A}} \otimes \rho_{\mathcal{B}}\right] = \rho_{\mathcal{B}}\mathrm{tr}_{\mathcal{A}}\left[\rho_{\mathcal{A}}\right] = \rho_{\mathcal{B}}.$$

This result reflects also our intuition; if indeed both the density operator $\rho_{\mathcal{A}}$ and the reduced density operator $\rho^{\mathcal{A}}$ characterize the same state of the system $\mathcal{A}$, then the average of an observable should be the same regardless whether it is computed using $\rho_{\mathcal{A}}$, or $\rho^{\mathcal{A}}$. This is true only for product states; if the system is in an entangled state we expect that the interaction of the two subsystems will affect the outcome of a measurement in a more subtle manner, Figure 52.

**Proposition.** *Let $M^{\mathcal{A}}$ be an observable of $\mathcal{A}$ and $M^{\mathcal{C}}$ the corresponding observable of $\mathcal{A}$ but performed on $\mathcal{C} = \mathcal{AB}$. Let $\mathbf{I}^{\mathcal{B}}$ be the identity operator in the state space of $\mathcal{B}$ and $\rho_{\mathcal{A}}$ and $\rho_{\mathcal{C}}$ be the density operators for the two systems, respectively. Then, the measurement operator for the observable $M^{\mathcal{C}}$ is*

$$\mathbf{M}^{\mathcal{C}} = \mathbf{M}^{\mathcal{A}} \otimes \mathbf{I}^{\mathcal{B}}.$$

*Proof:* Let $\mid \varphi\rangle$ be any state of $\mathcal{B}$; let $\mid m\rangle$ be an eigenstate of $\mathbf{M}^{\mathcal{A}}$ with eigenvalue $m$ and $\mathcal{P}_m$ the projector with outcome $m$. Then the spectral decomposition of the measurement operator $\mathbf{M}^{\mathcal{A}}$ for the observable $M^{\mathcal{A}}$ is

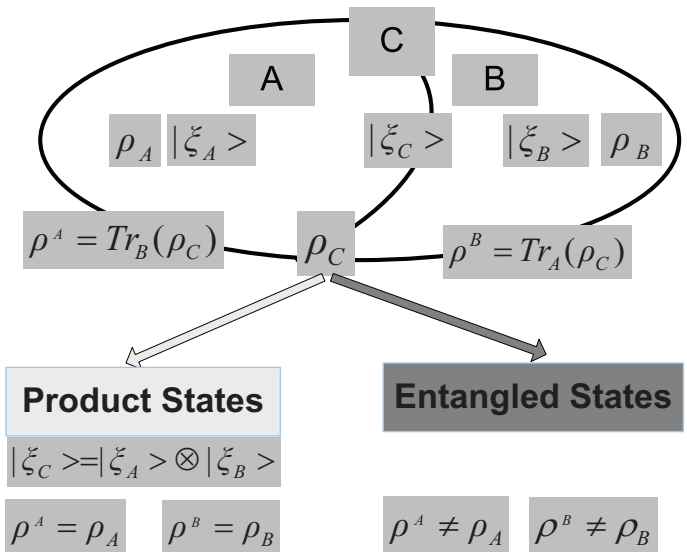



Figure 52: $\mathcal{C}$ is a bipartite system consisting of systems $\mathcal{A}$ and $\mathcal{B}$. The state and the density operators of the three systems are respectively: $\mid \xi_{\mathcal{C}}\rangle$ and $\rho_{\mathcal{C}}$, $\mid \xi_{\mathcal{A}}\rangle$ and $\rho_{\mathcal{A}}$, $\mid \xi_{\mathcal{B}}\rangle$ and $\rho_{\mathcal{B}}$. The bipartite system can be either in a product state, or in an entangled state. $\mathcal{C}$ is in a product state if $\mid \xi_{\mathcal{C}}\rangle = \mid \xi_{\mathcal{A}}\rangle \otimes \mid \xi_{\mathcal{B}}\rangle$; in this case the reduced density operators satisfy the equalities $\rho^{\mathcal{A}} = \rho_{\mathcal{A}}$ and $\rho^{\mathcal{B}} = \rho_{\mathcal{B}}$.

$$\mathbf{M}^{\mathcal{A}} = \sum_m m\mathcal{P}_m.$$

The projector of the measurement operator $\mathbf{M}^{\mathcal{C}}$ is

$$\mathbf{M}^{\mathcal{C}} = \mathbf{M}^{\mathcal{A}} \otimes \mathbf{I}^{\mathcal{B}} = \sum_m m\mathcal{P}_m \otimes I^{\mathcal{B}}.$$

A measurement of the state $\mid m\rangle \otimes \mid \varphi\rangle$ of $\mathcal{C}$ must produce the result $m$ with certainty. The average of the measurements performed on $\mathcal{A}$ and on $\mathcal{C}$ must be the same

$$\mathrm{tr}\left[\mathbf{M}^{\mathcal{A}}\rho_{\mathcal{A}}\right] = \mathrm{tr}\left[(\mathbf{M}^{\mathcal{A}} \otimes \mathbf{I}^{\mathcal{B}})\rho_{\mathcal{C}}\right]$$

The partial trace $\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{C}}\right]$ is the only function satisfying this equation, see [308].

**Proposition.** *If $\mid \psi_{\mathcal{C}}\rangle$ is a pure state of the bipartite system $\mathcal{C} = \mathcal{AB}$ then the eigenvalues of the reduced density operators of the two component systems are identical. Thus many properties of the two subsystems which are determined by the eigenvalues of the reduced density operators are the same for the two subsystems.*

*Proof:* The reduced density operator for the subsystem $\mathcal{A}$ can be transformed when we apply the Schmidt decomposition as

$$\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}}\left(\rho_{\mathcal{C}}\right) = \mathrm{tr}_{\mathcal{B}}\left(\mid \psi_{\mathcal{C}}\rangle\langle\psi_{\mathcal{C}} \mid\right) = \sum_{j=1}^{n} \lambda_j \mid e_j^A\rangle \sum_{j=1}^{n} \lambda_j^* \langle e_j^A \mid = \sum_{j=1}^{n} \mid \lambda_j \mid^2 \mid e_j^A\rangle\langle e_j^A \mid .$$

Recall that the eigenvalues of the density operator are the coefficients of the projectors in the sum; thus the eigenvalues of $\rho^{\mathcal{A}}$ are $\mid \lambda_i \mid^2$. Similarly,

$$\rho^{\mathcal{B}} = \mathrm{tr}_{\mathcal{A}}\left(\rho_{\mathcal{C}}\right) = \mathrm{tr}_{\mathcal{A}}\left(\mid \psi_{\mathcal{C}}\rangle\langle\psi_{\mathcal{C}} \mid\right) = \sum_{j=1}^{n} \lambda_j \mid e_j^B\rangle \sum_{j=1}^{n} \lambda_j^* \langle e_j^B \mid = \sum_{j=1}^{n} \mid \lambda_j \mid^2 \mid e_j^B\rangle\langle e_j^B \mid .$$

Communication over quantum channels involves the transport of quantum particles and could certainly benefit from the formalism described above. This formalism allows us to determine the state of $\mathcal{A}$, the original quantum particle(s) prepared in state $\mid \varphi_{\mathcal{A}}\rangle$ based on observations performed on the bipartite system $\mathcal{AB}$ in state $\mid \varphi_{\mathcal{AB}}\rangle$. Next we discuss quantum teleportation, an application of a measurement of one subsystem of a bipartite system.

**Quantum teleportation.** Quantum teleportation is the transfer of quantum state from one quantum system to another. We assume that Alice and Bob are given a pair of entangled particles called $P_1$ and $P_2$ in a maximally entangled state

$$\mid \psi_{12}\rangle = \mid \beta_{00}\rangle = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}}.$$

Eve asks Alice to deliver to Bob a message encoded as the state $\mid \psi_3\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ of particle $P_3$ Figure 53. Alice follows the following procedure: she first entangles her half of the pair with the particle from Eve; she applies a `CNOT` to this pair, using the state of $P_3$ as the control qubit and the state of $P_1$ as the target qubit. Then Alice performs a joint measurement of her two qubits, whereby she measures the state of $P_3$ and leaves $P_1$ unchanged; as a result of this measurement she obtains one of four values: $\mid 00\rangle, \mid 01\rangle, \mid 10\rangle$, and $\mid 11\rangle$ which correspond to classical information $00, 01, 10$, and $11$, respectively, with equal probability, $p = 1/4$. Alice sends the result of the measurement, $00, 01, 10$, or $11$, over a classical communication channel; upon receiving the results Bob applies one of four transformations to the state of $P_2$: $I$ if the string is 00, $X$ for 01, $Z$ for 10, or $Y$ for 11. After the transformation, the state of $P_2$ replicates to the state of $P_3$.

Quantum teleportation *does not involve an instantaneous exchange of information;* the state of Bob's qubit after Alice's measurement is not dependent on the state of $P_3$. No measurement performed by Bob after Alice's measurement and without knowing her result contains definite information about particle $P_3$. Alice *must use a classical communication channel* to transmit the result of her measurement to Bob; therefore, there is no instantaneous transfer of information. An analysis of this process follows:

Before the measurement $\mid \xi\rangle$, the joint state of particle $P_3$ and the entangled pair of particles $P_1$ and $P_2$, is

$$\mid \xi\rangle = \mid \psi_3\rangle \otimes \mid \psi_{12}\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ \alpha_0 \\ \alpha_1 \\ 0 \\ 0 \\ \alpha_1 \end{pmatrix},$$

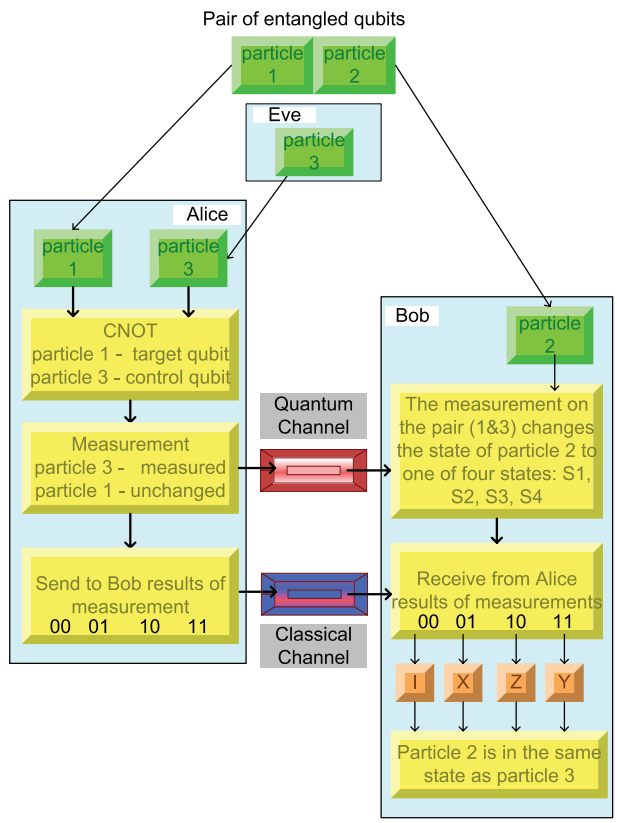



Figure 53: Quantum teleportation. Initially, Eve's particle is in state $\mid \psi_C\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$; the pair of particles shared by Alice and Bob are in a maximally entangled state $\mid \beta_{00}\rangle$. At the end of the experiment Bob's half of the entangled pair is in state $\mid \psi_C\rangle$, while the state of Eve's particle was affected by the quantum measurement and is no longer $\mid \psi_C\rangle$. The change of state of Bob's particle does not occur instantaneously, Alice must use a classical communication channel to transmit the results of her measurement to Bob.

$$\mid \xi\rangle = (1/\sqrt{2})(\alpha_0 \mid 000\rangle + \alpha_0 \mid 011\rangle + \alpha_1 \mid 100\rangle + \alpha_1 \mid 111\rangle).$$

Alice applies a `CNOT` to the pair; she uses Eve's qubit as a control and her own as a target

$$\mid \kappa\rangle = (G_{CNOT} \otimes I)(\mid \xi\rangle) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (\mid \xi\rangle),$$

$$\mid \kappa\rangle = (1/\sqrt{2})(\alpha_0 \mid 000\rangle + \alpha_0 \mid 011\rangle + \alpha_1 \mid 101\rangle + \alpha_1 \mid 110\rangle).$$

$\mid \zeta\rangle$, the state of the three particle system after Alice performs her measurement is

$$\mid \zeta\rangle = (H \otimes I \otimes I) \mid \kappa\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mid \kappa\rangle,$$

$$| \zeta\rangle = \frac{1}{2}[\alpha_0(| 000\rangle + | 011\rangle + | 100\rangle + | 111\rangle) + \alpha_1(| 001\rangle + | 010\rangle - | 101\rangle - | 110\rangle)].$$

We rewrite the expression for $| \zeta\rangle$ as

$$| \zeta\rangle = 1/2[| 00\rangle(\alpha_0 | 0\rangle + \alpha_1 | 1\rangle) \ + \ | 01\rangle(\alpha_0 | 1\rangle) + \alpha_1 | 0\rangle) +$$

$$| 10\rangle(\alpha_0 | 0\rangle - \alpha_1 | 1\rangle) \ + \ | 11\rangle(\alpha_0 | 1\rangle - \alpha_1 | 0\rangle)].$$

From this expression it follows that when Alice performs a joint measurement of her two qubits, she gets one of the four equally probable results $| 00\rangle$, or $| 01\rangle$, or $| 10\rangle$, or $| 11\rangle$. As we already know, the measurement forces the pair of qubits 1 and 3 to one of the four basis states, and transforms quantum information to classical one. Then, she sends Bob the result of her measurement, $00, 01, 10$, or $11$, over a classical communication channel. At the same time, the measurement performed by Alice forces the qubit $P_2$ in Bob's possession to change to one of the four states:

1. $| \eta_{00}\rangle = \alpha_0 | 0\rangle + \alpha_1 | 1\rangle$ when the result is 00.
2. $| \eta_{01}\rangle = \alpha_0 | 1\rangle + \alpha_1 | 0\rangle$ when the result is 01.
3. $| \eta_{10}\rangle = \alpha_0 | 0\rangle - \alpha_1 | 1\rangle$ when the result is 10.
4. $| \eta_{11}\rangle = \alpha_0 | 1\rangle - \alpha_1 | 0\rangle$ when the result is 11.

The density operator for state $| \zeta\rangle$ is

$$\begin{aligned}
\rho_{\mathcal{C}}(\zeta) = \textstyle\sum_{i=1}^{4} p_i\rho_i &= 1/4(\rho_0 + \rho_1 + \rho_2 + \rho_3) \\
&= 1/4\{| 00\rangle\langle 00 | [(\alpha_0 | 0\rangle + \alpha_1 | 1\rangle)(\alpha_0^*\langle 0 | + \alpha_1^*\langle 1 |)] \\
&\quad + | 01\rangle\langle 01 | [(\alpha_0 | 1\rangle + \alpha_1 | 0\rangle)(\alpha_0^*\langle 1 | + \alpha_1^*\langle 0 |)] \\
&\quad + | 10\rangle\langle 10 | [(\alpha_0 | 0\rangle - \alpha_1 | 1\rangle)(\alpha_0^*\langle 0 | - \alpha_1^*\langle 1 |)] \\
&\quad + | 11\rangle\langle 11 | [(\alpha_0 | 1\rangle - \alpha_1 | 0\rangle)(\alpha_0^*\langle 1 | - \alpha_1^*\langle 0 |)]\}.
\end{aligned}$$

Bob's qubit is the second of the pair and the reduced density operator of Bob's qubit is

$$\begin{aligned}
\rho^{\mathcal{B}} = \mathrm{tr}_{\mathcal{A}}\,[\rho_{\mathcal{C}}(\zeta)] &= 1/4[(\alpha_0 | 0\rangle + \alpha_1 | 1\rangle)(\alpha_0^*\langle 0 | + \alpha_1^*\langle 1 |) \\
&\quad + (\alpha_0 | 1\rangle + \alpha_1 | 0\rangle)(\alpha_0^*\langle 1 | + \alpha_1^*\langle 0 |) \\
&\quad + (\alpha_0 | 0\rangle - \alpha_1 | 1\rangle)(\alpha_0^*\langle 0 | - \alpha_1^*\langle 1 |) \\
&\quad + (\alpha_0 | 1\rangle - \alpha_1 | 0\rangle)(\alpha_0^*\langle 0 | - \alpha_1^*\langle 1 |)].
\end{aligned}$$

Then,

$$\rho^{\mathcal{B}} = \frac{1}{4}\ [| 0\rangle\langle 0 | 2(| \alpha_0 |^2 + | \alpha_1 |^2) + | 1\rangle\langle 1 | 2(| \alpha_0 |^2 + | \alpha_1 |^2)].$$

But we know that $| \alpha_0 |^2 + | \alpha_1 |^2 = 1$ thus,

$$\rho^{\mathcal{B}} = \frac{| 0\rangle\langle 0 | + | 1\rangle\langle 1 |}{2} = \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}I.$$

This confirms that the state of Bob's qubit after Alice's measurement, but before Bob has learned the measurement result is $I/2$ and it is independent on the state of particle $P_3$, as stated earlier. Note that Bob's qubit is in a mixed state; indeed:

$$\mathrm{tr}\left[\left(\frac{I}{2}\right)^2\right] = \frac{1}{2} < 1.$$

The particles $P_1$ and $P_3$ are in pure states, while $P_2$ is in a mixed state. Needless to say that the teleportation gedanken experiment described in this section does not violate the "no-cloning theorem." The state of $P_3$ has not been cloned, it has been altered in the measurement process; $P_2$ acquires the original state of $P_3$ as a result of teleportation.

## 2.14 Operator-Sum (Kraus) Representation

The state $\mid \psi\rangle$ of a quantum system with the density matrix $\rho$ can be transformed by applying the unitary operator $\mathbf{U}$ or a measurement operator $\mathbf{M}$. Call $\rho'$ the density matrix of the quantum system after the transformation and denote by $\mathcal{E}$ a generic quantum operation, be it a unitary transformation or a measurement, such that

$$\rho' = \mathcal{E}(\rho).$$

We distinguish open from closed quantum systems; an open quantum system consists of a *principal system* in contact and interacting with an *environment*. The ensemble (principal system + environment) forms a closed quantum system. We consider three cases: (i) a unitary transformation $\mathbf{U}$ is applied to a system in isolation; (ii) a measurement of the system; and (iii) a unitary transformation $\mathbf{U}$ applied to an *open* system.

When the quantum operation is a unitary transformation $\mathbf{U}$ applied to a system in isolation, Figure 54(a) then

$$\rho' = \mathcal{E}(\rho) = \mathbf{U}\rho\mathbf{U}^\dagger.$$

If the quantum operation is a measurement of the quantum system with density matrix $\rho$ and the measurement is described by the collection of measurement operators $\{\mathbf{M}_m\}$ then, the probability of the outcome labelled $m$ is $p(m) = \langle\psi \mid \mathbf{M}_m^\dagger\mathbf{M}_m \mid \psi\rangle$ and the normalized state after the measurement is carried out by the operator $\mathbf{M}_m$ is

$$\mid \varphi\rangle = \frac{\mathbf{M}_m \mid \psi\rangle}{\sqrt{\langle\psi \mid \mathbf{M}_m^\dagger\mathbf{M}_m \mid \psi\rangle}}.$$

It is easy to show that when the measurement is carried out using operator $\mathbf{M}_m$ then, immediately after the measurement:
(i) The density matrix of the quantum system is

$$\rho'_m = \mathcal{E}_m(\rho) = \mathbf{M}_m\rho\mathbf{M}_m^\dagger.$$

(ii) The probability of the outcome $m$ is

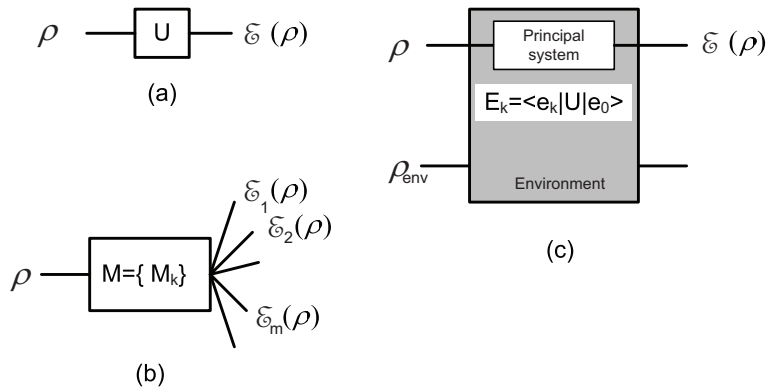



Figure 54: Quantum operations. (a) An isolated quantum system subject to unitary transformation $\mathbf{U}$; the density matrix after the operation is $\rho' = \mathcal{E}(\rho) = \mathbf{U}\rho\mathbf{U}^\dagger$. (b) A measurement described by the collection of measurement operators $\{\mathbf{M}_m\}$; if the measurement operator is $\mathbf{M}_m$ then the density matrix after the measurement is $\rho'_m = \mathcal{E}_m(\rho) = \mathbf{M}_m\rho\mathbf{M}_m^\dagger$ (c) An open quantum system with density matrix $\rho$ which interacts with an environment with density matrix $\rho_{env}$ subject to unitary transformation $\mathbf{U}$. If $\mathbf{E}_k = \langle e_k \mid \mathbf{U} \mid e_0\rangle$ is an operator on the state space of the principal system with $\mid e_0\rangle$ the initial state and $\{e_k\}$ an orthonormal basis in the Hilbert space of the environment then the density matrix after the measurement is $\mathcal{E}(\rho) = \sum_k \mathbf{E}_k\rho\mathbf{E}_k^\dagger$.

$$p(m) = \mathrm{tr}(\mathcal{E}_m(\rho)).$$

Lastly, we consider an open system consisting of a *principal system,* and the environment, Figure 54(c). The density matrices of the two systems are $\rho$ and $\rho_{env}$, respectively; we assume that the principal system and the environment are in a product state $\rho \times \rho_{env}$. Then the effect of the unitary transformation $\mathbf{U}$ on the open system can be represented as

$$\mathcal{E} = \mathrm{tr}_{env}\left[\mathbf{U}(\rho \times \rho_{env})\mathbf{U}^\dagger\right].$$

If $\mid e_0\rangle$ is the initial state and $\{e_k\}$ is an orthonormal basis in the Hilbert space of the environment then the density matrix of the environment in the pure state $e_0$ is $\rho_{env} = \mid e_0\rangle\langle e_0 \mid$. Call $\mathbf{E}_k = \langle e_k \mid \mathbf{U} \mid e_0\rangle$ an operator on the state space of the principal system. Then the effect of the unitary transformation $\mathbf{U}$ on the principal system is

$$\mathcal{E}(\rho) = \sum_k \langle e_k \mid \mathbf{U}[\rho \otimes (\mid e_0\rangle\langle e_0 \mid)]\mathbf{U}^\dagger \mid e_k\rangle.$$

This can be expressed in a form known as <u>operator-sum representation</u>

$$\mathcal{E}(\rho) = \sum_k \mathbf{E}_k\rho\mathbf{E}_k^\dagger.$$

The operator sum-representation can be seen as a transformation in which the original state $\rho$ is subject to random transformations

$$\rho_k = \frac{\mathbf{E}_k \rho \mathbf{E}_k^\dagger}{\mathrm{tr}[\mathbf{E}_k \rho \mathbf{E}_\mathbf{k}^\dagger]}$$

each occurring with probability $p_k = \mathrm{tr}[\mathbf{E}_k \rho \mathbf{E}_k^\dagger]$, or

$$\mathcal{E}(\rho) = \sum_k p_k \rho_k.$$

The so called trace-preserving transformations $\mathcal{E}$ have the property that

$$\mathrm{tr}\,[\mathcal{E}(\rho)] = 1.$$

This implies that for all density operators $\rho$ we have

$$\mathrm{tr}\,[\mathcal{E}(\rho)] = \mathrm{tr}[\sum_k \mathbf{E}_k \rho \mathbf{E}_k^\dagger] = \mathrm{tr}[\sum_k \mathbf{E}_k^\dagger \mathbf{E}_k \rho].$$

It follows immediately that to be trace-preserving the operators $\mathcal{E}$ must satisfy the condition

$$\sum_k \mathbf{E}_k^\dagger \mathbf{E}_k = \mathbf{I}.$$

A non-trace-preserving quantum operation is characterized by the inequality

$$\sum_i \mathbf{E}_i^\dagger \mathbf{E}_i \leq \mathbf{I}.$$

It corresponds to a transformation which involves measurement of the quantum system.

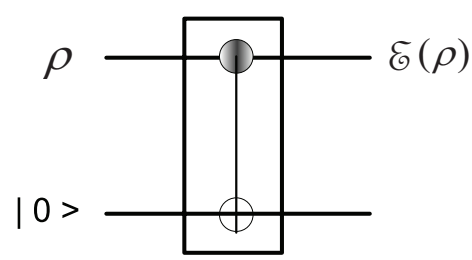



Figure 55: An open quantum system with density matrix $\rho$ which interacts with an environment with density matrix $\rho_{env}$ subject to unitary transformation `CNOT`. The principal system is the control qubit and the environment is the target qubit, initially in state $\mid 0\rangle$

**Example.** *Consider the circuit in Figure 55.* The principal system is the control qubit and the environment is initially in state $\mid 0\rangle$; the basis states of the environment $\mid e_k\rangle$ are $\mid 0_E\rangle$ and $\mid 1_E\rangle$. The unitary transformation $\mathbf{U}$ can be written as

$$\mathbf{U} = \mid 0_P 0_E\rangle\langle 0_P 0_E \mid + \mid 0_P 1_E\rangle\langle 0_P 1_E \mid + \mid 1_P 1_E\rangle\langle 1_P 0_E \mid + \mid 1_P 0_E\rangle\langle 1_P 1_E \mid,$$

where the subscripts $E$ and $P$ refer to environment and principal system, respectively.

Here, the operation elements are

$$\mathbf{E}_0 = \langle 0_E \mid \mathbf{U} \mid 0_E \rangle = \mid 0_P \rangle \langle 0_P \mid \quad \text{and} \quad \mathbf{E}_1 = \langle 1_E \mid \mathbf{U} \mid 1_E \rangle = \mid 1_P \rangle \langle 1_P \mid .$$

Thus,

$$\mathcal{E}(\rho) = \mathbf{E}_0 \rho \mathbf{E}_0^\dagger + \mathbf{E}_1 \rho \mathbf{E}_1^\dagger = \mathbf{P}_0 \rho \mathbf{P}_0 + \mathbf{P}_1 \rho \mathbf{P}_1$$

where $\mathbf{P}_0 = \mid 0 \rangle \langle 0 \mid$ and $\mathbf{P}_1 = \mid 1 \rangle \langle 1 \mid$ are the projectors.

The operator sum representation *is not unique.* We show that given a system with density matrix $\rho$, two operations $\mathcal{E}$ and $\mathcal{F}$, with distinct operation elements $\mathbf{E}_1, \mathbf{E}_2$ and $\mathbf{F}_1, \mathbf{F}_2$, lead to the same state of the system, $\mathcal{E}(\rho) = \mathcal{F}(\rho)$. Consider the following operations and elements

| Operation | Operation Element 1 | Operation Element 2 |
|---|---|---|
| $\mathcal{E}$ | $\mathbf{E}_1 = \frac{1}{\sqrt{2}}\mathbf{I} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ | $\mathbf{E}_2 = \frac{1}{\sqrt{2}}\mathbf{X} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ |
| $\mathcal{F}$ | $\mathbf{F}_1 = \frac{1}{\sqrt{2}}(\mathbf{E}_1 + \mathbf{E}_2) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ | $\mathbf{F}_2 = \frac{1}{\sqrt{2}}(\mathbf{E}_1 - \mathbf{E}_2) = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ |

We start from the definition

$$\mathcal{E}(\rho) = \mathbf{E}_1 \rho \mathbf{E}_1^\dagger + \mathbf{E}_2 \rho \mathbf{E}_2^\dagger.$$

Then,

$$\mathcal{F}(\rho) = \mathbf{F}_1 \rho \mathbf{F}_1^\dagger + \mathbf{F}_2 \rho \mathbf{F}_2^\dagger = \frac{1}{2}[(\mathbf{E}_1 + \mathbf{E}_2)\rho(\mathbf{E}_1^\dagger + \mathbf{E}_2^\dagger)] + [(\mathbf{E}_1 - \mathbf{E}_2)\rho(\mathbf{E}_1^\dagger - \mathbf{E}_2^\dagger)]$$

or, after expansion we get

$$\mathcal{F}(\rho) = \mathbf{E}_1 \rho \mathbf{E}_1^\dagger + \mathbf{E}_2 \rho \mathbf{E}_2^\dagger = \mathcal{E}(\rho).$$

The next question is how can the unitary transformation $\mathbf{U}$ be applied to an open system when the environment has a very large number of degrees of freedom. The answer, discussed in Section 3.18, is that *when the principal system is specified in a Hilbert space of dimension $d$, then it is sufficient to model the environment in a Hilbert space of dimension $d^2$.*

We now switch topics and discuss entanglement and its applications for Quantum Information Processing.

## 2.15 Entanglement; Monogamy of Entanglement

Entanglement plays a critical role in quantum computing and quantum information theory as we shall see when we discuss the quantum communication channels and quantum error correction; for example when two parties share an entangled quantum state they are able to communicate securely. In this section we present some basic facts about entanglement and introduce the *monogamy of entanglement* a property of quantum systems with important

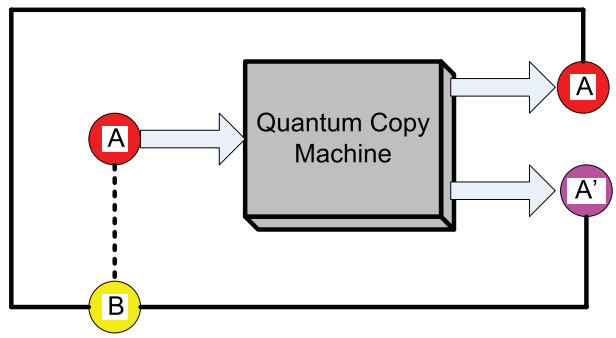



Figure 56: In violation of the monogamy of entanglement a quantum copy machine would allow a quantum state $\mathcal{B}$ maximally entangled with another quantum state $\mathcal{A}$ to end up entangled with the two copies of state $\mathcal{A}$.

consequences for many applications including quantum cryptography where it quantifies how much information an eavesdropper could potentially obtain.

The monogamy of entanglement is the deeper root of our inability to clone quantum states. Figure 56 illustrates why the monogamy of entanglement prevents quantum states to be cloned. Consider two maximally entangled quantum states, $\mathcal{A}$ and $\mathcal{B}$. Assume that we have a quantum copy machine able to clone quantum states. If the input to this quantum copy machine is system $\mathcal{A}$, then the output will be the original, system $\mathcal{A}$, and a perfect replica of it, $\mathcal{A}'$. Thus, the quantum system $\mathcal{B}$ would end up being entangled with both systems $\mathcal{A}$ and $\mathcal{A}'$, in violation of the monogamy of entanglement. The dotted line represents the original entanglement of $\mathcal{B}$ with $\mathcal{A}$, and the solid lines the entanglement of $\mathcal{B}$ with $\mathcal{A}$ and its clone $\mathcal{A}'$.

The joint state of a maximally entangled pair system is a pure state, while individual particles are in mixed state; thus, the state of individual particles cannot be known with certainty and individual particles cannot be cloned. Recall from Section 1.12 that the Bell states are *entangled* states of a Bell pair:

$$|\,\beta_{00}\rangle = \frac{|\,00\rangle + |\,11\rangle}{\sqrt{2}}, \quad |\,\beta_{01}\rangle = \frac{|\,01\rangle + |\,10\rangle}{\sqrt{2}}, \quad |\,\beta_{10}\rangle = \frac{|\,00\rangle - |\,11\rangle}{\sqrt{2}}, \quad |\,\beta_{11}\rangle = \frac{|\,01\rangle - |\,10\rangle}{\sqrt{2}}.$$

The Bell states form an orthonormal basis and can be distinguished from one another. All four states are called *maximally entangled* states. The last one is called *anti-correlated* state.

**Proposition.** *Bell states are pure states; the joint state of an EPR pair is known exactly (it is a pure state), while the state of either qubit of the pair is not (it is a mixed state).*

*Proof:* Let us pick one of the Bell states, say $|\,\beta_{10}\rangle$ and compute its density operator, $\rho_{(\beta_{10})}$. Then we compute the density operator of one of the qubits of the EPR pair, say the second one, $\rho_{(\beta_{10},second)}$. We expect that the traces of the two density operators satisfy the known relations for pure and respectively mixed states

$$\mathrm{tr}\left[\rho^2_{(\beta_{10})}\right] = 1 \qquad\qquad \mathrm{tr}\left[\rho^2_{(\beta_{10},second)}\right] < 1.$$

First, we compute the density operator for the pair:

$$\rho_{(\beta_{10})} = \mid \beta_{10}\rangle\langle\beta_{10} \mid = \frac{(\mid 00\rangle - \mid 11\rangle)}{\sqrt{2}} \frac{(\langle 00 \mid - \langle 11 \mid)}{\sqrt{2}}$$

$$\rho_{(\beta_{10})} = \frac{1}{2}(\mid 00\rangle\langle 00 \mid - \mid 00\rangle\langle 11 \mid - \mid 11\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid)$$

$$\rho_{(\beta_{10})} = \frac{1}{2}\left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right]$$

$$\rho_{(\beta_{10})} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Then,

$$\rho^2_{(\beta_{10})} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 2 \end{pmatrix}$$

$$\rho^2_{(\beta_{10})} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \qquad \Longrightarrow \qquad \mathrm{tr}\left[\rho^2_{(\beta_{10})}\right] = \frac{1}{2}(1+0+0+1) = 1$$

Now we compute the reduced density operator and the partial trace for the second qubit of the EPR pair by tracing the first qubit

$$\rho_{(\beta_{10},second)} = \mathrm{tr}_{first}\left[\rho_{\beta_{10}}\right] = \frac{\mathrm{tr}_{first}\left[\mid 00\rangle\langle 00 \mid - \mid 00\rangle\langle 11 \mid - \mid 11\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid\right]}{2}$$

$$\rho_{(\beta_{10},second)} = \frac{\mid 0\rangle\langle 0 \mid \langle 0 \mid 0\rangle - \mid 0\rangle\langle 1 \mid \langle 0 \mid 1\rangle - \mid 1\rangle\langle 0 \mid \langle 1 \mid 0\rangle + \mid 1\rangle\langle 1 \mid \langle 1 \mid 1\rangle}{2}$$

But,

$$\langle 0 \mid 0\rangle = 1, \quad \langle 0 \mid 1\rangle = 0, \quad \langle 1 \mid 0\rangle = 0, \quad \langle 1 \mid 1\rangle = 1$$

and

$$\mid 0\rangle\langle 0 \mid = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mid 1\rangle\langle 1 \mid = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

It follows that

$$\rho_{(\beta_{10},second)} = \frac{\mid 0\rangle\langle 0 \mid + \mid 1\rangle\langle 1 \mid}{2} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{I}{2}$$

$$\rho^2_{(\beta_{10},second)} = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \longrightarrow \quad \mathrm{tr}\left[\rho^2_{(\beta_{10},second)}\right] = \frac{1}{4}(1+1) = \frac{1}{2}.$$

To complete the proof of this proposition we have to repeat the calculation for the first qubit of the $\beta_{10}$ pair following the pattern presented above. Then we have to redo the calculations for the other three Bell states $\beta_{00}, \beta_{01}$ and $\beta_{11}$.

One of the most intriguing properties of quantum information is the shareability of quantum correlations. While classical correlations can be shared among many parties, quantum correlations cannot be shared freely, quantum correlations of pure states are monogamous. As we know from statistics, two random variables $X$ and $Y$ correlated with each other may also be individually correlated with other random variables. In contrast, if two quantum systems $\mathcal{A}$ and $\mathcal{B}$ are in a maximally entangled pure state, then neither of them can be correlated with any other system in the universe. Otherwise, there is tradeoff between the amount of entanglement of two qubits and the quantum correlation each of the two qubits could share with a third one. It is believed that if the two qubits are as much entangled with each other as it is possible, they cannot be entangled or even classically correlated with another qubit.

Consider three qubits $a, b, c$ with the first two in a maximally entangled pure state

$$\mid \psi_{ab}\rangle = \frac{1}{\sqrt{2}} \mid 0_a \otimes 0_b\rangle + \frac{1}{\sqrt{2}} \mid 1_a \otimes 1_b\rangle$$

and the third is in state $\mid \psi_c\rangle$, [419]. There is no quantum state shared by the three qubits, such that when we remove the third qubit, $c$, we get the joint state of qubits $a$ and $b$, $\mid \psi_{ab}\rangle$, and, at the same time, the three-qubit state does not change when we interchange qubits $a$ and $b$. Thus, the joint state of the three qubits is

$$\mid \psi_{abc}\rangle = \mid \psi_c\rangle \otimes [\frac{1}{\sqrt{2}} \mid 0_a \otimes 0_b\rangle + \frac{1}{\sqrt{2}} \mid 1_a \otimes 1_b\rangle].$$

There is no symmetry, between qubits $b$ and $c$, $b$ is maximally entangled with $a$ while $c$ is not entangled with $a$.

The next question is if all entangled states are monogamous; for example, we may inquire if entangled mixed states are monogamous. Schumacher introduced the term *shareable quantum states* and Bennett *et al.* [38] gave an example of a mixed entangled state that is shareable rather than monogamous. Consider a noisy quantum channel shared by Alice, Bob and Eve; with probability 1/2 a qubit sent by Alice is transmitted unchanged to Bob and with probability 1/2 the qubit is intercepted and Bob gets a random qubit. When Alice sends to Bob half of a maximally entangled pair, the state shared by Alice and Bob is still entangled while the state shared by the three is always symmetric with respect to Eve and Bob, thus Eve is entangled with Alice as well.

## 2.16 Einstein-Podolski-Rosen (EPR) Thought Experiment

"Liebe Gott würfelt nicht"[19], Einstein's famous pronouncement, reflects Einstein's skepticism shared by some physicists and philosophers regarding the nondeterminism of quantum theory. In 1935, in the now famous paper "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" [134] Einstein, Podolsky, and Rosen proposed a thought experiment challenging the very foundations of quantum theory. They concluded that we have a choice, either to accept the principle of locality, or to accept the postulates of Quantum Mechanics. In his autobiography Einstein wrote [135]: "....the paradox forces us to relinquish one of the following two assertions: (i) the description by means of the wave function $\psi$ is complete, (ii) the real states of spatially separated objects are independent of each other." Einstein concluded that the quantum mechanical description by means of the wave function is not complete, he had no doubts that the second assertion is true. Non-locality would violate one of the basic postulates of the Special Theory of Relativity[20].

Before presenting the EPR concepts we describe a relatively simple experiment known as the ARC which seems to illustrate that the randomness of the quantum theory is illusory.

**The atomic radiation cascade (ARC) experiment.** The ARC source in Figure 57 produces two polarization-entangled photons which move in opposite directions. An atom, e.g., $Ca^{40}$, initially laser pumped to an excited $S$ state, the upper level of the cascade state, emits two photons that appear to have the same polarization. Let $\mid x_1\rangle, \mid x_2\rangle, \mid y_1\rangle$ and $\mid y_2\rangle$ denote the states of the two photons with linear polarization along the $x$- and respectively $y$-axis orthogonal to the $z$-axis. It can be shown [326] that the state of the entangled system of the two photons can only be

$$\mid \psi_+\rangle = 1/\sqrt{2}(\mid x_1\rangle \mid x_2\rangle + \mid y_1\rangle \mid y_2\rangle).$$

One of the photons is detected by a polarizing cube[21] located in the immediate vicinity of the source and the other is reflected by a distant mirror and then detected by a second calcite crystal. Each detector is connected to a device which prints on a paper tape the symbols $+$ and $-$ depending on the polarization of the photon it has detected, Figure 57. There is a delay between the measurement performed by the upper detector located in the immediate vicinity of the source and the measurement performed by the second detector on the photon whose path is reflected by the distant mirror.

As expected, each paper tape contains a complectly random sequence of $+$ and $-$ symbols. If we compare the two paper tapes we see that they are perfectly correlated, an observer able to see the first tape can predict with certainty the next symbol to be printed on the second tape. Amazingly, the second tape which looks perfectly random when examined in isolation is

[19] Dear God does not play dice.

[20] The special theory of relativity of Einstein describes the motion of particles moving at close to the speed of light. The two basic postulates of special relativity are: (1) The speed of light is the same for all observers, no matter what their relative speeds. (2) The laws of physics are the same in any inertial (that is, non-accelerated) frame of reference. This means that the laws of physics observed by a hypothetical observer travelling with a relativistic particle must be the same as those observed by an observer who is stationary in the laboratory.

[21] A polarizing cube is made of two prisms with a dielectric film on the sides which are stuck together. The polarizing cube transmits photons with horizontal polarization and reflects photons with vertical polarization.

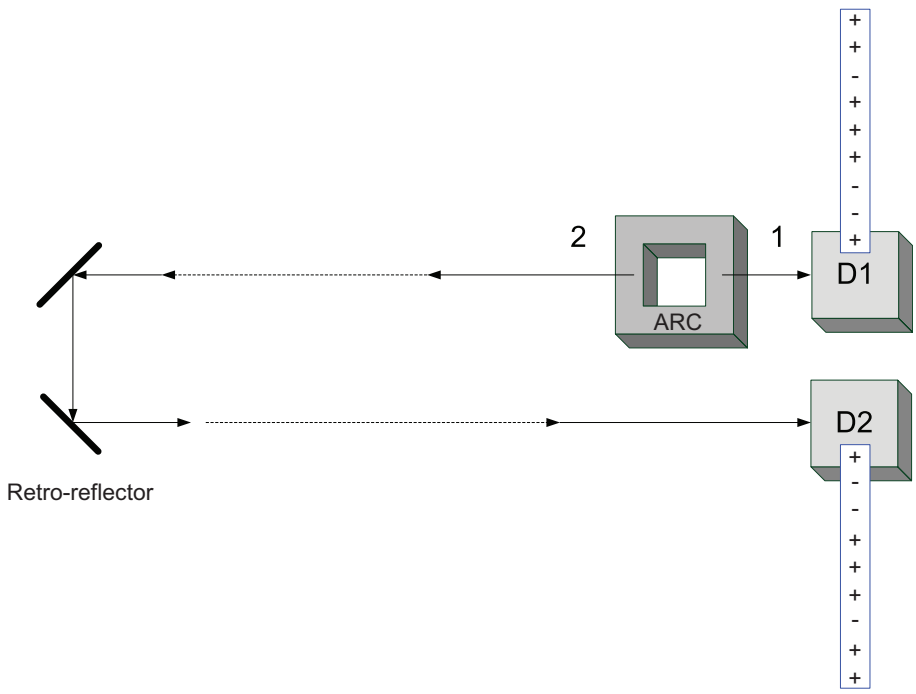



Figure 57: An ARC source produces two polarization entangled photons. Detectors $D1$, located in the immediate vicinity of the source, and $D2$ detecting a photon reflected by a distant mirror, are connected to output devices which print one of tho symbols $+$ or $-$ depending on the polarization of the photons. The two paper tapes contain random sequences of $+$s and $-$s but they are perfectly correlated. The second tape which looks perfectly random when examined in isolation, is in fact deterministic, we could have predicted with absolute certainty the sequence of symbols on that tape.

in fact deterministic, we could have predicted with absolute certainty the sequence of symbols on that tape.

**The EPR experiment.** EPR is a thought (gedanken) experiment; the EPR argument is based on the concept of "element of physical reality" defined as follows: *If without in any way disturbing a system we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.*

We consider a bipartite system consisting of two particles, $P_1$ and $P_2$, prepared in such a state that their total momentum is close to zero and their relative distance is close to $L$, where $L$ is much larger than the distance that allows the two particles to interact with each other, as in Figure 58(a). If we denote by $\gamma$ a normalizable function with a very high and very narrow peak and by $x_1, x_2, p_1, p_2$ the positions and the momenta of the two particles, then the state of the system is described by an entangled wave function

$$\psi = \gamma(x_1 - x_2 - L)\gamma(p_1 + p_2).$$

We do not know anything about the position of the particles or about their individual momenta, we only know that they are at distance $L$ from one another and that the total momentum is equal to zero.

If we measure the position of the first particle, $x_1$, the wave function allows us to predict with certainty the position of the second particle, $x_2$. It is easy to argue that both $x_1$ and $x_2$ are elements of the physical reality. Indeed, the measurements of the two systems do not affect each other because the distance $L$ was chosen to be large enough to prevent such interactions. Thus, no change may take place in the second system as a result of the measurement performed in the first system.

We could have measured the momentum of the first particle, $p_1$, and then we would have been able to predict the momentum of the second one, $p_2$. Similar arguments indicate that both $p_1$ and $p_2$ are elements of the physical reality. Yet, Heisenberg inequality precludes the simultaneous assignment of precise values to both the position $x_2$ and the momentum, $p_2$ of the second particle because the two operators corresponding to the two measurements of observables do not commute.

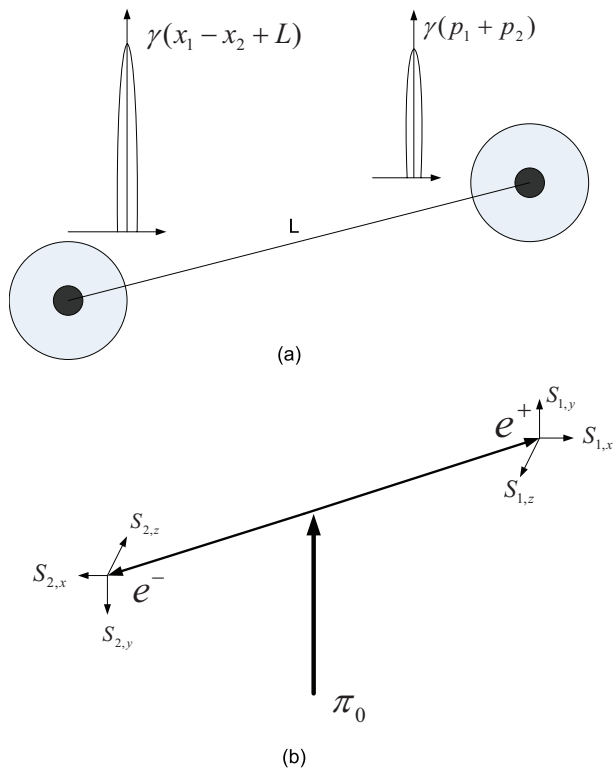



Figure 58: (a) The original EPR experiment. (b) Bohm version of the EPR experiment

**Bohm version of EPR.** David Bohm gives a simpler example of the same paradox [64], Figure 58(b). Recall that the "spin" is the quantum number characterizing the intrinsic angular momentum of a boson or fermion and that Quantum Mechanics asserts that at most one component of the spin of a particle may be definite.

Consider a spinless system such as the neutral pion, $\pi^0$, which can decay in two 1/2-spin fermions, a positron, $e^+$, and an electron, $e^-$, which fly away from each other, and a photon $\pi^0 \rightarrow e^+ + e^- + \gamma$. When they are very far apart we measure the component of the spin of the electron along the $x$-direction, $S_{1,x}$, and find it equal to $\hbar/2$. Then we know that the component of the spin of the positron along the $x$-axis must have opposite orientation,

$S_{2,x} = -\hbar/2$; moreover it must have been in that state since it was emitted because it did not suffer any interactions. Similarly, we could have measured the component of the spin of the electron along the $y$-axis, $S_{1,y}$, and this measurement would have allowed us to predict with certainty the moment of the positron along the $y$-axis, $S_{2,y}$. This same argument applies for the measurement of the momentum along the $z$-axis, $S_{1,z}$ and $S_{2,z}$.

We do not claim that we actually measure the three components of the spin of the electron; all we say is *we could have* measured all of them and in that case all three components of the spin of the positron *would have* been predictable with certainty. In this case the three components of the spin of the positron, $S_{2,x}, S_{2,y}$, and $S_{2,z}$ would have been "elements of reality," and that would have been in stark contradiction with the laws of Quantum Mechanics.

**Peres version of EPR.** Asher Peres gives an argument in the spirit of EPR and reformulates the EPR paradox as an algebraic contradiction [326]. Given two commuting operators, $\mathbf{A}$ and $\mathbf{B}$ for observables $A$ and $B$ Quantum Mechanics allows us to discover "primitive elements of reality." We can measure the observables $A$ and $B$ simultaneously and get as results of these measurements the eigenvalues of these operators, $\alpha$ and $\beta$, respectively. Let us now assume that Quantum Mechanics allows us to define recursively "secondary elements of reality." This means that any algebraic function $f(\mathbf{A}, \mathbf{B})$ corresponds to an "element of reality" and the result of such a measurement is $f(\alpha, \beta)$.

Let us apply this argument to the spinless system in Bohm reformulation of the EPR experiment. Call $m_{1,x}$ and $m_{2,x}$, $m_{1,y}$ and $m_{2,y}$ the values of the projections of the spin on the $x$- and $y$-axis for the electron and positron, $S_{1,x}$, $S_{2,x}$ and $S_{1,y}$, $S_{2,y}$, respectively. We know that $m_{1,x} = -m_{2,x}$ and $m_{1,y} = -m_{2,y}$. The spin observables $S_{1,x}$ and $S_{2,y}$ commute, as well as $S_{1,y}$ and $S_{2,x}$. Since they correspond to elements of reality, their products $S_{1,x}S_{2,y}$ and $S_{1,y}S_{2,x}$ also correspond to elements of reality; their numerical values are given by the products of individual numerical values $m_{1,x}m_{2,y}$ and $m_{1,y}m_{2,x}$. These numerical products are equal in our case, thus:

$$S_{1,x}S_{2,y} = S_{1,y}S_{2,x} = m_{1,x}m_{2,y}.$$

This contradicts Quantum Mechanics which asserts that these products should have opposite values because the singlet state[22] $\pi^0$ satisfies the equation:

$$(S_{1,x}S_{2,y} + S_{1,y}S_{2,x})\psi = 0.$$

In summary, the EPR is a gedanken, or thought experiment challenging the dogma that the description of reality given by the wave function in Quantum Mechanics is complete. Einstein believed that EPR forces a choice "either (1) the quantum-mechanical description of reality given by the wave function is not complete; or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality."

Some of those questioning the non-determinism of Quantum Mechanics attempted to deduce the properties of quantum systems from those of a yet unknown, but deterministic sub-quantum world and found refuge in the theory of "hidden variables." In the next section we show that the hidden variable theory is not sound.

[22] Singlet electron state: anti-symmetric state of a pair of electrons with *anti-parallel* spins $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. The electrons have different spin quantum numbers $+1/2$ and $-1/2$, and the total spin of the state is zero.

## 2.17 Hidden Variables

The EPR experiment led some physicists to the belief that the nondeterminism of Quantum Mechanics could be explained by the existence of "hidden variables." The proponents of this theory believed that if we knew the exact values of hidden variables then we would have a fully deterministic view of the world.

A suggestive analogy was proposed by one of our students: "Imagine that we are behind a wall that obscures the view of the other side where several machines throw tennis balls over the wall. The trajectory of each ball depends on the setting of each machine; assuming that the tennis balls are identical and their weight and diameter are known and that there is no wind, the trajectory of each one is perfectly deterministic. Yet, to us the trajectory of a ball appears to be random. The nondeterminism is due to the lack of knowledge of the initial conditions for each trajectory, the hidden variables of this game."

We now examine more closely the hidden variable theory and show that it cannot reproduce the statistical properties of Quantum Mechanics evidenced by countless experiments.

**The argument of von Neumann.** In 1932, von Neumann attempted to prove that quantum theory is incompatible with the existence of dispersion-free ensembles (ensembles which have dispersion equal to zero for all their observables) [442]. He considered $N$ independent versions of a quantum system, each in a quantum state $\mid \psi_i\rangle$, with $1 \leq i \leq N$. A measurement of the same observable $A$ on all $N$ versions of the system allows us to compute the ensemble average of the operator $\mathbf{A}$. Call this ensemble average $\langle \mathbf{A} \rangle$:

$$\langle \mathbf{A} \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \psi_i \mid \mathbf{A} \mid \psi_i \rangle.$$

The proof requires an ensemble of physical systems such that for any observable $A$, the associated operator $\mathbf{A}$ is self-adjoint and satisfies the zero-dispersion condition:

$$\langle \mathbf{A}^2 \rangle = \langle \mathbf{A} \rangle^2.$$

Yet it is impossible to prepare an ensemble of physical systems such that any observable $A$ has an associated measurement operator which satisfies the condition $\langle \mathbf{A}^2 \rangle = \langle \mathbf{A} \rangle^2$. The proof requires also that for any two observables, $A$ and $B$, the corresponding operators $\mathbf{A}$ and $\mathbf{B}$ satisfy the equation:

$$\langle \mathbf{A} + \mathbf{B} \rangle = \langle \mathbf{A} \rangle + \langle \mathbf{B} \rangle.$$

Consider the case when the operators $A$ and $B$ do not commute and cannot be measured simultaneously. If $\mathbf{A}$ and $\mathbf{B}$ do not commute we need three experimental setups to measure independently the observables $A$, $B$, and $A + B$. The three setups could be very different and the measurements would not justify the equality $\langle \mathbf{A} + \mathbf{B} \rangle = \langle \mathbf{A} \rangle + \langle \mathbf{B} \rangle$.

**John Bell's proof of incompatibility between Quantum Mechanics and the hidden variable theory.** John Bell has devoted his scientific life to the philosophy of the measurements [26] and his results are critical for understanding the subtle differences between classical and quantum information. In 1964, 32 years after von Neumann published his result,

John Bell provided a rigorous proof that a non-deterministic model based on hidden variables is incompatible with Quantum Mechanics [25].

Consider a pair of spin one-half particles in the singlet state $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$ moving freely in opposite directions. Let $\bar{\alpha}$ and $\bar{\beta}$ be two unit vectors with arbitrary orientations. We can measure the projections of the spins of the two particles, $\bar{\sigma_1}$ and $\bar{\sigma_2}$ along $\bar{\alpha}$ using two Stern-Gerlach setups. The two measurements produce opposite results: if $\bar{\sigma_1} \cdot \bar{\alpha} = +1$ then $\bar{\sigma_2} \cdot \bar{\alpha} = -1$ and viceversa.

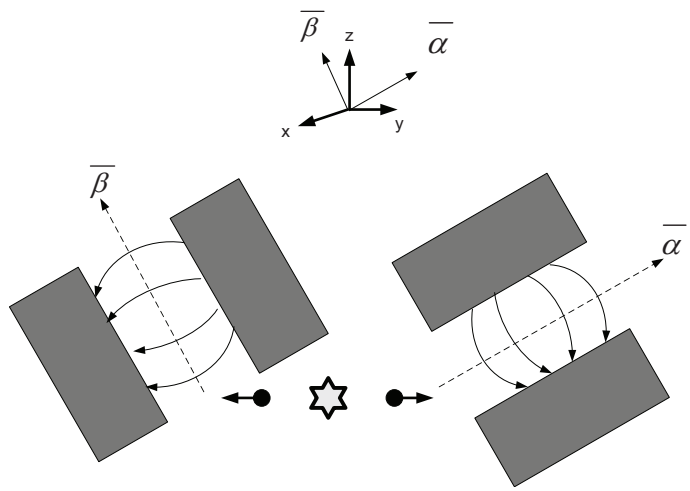



Figure 59: A pair of spin one-half particles in the singlet state $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$ move freely in opposite directions. We measure the spins along two arbitrary directions, $\bar{\alpha}$ and $\bar{\beta}$ using two Stern-Gerlach experimental setups.

The question posed is: "Is there a more complete specification of the state of the two particles from the one given by the wave function?" To answer the question we assume that we measure the spins of the particles along two independently chosen directions given by the unitary vectors $\bar{\alpha}$ and $\bar{\beta}$ using two Stern-Gerlach setups as in Figure 59. We also assume that there is a hidden continuous random variable $\lambda$ that would enable a more complete specification of the state; the probability density function of this random variable is $\rho(\lambda)$ and satisfies the relationship $\int \rho(\lambda) d\lambda = 1$.

Let $a$ and $b$ be the outcomes of the measurements of the spins of the two particles, $\bar{\sigma_1} \cdot \bar{\alpha}$ and $\bar{\sigma_2} \cdot \bar{\beta}$ respectively. We assume that the outcome $b$ of the spin measurement for the second particle does not depend on the direction $\bar{\alpha}$ of the Stern-Gerlach apparatus setting for the measurement of the first particle, and that the outcome $a$ of the spin measurement for the first particle does not depend on the direction $\bar{\beta}$ of the Stern-Gerlach apparatus setting for the measurement of the spin of the second particle. Thus $a$ and $b$ are random variables and functions of

(i) the direction we project the spin onto, and

(ii) the value of the hidden random variable $\lambda$.

$$a = a(\bar{\alpha}, \lambda) = \pm 1 \qquad \text{and} \qquad b = b(\bar{\beta}, \lambda) = \pm 1.$$

The expected values of $a$ and $b$ when $\lambda$ takes on different values are functions of the two orientations $\bar{\alpha}$ and $\bar{\beta}$, respectively:

$$E(a) = f_1(\bar{\alpha}) = \int a(\bar{\alpha}, \lambda)\rho(\lambda)d\lambda \qquad \text{and} \qquad E(b) = f_2(\bar{\beta}) = \int b(\bar{\beta}, \lambda)\rho(\lambda)d\lambda.$$

The expected value of the product of the two random variables $a$ and $b$ is a function of both $\bar{\alpha}$ and $\bar{\beta}$:

$$E(ab) = f(\bar{\alpha}, \bar{\beta}) = \int a(\bar{\alpha}, \lambda)b(\bar{\beta}, \lambda)\rho(\lambda)d\lambda.$$

On the other hand, the expected value of the product of the two measurements predicted by Quantum Mechanics is a function of the angle between the two directions $\bar{\alpha}$ and, respectively, $\bar{\beta}$ on which we project the spins of the two particles:

$$\langle(\bar{\sigma_1} \cdot \bar{\alpha})(\bar{\sigma_2} \cdot \bar{\beta})\rangle = (\bar{\sigma_1} \cdot \bar{\sigma_2})(\bar{\alpha} \cdot \bar{\beta}) = -(\bar{\alpha} \cdot \bar{\beta}).$$

In this expression $\bar{\sigma_1} \cdot \bar{\sigma_2} = -1$ because the spins of the two particles are opposite.

Our original question can now be formulated as: "Are the two averages, $f(\bar{\alpha}, \bar{\beta})$, calculated based on a hidden variable model, and $-(\bar{\alpha} \cdot \bar{\beta})$, predicted by Quantum Mechanics, equal?" John Bell proves that they are not [25, 26] and we summarize Bell's arguments in the following propositions.

**Proposition.** *The average of the two observables calculated based on a hidden variable model satisfies the following five conditions:*

**1.** Has a lower bound $f(\bar{\alpha}, \bar{\beta}) \geq -1$; this follows from the definition,

$$f(\bar{\alpha}, \bar{\beta}) = \int a(\bar{\alpha}, \lambda)b(\bar{\beta}, \lambda)\rho(\lambda)d\lambda \text{ with } a(\bar{\alpha}, \lambda) = \pm 1,\ b(\bar{\beta}, \lambda) = \pm 1, \text{ and } \int \rho(\lambda)d\lambda = 1.$$

**2.** $f(\bar{\alpha}, \bar{\alpha}) = -1$ only if $a(\bar{\alpha}, \lambda) = -b(\bar{\alpha}, \lambda)$. When the two directions coincide, $\bar{\alpha} = \bar{\beta}$ then,

$$f(\bar{\alpha}, \bar{\alpha}) = \int a(\bar{\alpha}, \lambda)b(\bar{\alpha}, \lambda)\rho(\lambda)d\lambda.$$

If $a(\bar{\alpha}, \lambda) = -b(\bar{\alpha}, \lambda)$ then,

$$f(\bar{\alpha}, \bar{\alpha}) = -\int a(\bar{\alpha}, \lambda)a(\bar{\alpha}, \lambda)\rho(\lambda)d\lambda.$$

But,

$$a(\bar{\alpha}, \lambda) = \frac{1}{a(\bar{\alpha}, \lambda)} = \pm 1.$$

Then $a(\bar{\alpha}, \lambda)a(\bar{\alpha}, \lambda) = 1$; we also know that $\int \rho(\lambda)d\lambda = 1$ thus, $f(\bar{\alpha}, \bar{\alpha}) = -1$.

**3.** An alternative expression for $f(\bar{\alpha}, \bar{\beta})$ is

$$f(\bar{\alpha},\bar{\beta}) = -\int a(\bar{\alpha},\lambda)a(\bar{\beta},\lambda)\rho(\lambda)d\lambda.$$

If we substitute $b(\bar{\beta},\lambda)$ with $-a(\bar{\beta},\lambda)$ in the definition of $f(\bar{\alpha},\bar{\beta})$ then we obtain the desired result, the outcome $a$ along direction $\bar{\beta}$.

**4.** Consider three arbitrary directions in space and let $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$ be three unitary vectors along these three directions. Then,

$$1 + f(\bar{\beta},\bar{\gamma}) \geq \mid f(\bar{\alpha},\bar{\beta}) - f(\bar{\alpha},\bar{\gamma}) \mid .$$

To show this we transform the following expression:

$$\begin{aligned} f(\bar{\alpha},\bar{\beta}) - f(\bar{\alpha},\bar{\gamma}) &= -\textstyle\int a(\bar{\alpha},\lambda)a(\bar{\beta},\lambda)\rho(\lambda)d\lambda + \int a(\bar{\alpha},\lambda)a(\bar{\gamma},\lambda)\rho(\lambda)d\lambda \\ &= \textstyle\int a(\bar{\alpha},\lambda)a(\bar{\beta},\lambda)[a(\bar{\gamma},\lambda)/a(\bar{\beta},\lambda) - 1]\rho(\lambda)d\lambda \\ &= \textstyle\int a(\bar{\alpha},\lambda)a(\bar{\beta},\lambda)[a(\bar{\gamma},\lambda)a(\bar{\beta},\lambda) - 1]\rho(\lambda)d\lambda. \end{aligned}$$

The last equality is due to the fact that

$$a(\bar{\beta},\lambda) = \pm 1, \quad \text{therefore} \quad a(\bar{\beta},\lambda) = \frac{1}{a(\bar{\beta},\lambda)}.$$

Considering also that $a(\bar{\alpha},\lambda) = \pm 1$, we can write

$$\begin{aligned} \mid f(\bar{\alpha},\bar{\beta}) - f(\bar{\alpha},\bar{\gamma}) \mid &\leq \textstyle\int -\left[a(\bar{\beta},\lambda)a(\bar{\gamma},\lambda) - 1\right]\rho(\lambda)d\lambda \\ &\leq \textstyle\int \left[1 - a(\bar{\beta},\lambda)a(\bar{\gamma},\lambda)\right]\rho(\lambda)d\lambda \\ &\leq 1 - \textstyle\int (a(\bar{\beta},\lambda)a(\bar{\gamma},\lambda)\rho(\lambda)d\lambda. \end{aligned}$$

But $f(\bar{\beta},\bar{\gamma}) = -\int (a(\bar{\beta},\lambda)a(\bar{\gamma},\lambda)\rho(\lambda)d\lambda$ which proves that

$$1 + f(\bar{\beta},\bar{\gamma}) \geq \mid f(\bar{\alpha},\bar{\beta}) - f(\bar{\alpha},\bar{\gamma}) \mid .$$

**5.** $f(\bar{\beta},\bar{\gamma})$ cannot be stationary when it reaches the minimum value of $-1$ for $\bar{\beta} = \bar{\gamma}$.

If $f(\bar{\beta},\bar{\gamma}) \neq$ constant and if $\mid \bar{\beta} - \bar{\gamma} \mid$ is small, then

$$\mid f(\bar{\alpha},\bar{\beta}) - f(\bar{\alpha},\bar{\gamma}) \mid \approx \mid \bar{\beta} - \bar{\gamma} \mid .$$

Thus, $1 + f(\bar{\beta},\bar{\gamma}) \geq \mid \bar{\beta} - \bar{\gamma} \mid$ and $f(\bar{\beta},\bar{\gamma}) = -1$ when the two directions coincide, $\bar{\beta} = \bar{\gamma}$.

Until now we have considered expected values relative to the hidden random variable $\lambda$ with the density function $\rho(\lambda)$ and we have denoted the expected value of a function $g(\bar{\theta},\lambda)$ as $E\left[g(\bar{\theta},\lambda)\right] = \int g(\bar{\theta},\lambda)\rho(\lambda)d\lambda$. Now we assume that the unit vectors $\bar{\alpha}$ and $\bar{\beta}$ rotate with a small angle around their current position. We denote by $\mathcal{E}\left[g(\bar{\theta})\right]$ the expected value of a function $g(\bar{\theta})$ when the unit vector $\bar{\theta}$ rotates with a small angle around its current value.

**Proposition.** *The expected value $\mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right]$ relative to $\bar{\alpha}$ and $\bar{\beta}$ of the observation based on a hidden variable model and the expected value $\mathcal{E}\left[\bar{\alpha}\cdot\bar{\beta}\right]$ predicted by Quantum Mechanics cannot be arbitrarily close to one another.*

*Proof:* Let us assume that there exist two positive arbitrarily small real numbers $\epsilon > 0$ and $\delta > 0$ such that

$$\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[\bar{\alpha}\cdot\bar{\beta}\right] \mid \leq \epsilon \quad \text{and} \quad \mid \mathcal{E}\left[\bar{\alpha}\cdot\bar{\beta}\right] - \bar{\alpha}\cdot\bar{\beta}) \mid \leq \delta.$$

First, we notice that these two conditions imply that

$$\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] + \bar{\alpha}\cdot\bar{\beta}) \mid \;\leq\; \epsilon + \delta.$$

When $\bar{\alpha} = \bar{\beta}$ then $\bar{\alpha}\cdot\bar{\beta} = \bar{\beta}\cdot\bar{\beta} = +1$ and:

$$\mid \mathcal{E}\left[f(\bar{\beta},\bar{\beta})\right] + 1 \mid \;\leq\; \epsilon + \delta,$$

or

$$\mid \mathcal{E}\left[f(\bar{\beta},\bar{\beta})\right] \mid \;\leq\; \epsilon + \delta - 1.$$

Recall that $a(\bar{\alpha},\lambda)$ and $b(\bar{\beta},\lambda)$ are outcomes of the measurements of the projections of the spins of the two particles along the two arbitrary directions $\bar{\alpha}$ and $\bar{\beta}$, respectively, thus $a(\bar{\alpha},\lambda) = \pm 1$ and $b(\bar{\beta},\lambda) = \pm 1$. It follows that

$$\mid \mathcal{E}[a(\bar{\alpha},\lambda)] \mid \;\leq 1 \quad \text{and} \quad \mid \mathcal{E}[b(\bar{\beta},\lambda)] \mid \;\leq 1.$$

Also,

$$1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)] \geq 0 \quad \text{and} \quad 1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)] \geq 0.$$

Now we rewrite the expression for the expected value of $f(\bar{\alpha},\bar{\beta})$

$$\mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] = \mathcal{E}\left(\int a(\bar{\alpha},\lambda)b(\bar{\beta},\lambda)\rho(\lambda)d\lambda\right) = \int \rho(\lambda)d\lambda\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)].$$

Then

$$\begin{aligned}
\mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] &= \textstyle\int \rho(\lambda)d\lambda\left(\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)] - \mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right) \\
&= \textstyle\int \rho(\lambda)d\lambda\left(\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)]\left[1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right]\right) \\
&- \textstyle\int \rho(\lambda)d\lambda\left(\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\left[1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)]\right]\right).
\end{aligned}$$

But $\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)] \leq 1$ and $\mathcal{E}[a(\bar{\alpha},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)] \leq 1$ thus,

$$\begin{aligned}
\mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] &\leq \textstyle\int \rho(\lambda)d\lambda\left(\left[1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right]\right) \\
&- \textstyle\int \rho(\lambda)d\lambda\left(\left[1 + \mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)]\right]\right)
\end{aligned}$$

It follows that

$$\begin{aligned} \mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] \mid &\leq \int \rho(\lambda)d\lambda \left(\left[1+\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right]\right) \\ &+ \int \rho(\lambda)d\lambda \left(\left[1+\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)]\right]\right). \end{aligned}$$

The first term on the right hand side of this inequality can be expressed as

$$\begin{aligned} \int \rho(\lambda)d\lambda \left(\left[1+\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right]\right) &= \int \rho(\lambda)d\lambda + \int \rho(\lambda)d\lambda \left(\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\gamma},\lambda)]\right) \\ &= 1+\mathcal{E}[f(\bar{\beta},\bar{\gamma})]. \end{aligned}$$

We also notice that

$$\int \rho(\lambda)d\lambda\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)] = \mathcal{E}\left[f(\bar{\beta},\bar{\beta})\right] \leq \epsilon+\delta-1.$$

Let us denote $\epsilon' = \epsilon+\delta \geq 0$; then

$$\int \rho(\lambda)d\lambda \left(\left[1+\mathcal{E}[a(\bar{\beta},\lambda)]\mathcal{E}[b(\bar{\beta},\lambda)]\right]\right) \leq \epsilon'.$$

The inequality becomes

$$\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] \mid \leq 1+\mathcal{E}[f(\bar{\beta},\bar{\gamma})]+\epsilon'.$$

We now evaluate the averages on both sides of this inequality. First, we notice that $\mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right]$ has an upper bound:

$$\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] + \bar{\alpha}\cdot\bar{\beta} \mid \;\leq\; \epsilon' \;\Longrightarrow\; \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] \leq \epsilon' - \bar{\alpha}\cdot\bar{\beta} \;\Longrightarrow\; \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] \leq \epsilon' - \mid \bar{\alpha}\cdot\bar{\beta} \mid .$$

Similarly,

$$\mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] \leq \epsilon' - \mid \bar{\alpha}\cdot\bar{\gamma} \mid \quad \text{and} \quad \mathcal{E}\left[f(\bar{\beta},\bar{\gamma})\right] \leq \epsilon' - \mid \bar{\beta}\cdot\bar{\gamma} \mid .$$

It follows that

$$\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] \mid \;\geq\; \mid \bar{\alpha}\cdot\bar{\gamma} - \bar{\alpha}\cdot\bar{\beta} \mid -2\epsilon'.$$

Finally, the inequality $\mid \mathcal{E}\left[f(\bar{\alpha},\bar{\beta})\right] - \mathcal{E}\left[f(\bar{\alpha},\bar{\gamma})\right] \mid \leq 1+\epsilon'+\mathcal{E}[f(\bar{\beta},\bar{\gamma})]$ becomes

$$\mid \bar{\alpha}\cdot\bar{\gamma} - \bar{\alpha}\cdot\bar{\beta} \mid -2\epsilon' \leq 1+\epsilon'+(\epsilon'-\bar{\beta}\cdot\bar{\gamma})$$

or,

$$4(\epsilon+\delta) \geq \mid \bar{\alpha}\cdot\bar{\gamma} - \bar{\alpha}\cdot\bar{\beta} \mid +\bar{\beta}\cdot\bar{\gamma} - 1.$$

This contradicts our hypothesis that given a small $\delta$ we can find an arbitrarily small $\epsilon$.

**Example.** *Assume* $\bar{\alpha} \cdot \bar{\gamma} = 0$, $\bar{\alpha} \cdot \bar{\beta} = \bar{\beta} \cdot \bar{\gamma} = 1/\sqrt{2}$. Then,

$$4(\epsilon + \delta) \geq | \, 0 - 1/\sqrt{2} \, | + 1/\sqrt{2} - 1,$$

or $\epsilon + \delta \geq 1/4(\sqrt{2} - 1)$.

We conclude that the expected value of the observation based on a hidden variable model and the expected values predicted by Quantum Mechanics cannot be arbitrarily close to one another. Next we discuss Bell inequality which provides deeper insights in the fallacy of hidden variable model.

## 2.18 Bell and CHSH Inequalities

John Bell investigated the separability of hidden variables and derived an inequality for the outcomes reported by two observers of an experiment in a setup similar to the one used for the EPR experiments. The more general CHSH inequality, named after Clauser, Horn, Shimony, and Holt [100] is related to Bell inequality.

Bell and CHSH inequalities establish an upper bound for the correlation of distant events based on the assumption that the principle of locality is valid. Bell inequality is based on two common sense assumptions:

1. *Locality* — measurements of different physical properties of different objects, carried out by different individuals, at distinct locations, cannot influence each other.
2. *Realism* — physical properties are independent of observations.

**Bell inequality.** We revisit the EPR experiment and consider a pair of photons emitted in opposite directions, Figure 57. There are now two independent observers, each using a polarization analyzer as the measurement apparatus; the first one may choose for her analyzer one of two possible settings, the orientation $\bar{\alpha}$ or $\bar{\gamma}$ relative to an arbitrary direction. For each orientation the outcome is unpredictable; if she chooses the orientation to be $\bar{\alpha}$, then the measured observable is called $\sigma_{\bar{\alpha}}$ and the two possible outcome values are $a = \pm 1$; if she chooses $\bar{\gamma}$ then the measured observable is $\sigma_{\bar{\gamma}}$ and the two possible outcome values are $c = \pm 1$. A similar scenario is true for the second observer; she may choose two possible settings: a new orientation, $\bar{\beta}$, or the same orientation $\bar{\gamma}$ as the first observer. If she chooses $\bar{\beta}$, then the measured observable is $\sigma_{\bar{\beta}}$ and the two possible outcome values are $b = \pm 1$; if she chooses $\bar{\gamma}$, then the measured observable is $\sigma_{\bar{\gamma}}$ and the outcome values are $c = \pm 1$. If both observers choose the same orientation $\bar{\gamma}$, then they obtain the same result, $c$. Out of the three possible measurement results, $a, b$, and $c$ only the pair $(a, b)$ or the single $c$ can be observed.

We wish to test the following hypothesis: the outcome of the experiment for the first observer depends only on *local hidden variables* related *only* to the first photon and to the analyzer used by the first observer. This property called *separability of hidden variables* is a consequence of Einstein locality principle which conjectures that "events occurring in a given space-time region are independent of the external parameters that may be controlled, at the same moment, by agents located in distant space-time regions." If the hidden variables have the separability property, then the outcome of the measurement performed by the first observer does not depend on the orientation of the analyzer used by the second observer.

First, we show that the three possible results $a$, $b$, and $c$ satisfy the equation

$$ab - ac \equiv \pm(1 - bc).$$

We evaluate separately the left- and the right-hand side of the equation for the two cases, correlated, $b = c$, and anti-correlated, $b \neq c$, photons:

$$\begin{array}{llll} (b = c) \;\Rightarrow & b - c = 0 \;\Rightarrow\; a(b-c) = 0 & \text{and} & 1 - bc = 1 - (\pm 1)(\pm 1) = 1 - 1 = 0 \\ (b \neq c) \;\Rightarrow & b - c = \pm 2 \;\Rightarrow\; a(b-c) = \pm 2 & \text{and} & 1 - bc = 1 - (\pm 1)(\mp 1) = 1 + 1 = 2. \end{array}$$

We repeat the experiment for many pairs of photons; if we accept the existence of variables we do not control, then the hidden variable for the experiment with the $i$-th photon must be different from the one for the $j$-th photon. From the previous equality it follows that the results of measurements for the $i$-th photon satisfy the relationship

$$a_i b_i - a_i c_i \equiv \pm(1 - b_i c_i).$$

For correlated and anti-correlated photons we have, respectively

$$b_i c_i = +1 \Longrightarrow a_i b_i - a_i c_i = \pm(1-1) = 0 \;\text{ and }\; b_i c_i = -1 \Longrightarrow a_i b_i - a_i c_i = \pm(1+1) = \pm 2.$$

The hidden variables are different for each value of the index $i$; the average of the products of pairs of hidden variable over the index $i$ gives us the correlation of hidden variables:

$$\langle ab \rangle = \frac{1}{N}\sum_{i=1}^{N} a_i b_i, \qquad \langle ac \rangle = \frac{1}{N}\sum_{i=1}^{N} a_i c_i, \qquad \text{and} \qquad \langle bc \rangle = \frac{1}{N}\sum_{i=1}^{N} b_i c_i.$$

Then $\langle ab \rangle - \langle ac \rangle = \pm(1 - \langle bc \rangle)$ and the well-known Bell inequality follows immediately

$$\mid \langle ab \rangle - \langle ac \rangle \mid \leq 1 - \langle bc \rangle.$$

**Clauser, Horn, Shimony, and Holt (CHSH) inequality.** We consider a similar setup as the one for Bell inequality, but the two observers do not share a common direction. The tests are performed for four settings: along the directions $\bar{\alpha}$ and $\bar{\gamma}$ for the first observer and $\bar{\beta}$ and $\bar{\delta}$ for the second one; the measurements produce the four pairs of outcome values, $a = \pm 1$, $c = \pm 1$, respectively, for the first observer and $b = \pm 1$, $d = \pm 1$, respectively, for the second observer. The four random variables, $a, c, b, d$ have a joint probability distribution function

$$p(q, r, s, t) = Prob(a = q, c = r, b = s, d = t).$$

Recall that the results of the measurements performed by two distinct observers can only be $a, c, b, d = \pm 1$ thus, we have two alternatives for observer 1

$$a + c = 0 \;\text{ and }\; a - c = \pm 2 \qquad \text{or} \qquad a + c = \pm 2 \;\text{ and }\; a - c = 0.$$

The identity $(a + c)b - (a - c)d \equiv \pm 2$ follows from the listing all possible combination of values of observable $a, b, c$, and $d$ in Table 1.

We repeat the experiment $N$ times; for the $i$-th pair of photons we have

$$(a_i + c_i)b_i - (a_i - c_i)d_i \equiv \pm 2.$$

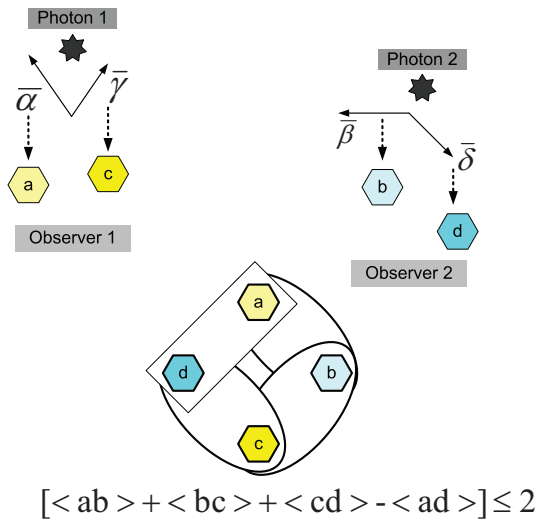



Figure 60: Two observers measure the polarization of two photons along four directions, $\bar{\alpha}$ and $\bar{\gamma}$ for the first observer and $\bar{\beta}$ and $\bar{\delta}$ for the second one, and produce the results $a = \pm 1$, $c = \pm 1$, respectively, for the first observer and $b = \pm 1$, $d = \pm 1$, respectively, for the second observer. The expected values of the four observables satisfy the CHSH inequality.

Table 1: Truth table listing possible combination of values of observable for the CHSH inequality.

| a | c | b | d | a+c | a-c | (a+c)b | (a-c)d | (a+c)b - (a-c)d |
|---|---|---|---|---|---|---|---|---|
| +1 | +1 | +1 | +1 | +2 | 0 | +2 | 0 | + 2 |
| +1 | +1 | +1 | -1 | +2 | 0 | +2 | 0 | + 2 |
| +1 | +1 | -1 | +1 | +2 | 0 | -2 | 0 | - 2 |
| +1 | +1 | -1 | -1 | +2 | 0 | -2 | 0 | - 2 |
| +1 | -1 | +1 | +1 | 0 | +2 | 0 | +2 | - 2 |
| +1 | -1 | +1 | -1 | 0 | +2 | 0 | -2 | + 2 |
| +1 | -1 | -1 | +1 | 0 | +2 | 0 | +2 | - 2 |
| +1 | -1 | -1 | -1 | 0 | +2 | 0 | -2 | + 2 |
| -1 | +1 | +1 | +1 | 0 | -2 | 0 | -2 | + 2 |
| -1 | +1 | +1 | -1 | 0 | -2 | 0 | +2 | - 2 |
| -1 | +1 | -1 | +1 | 0 | -2 | 0 | -2 | + 2 |
| -1 | +1 | -1 | -1 | 0 | -2 | 0 | +2 | - 2 |
| -1 | -1 | +1 | +1 | -2 | 0 | -2 | 0 | - 2 |
| -1 | -1 | +1 | -1 | -2 | 0 | -2 | 0 | - 2 |
| -1 | -1 | -1 | +1 | -2 | 0 | +2 | 0 | + 2 |
| -1 | -1 | -1 | -1 | -2 | 0 | +2 | 0 | + 2 |

The average of the $N$ measurements is

$$\frac{1}{N}\sum_{i=i}^{N}[(a_i+c_i)b_i-(a_i-c_i)d_i]\equiv\pm 2,$$

or $\langle ab\rangle+\langle bc\rangle+\langle cd\rangle-\langle ad\rangle\equiv\pm 2$. After takeing the modulus we obtain the CHSH inequality

$$\mid\langle ab\rangle+\langle bc\rangle+\langle cd\rangle-\langle ad\rangle\mid\ \leq\ 2.$$

We conclude that when a quantum system consists of several disjoint systems, as in the case of the EPR experiment, the hidden variables do not satisfy the *separability* requirement, there is an upper bound for their correlation thus, they do not fall in disjoint subsets.

## 2.19 Violation of Bell Inequality

In the previous section we have seen that if we accept two common sense assumptions collectively known as *local realism* then the results of the measurements must satisfy Bell and CHSH inequalities. We show now that *a pair of maximally entangled qubits violates Bell inequality.* Consider a pair of entangled qubits in the singlet state $\mid\beta_{11}\rangle=\mid\psi\rangle=(1/\sqrt{2})(\mid 01\rangle-\mid 10\rangle)$. Alice has one of the entangled particles and Bob the other. We have several one qubit gates used to transform the state of one qubit; among them are the $X$ gate (it transposes the components of a qubit) and the $Z$ gate (it flips the sign of a qubit). The Pauli transformations performed by the two gates are respectively:

$$\sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix}\qquad\text{and}\qquad\sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

Alice measures two observables, $a$ and $c$ by applying the operators $\mathbf{A}=\sigma_z$ and $\mathbf{C}=\sigma_x$. Alice measures her qubit in the eigenbasis of $\sigma_z$ after it has been acted on by the $Z$ gate and in the eigenbasis of $\sigma_x$ after it has been acted on by the $X$ gate. Similarly, Bob measures two observables $b$ and $d$ by applying the operators $\mathbf{B}=-1/\sqrt{2}(\sigma_z+\sigma_x)$ and $\mathbf{D}=1/\sqrt{2}(\sigma_z-\sigma_x)$, Figure 61.

Recall that the outcomes of measurements are eigenvalues of self-adjoint operators (projectors), and are real numbers; quantum transformations are described by unitary operators. Pauli transformations are both self-adjoint (thus, potentially physical observables) and unitary (thus, potentially gates). The eigenvectors of the transformation $\sigma_z$ are $\mid 0\rangle$ and $\mid 1\rangle$ and the corresponding eigenvalues are $+1$ and $-1$, respectively. The eigenvectors of the transformation $\sigma_x$ are $1/\sqrt{2}(\mid 0\rangle+\mid 1\rangle)$ and $1/\sqrt{2}(\mid 0\rangle-\mid 1\rangle)$ and the corresponding eigenvalues are $+1$ and $-1$, respectively.

**Proposition.** *The average values of the observables are*

$$\langle ab\rangle=\langle cb\rangle=\langle cd\rangle=\frac{1}{\sqrt{2}}\qquad\textit{and}\qquad\langle ad\rangle=-\frac{1}{\sqrt{2}}.$$

*Thus,* $\langle ab\rangle+\langle cb\rangle+\langle cd\rangle-\langle ad\rangle=2\sqrt{2}$ *in violation of Bell inequality.*

*Proof:* Indeed,

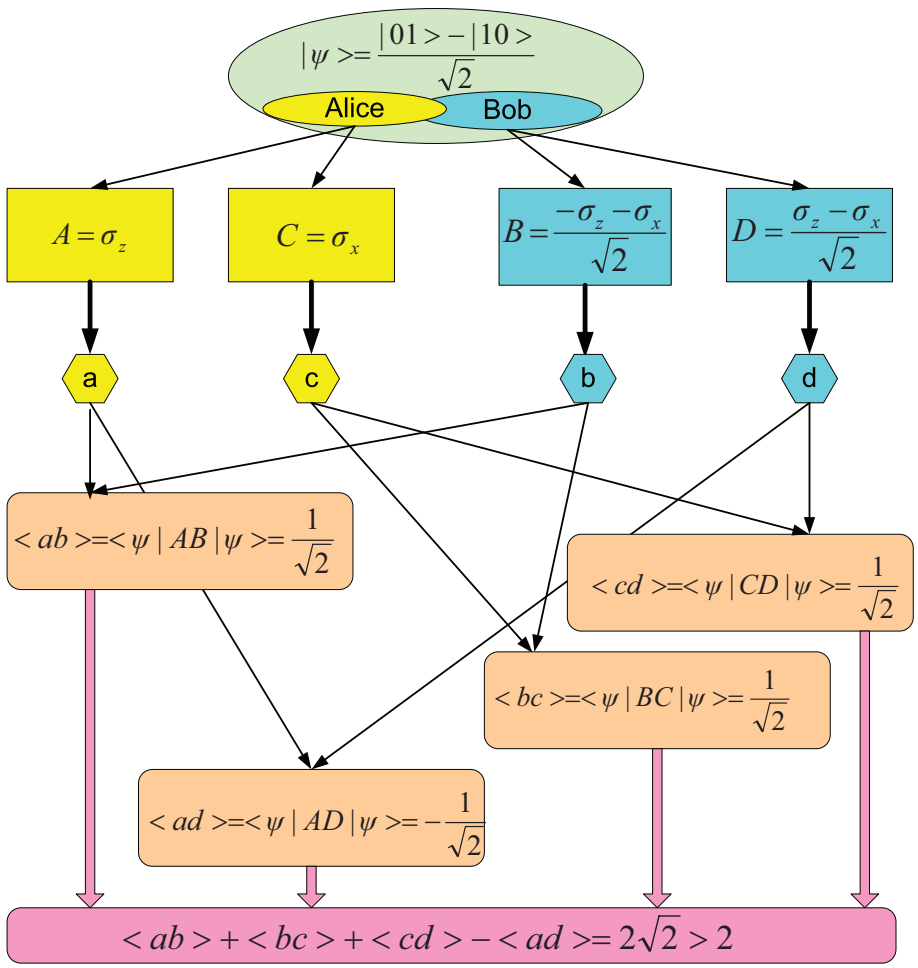



Figure 61: A pair of entangled particles, one belonging to Alice and the other to Bob, are in state $|\ \psi\rangle$. Alice measures two observables $a$ and $c$ by applying the operators $\mathbf{A}=\sigma_z$ and $\mathbf{C}=\sigma_x$. Bob measures two observables $b$ and $d$ by applying the operators $\mathbf{B}=-1/\sqrt{2}(\sigma_z+\sigma_x)$ and $\mathbf{D}=1/\sqrt{2}(\sigma_z-\sigma_x)$. The expected values of the four observables violate Bell inequality.

$$-\sigma_z-\sigma_x=\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \qquad \text{and} \qquad \sigma_z-\sigma_x=\begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix}.$$

It follows that

$$\mathbf{AB}=\sigma_z\otimes\frac{1}{\sqrt{2}}(-\sigma_z-\sigma_x)=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\otimes\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix} -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

$$\mathbf{CB} = \sigma_x \otimes \frac{1}{\sqrt{2}}(-\sigma_z - \sigma_x) = -\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -\frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix},$$

$$\mathbf{CD} = \sigma_x \otimes \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix},$$

and

$$\mathbf{AD} = \sigma_z \otimes \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Now we compute $\langle ab \rangle$, $\langle cb \rangle$, $\langle cd \rangle$, and $\langle ad \rangle$ the average values of the four observables, $ab, cb, cd$, and $ad$, respectively. Recall that when we have $n$ identical copies of a quantum system in state $\mid \psi \rangle$ the expected value of the observable associated with operator $\mathbf{P}$ applied to the systems in state $\mid \psi \rangle$ is

$$\langle \mathbf{P} \rangle = \langle \psi \mid \mathbf{P} \mid \psi \rangle.$$

In our case the state of the two entangled qubits is

$$\mid \psi \rangle = \frac{\mid 01 \rangle - \mid 10 \rangle}{\sqrt{2}} \quad \Longrightarrow \quad \mid \psi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

It follows that

$$\langle ab \rangle = \langle \psi \mid \mathbf{AB} \mid \psi \rangle.$$

Call

$$\mid \xi \rangle = \mathbf{AB} \mid \psi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

Then,

$$\langle ab \rangle = \langle \psi \mid \xi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

Similarly,

$$\langle cb\rangle = \langle\psi \mid \mathbf{CB} \mid \psi\rangle.$$

Call

$$\mid \xi'\rangle = \mathbf{CB} \mid \psi\rangle = -\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

Then,

$$\langle cb\rangle = \langle\psi \mid \xi'\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}(-\frac{1}{2})\begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

For the third average we have

$$\langle cd\rangle = \langle\psi \mid \mathbf{CD} \mid \psi\rangle.$$

Call

$$\mid \xi''\rangle = \mathbf{CD} \mid \psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

Then,

$$\langle cd\rangle = \langle\psi \mid \xi''\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}}.$$

Finally,

$$\langle ad\rangle = \langle\psi \mid \mathbf{AD} \mid \psi\rangle$$

Call

$$\mid \xi'''\rangle = \mathbf{AD} \mid \psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

Then,

$$\langle ad\rangle = \langle\psi \mid \xi'''\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & -1 & 0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = -\frac{1}{\sqrt{2}}.$$

It follows that

$$\langle ab\rangle + \langle bc\rangle + \langle cd\rangle - \langle ad\rangle = 2\sqrt{2}.$$

Thus, Quantum Mechanics predicts a value for the sum of the average values of observables in violation of Bell inequality.

An interesting question is if mixed entangled states violate Bell inequality. There are some mixed entangled states called Werner states [454] which do not violate Bell inequality for any number of local measurements.

## 2.20 Entanglement and Hidden Variables

John Bell provided deeper insights in the limitations of a hidden variables theory; he showed that *it is possible to mimic the statistical properties of quantum theory by a deterministic model based on hidden variables.* He constructed such a deterministic model generating results whose averages are identical to those predicted by Quantum Mechanics. When the quantum system is made up of disjoint subsystems entangled like those in the EPR experiment, then the hidden variables cannot be separated as disjoint subsets, local to each subsystem. Let us follow the arguments of John Bell.

Consider a spin one-half particle and an observable $A$ such that the associated measurement operator is $\mathbf{A} = \mathbf{m}\cdot\sigma$, where $\sigma$ are the Pauli matrices and $\mathbf{m}$ some arbitrary real numbers. The model is based on the assumption that the outcome of an experiment is determined by:

1. Two elements we can control:

1.1 The state $\mid \psi\rangle$ in which we prepare the system, and

1.2 The eigenvalues $\mathbf{m}$ of the operator $\mathbf{A}$, which can take values $\pm m$; they are parameters of the measuring apparatus.

2. One element we cannot control, a hidden variable uniformly distributed in the interval $-1 \leq \lambda \leq +1$.

The model is deterministic, *once we choose a value for the hidden variable, $\lambda$, the result of the measurement is certain.* The following table gives the result of a measurement and the probability of that result for a given range of the hidden variable $\lambda$:

| Range of $\lambda$ | Result of measurement | Probability |
|---|---|---|
| $-1 < \lambda < -\langle\psi \mid \mathbf{A} \mid \psi\rangle/m$ | $-m$ | $p_{(-m)} = (1 - \langle\psi \mid \mathbf{A} \mid \psi\rangle/m)/2$ |
| $-\langle\psi \mid \mathbf{A} \mid \psi\rangle/m < \lambda < +1$ | $+m$ | $p_{(+m)} = (1 + \langle\psi \mid \mathbf{A} \mid \psi\rangle/m)/2$ |

The average value of the result of a measurement based on this deterministic model which includes a uniformly distributed hidden variable is

$$\begin{aligned}\langle\mathbf{A}\rangle &= -mp_{-m} + mp_{+m} = -m[(1 - \langle\psi \mid A \mid \psi\rangle/m)/2] + m[(1 + \langle\psi \mid \mathbf{A} \mid \psi\rangle/m)/2] \\ &= \langle\psi \mid \mathbf{A} \mid \psi\rangle,\end{aligned}$$

in perfect agreement with Quantum Mechanics.

Bell showed that his model does not satisfy the separability requirement when the quantum system is in an entangled state. If the quantum system is made up of disjoint subsystems, entangled like those in the EPR experiment, then the hidden variables cannot be separated as disjoint subsets, local to each subsystem.

Is there any relationship between the local hidden variable models and shareability of entanglement? This question was answered recently [420]. A local hidden variable model assumes the existence of a random variable shared by two observers, Alice, $A$, and Bob, $B$; this variable is used to locally generate a measurement outcome. The outcome depends only on the choice of the local measurement and not on the choice of the remote observer's measurement. If each observer has a set of local measurements, $\{i = 1, 2, \ldots, m_A\}$ for Alice with $\lambda_A(i)$ possible outcomes per measurement and $\{k = 1, 2, \ldots, m_B\}$ measurements with $\lambda_B(k)$ outcomes per measurement for Bob, a local hidden variable model generates all the probability values $P_{ij,kl}$ that Alice's $i$-th measurement has outcome $j$ and Bob's $k$-th measurement has outcome $l$. Each local measurement has a single outcome with probability one independently of the measurement made by the other observer. For a composite quantum system described by the density matrix $\rho$ in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ the probability $P_{ij,kl}$ is given by

$$P_{ij,kl}(\rho) = \mathrm{tr}[(\mathbf{M}^A_{ij} \otimes \mathbf{M}^B_{kl})\rho]$$

where $\{\mathbf{M}^A_{ij}\}$, with $\mathbf{M}^A_{ij} \geq 0$ and $\sum_j \mathbf{M}^A_{ij} = I_A$, are POVM elements for the $i$-th measurement of observer $A$ and $\{\mathbf{M}^B_{kl}\}$, with $\mathbf{M}^B_{kl} \geq 0$ and $\sum_l \mathbf{M}^B_{kl} = I_B$, are POVM elements of the $k$-th measurement of observer $B$.

A Bell inequality is violated if and only if the probability value cannot be generated by a local hidden variable model. We can say that Bell inequality quantifies how measurements [23] on entangled quantum systems can invalidate local classical models of reality. It is shown [420] that violations of Bell inequalities can be connected to the existence of a symmetric extension/quasiextension of a quantum state. An extension of a quantum state $\rho$ of a system $AB$ is another quantum state defined on a system $ABC$ such that when we trace over $C$ we obtain the initial quantum state $\rho$. The system $C$ must be of the form

$$C = A^{\otimes(m_A-1)} \otimes B^{\otimes(m_B-1)}$$

and the extension state must be invariant under all permutations of the $m_A$ copies of $A$ among each other and under any permutation of the $m_B$ copies of $B$. Such an extension exists if the quantum state $\rho$ of the system $AB$ is separable, that is

$$\rho = \sum_i p_i(\mid \psi_i\rangle\langle\psi_i \mid)_A \otimes (\mid \varphi_i\rangle\langle\varphi_i \mid)_B.$$

The extension state, $\rho'$, is obtained by copying the individual product states onto the other spaces:

$$\rho' = \sum_i p_i(\mid \psi_i\rangle\langle\psi_i \mid)^{\otimes m_A} \otimes (\mid \varphi_i\rangle\langle\varphi_i \mid)^{\otimes m_B}.$$

[23] The types and number of local measurements the observers can perform are not bounded and, correspondingly, the enumeration of Bell inequalities is computationally hard.

**Proposition.** *Consider two Hilbert spaces, $\mathcal{H}_A$ and $\mathcal{H}_{B_1}$; let $\rho$ be a density matrix on $\mathcal{H}_A \otimes \mathcal{H}_{B_1}$. If $\rho$ has a symmetric extension, $\rho'$, on $\mathcal{H}_A \otimes \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2} \ldots \otimes \mathcal{H}_{B_m}$ then there exists a local hidden variable description of $\rho$ when one of the observers performs an arbitrary number of measurements and the other observer performs $m$ measurements.*

The intuition behind this result comes from several properties of hidden variables and measurements. When one of the two parties performs only one measurement of a quantum state then there is a local hidden variable model describing the quantum state. The second observer may carry out her $m$ measurements $\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_m$ in the $m$ Hilbert spaces, $\mathcal{H}_{B_1}, \mathcal{H}_{B_2} \ldots \mathcal{H}_{B_m}$, respectively. The joint probability of an outcome of a measurement $\mathbf{M}_A$ and some $\mathbf{M}_{B_i}, 1 \leq i \leq m$ are the same for $\rho$ and $\rho'$ because $\rho'$ is a symmetric extension of $\rho$. The $m$ measurements can be viewed as a single measurement on $\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2} \ldots \otimes \mathcal{H}_{B_m}$. Thus, the measurements on $\rho'$ have a local hidden variable description from which we can deduce the local hidden variable description of the original measurements on $\rho$.

If the quantum state $\rho$ is a pure entangled state (non-separable), such a symmetric extension does not exist since pure entanglement is monogamous. Therefore, a violation of a Bell inequality signals that the system is in a pure entangled state; the entanglement in this state of the composite quantum system is monogamous.

## 2.21 Quantum and Classical Correlations

To answer the question whether the correlation of quantum measurements is stronger than the correlation of classical measurements we consider two simple systems exhibiting some degree of similarity:

1. a quantum system consisting of two one-half spin particles in a singlet state; and
2. a classical system consisting of a projectile initially at rest, which explodes and produces two fragments carrying opposite momenta.

We assume that we have $N$ copies of each system and we are able to measure in each case two values $a$ and $b$ of an observable, compute their correlation function, $ab$, and obtain ensemble averages, $\langle a \rangle$, $\langle b \rangle$, and $\langle ab \rangle$. We show that the expected value of the correlation of the quantum measurements is stronger than the correlation of classical ones.

We examine first the quantum system; let $\bar{\alpha}$ and $\bar{\beta}$ two arbitrarily chosen unit vectors. The observables for the two particles are respectively $\bar{\alpha} \cdot \sigma_1$ and $\bar{\beta} \cdot \sigma_2$, with $\sigma_1$ and $\sigma_2$ the Pauli spin matrices for the two particles. The system is in a singlet state thus:

$$\sigma_2 \mid \psi \rangle = -\sigma_1 \mid \psi \rangle.$$

The results of the measurements are $a, b = \pm 1$. If we repeat an experiment similar to the one suggested by Bohm then quantum mechanics predicts that the expected value of the results is equal to zero

$$\langle a \rangle = 0 \quad \text{and} \quad \langle b \rangle = 0.$$

The expected value of the quantum correlation is

$$\langle ab \rangle^{quantum} = \langle \psi^\dagger \mid (\bar{\alpha} \cdot \sigma_1)(\bar{\beta} \cdot \sigma_2) \mid \psi \rangle.$$

We apply the identity [326]

$$(\bar{\alpha}\cdot\sigma)(\bar{\beta}\cdot\sigma) \equiv \bar{\alpha}\cdot\bar{\beta} + i(\bar{\alpha}\times\bar{\beta})\cdot\sigma$$

to the system in a singlet state, $\sigma_2 \mid \psi\rangle = -\sigma_1 \mid \psi\rangle$, and conclude that

$$\langle ab\rangle^{quantum} = -\bar{\alpha}\cdot\bar{\beta} = -\cos\theta$$

with $\theta$ the angle between the directions of the two vectors $\bar{\alpha}$ and $\bar{\beta}$ freely chosen by the observer. The largest correlation occurs for $\theta = \pi$ and then

$$\max_\theta \langle ab\rangle^{quantum} = -\cos\pi = +1.$$

Let us now discuss the classical system; the angular momentum of the projectile at rest is $\mathbf{J}$ and the angular momenta of the two fragments after the explosion are: $\mathbf{J}_2 = -\mathbf{J}_1$. As before, $\bar{\alpha}$ and $\bar{\beta}$ are two freely chosen unit vectors and $\theta$ the angle between them. The experiment is repeated $N$ times; the observer of the first fragment measures the observable $\mathrm{sign}(\bar{\alpha}\cdot\mathbf{J}_1)$ and obtains the results $a_i = \pm 1$, while the observer of the second fragment measures the observable $\mathrm{sign}(\bar{\beta}\cdot\mathbf{J}_2)$ and obtains the results $b_i = \pm 1,\ 1 \le i \le N$. The expected values of the outcomes of the two measurements are respectively

$$\langle a\rangle = \frac{1}{N}\sum_{i=1}^{N} a_i \quad \text{and} \quad \langle b\rangle = \frac{1}{N}\sum_{i=1}^{N} b_i.$$

These values are typically of the order $1/\sqrt{N}$, thus very close to zero. The correlation of the two measurements is given by

$$\langle ab\rangle^{classical} = \frac{1}{N}\sum_{i=1}^{N} a_i b_i,$$

an expression which does not vanish to zero. Simple mechanical considerations show that if the momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ are uniformly distributed then,

$$\langle ab\rangle^{classical} = \frac{\theta - (\pi - \theta)}{\pi} = -1 + \frac{2\theta}{\pi}.$$

The largest correlation occurs when $\theta = \pi$

$$\max_\theta \langle ab\rangle^{classical} = -1 + \frac{2\theta}{\pi} = -1 + 2 = +1.$$

In the interval $\theta \in [0, \pi]$ the two correlations are equal, with values 0 or $\pm 1$, when $\theta$ takes the values $\frac{\pi}{2}$ or 0 and $\pi$, respectively. In the rest of the interval, the classical correlation is a linear function, while the quantum correlation is a cosine function which runs below the straight line in the negative value range for $\theta \in (0, \frac{\pi}{2})$ and, respectively, above it in the positive value range for $\theta \in (\frac{\pi}{2}, \pi)$.

Thus, the inequality

$$\mid \langle ab\rangle^{quantum} \mid \ge \langle ab\rangle^{classical}$$

leads us to the conclusion that *Quantum Mechanics, based on a nondeterministic model, allows a stronger correlation of measurements than Classical Mechanics, based on a deterministic model of the physical world.*

We now switch our focus from theoretical considerations regarding the measurement process to practical applications and start our analysis with a discussion of the measurements of quantum circuits.

## 2.22 Measurements and Quantum Circuits

To report the results of any transformation produced by quantum circuits we have to carry out a measurement; we start our discussion with an application of quantum measurements to circuits used to detect errors during a quantum error correction process and then present two principles important for the measurement of quantum circuits.

Quantum error correction is based on our ability to detect if a quantum state is in error without actually measuring the system and thus altering its state. This topic is presented at length in Chapter 5 where we discuss how to determine the *syndrome* associated with a quantum state. The stabilizer formalism allows us to establish if a system of $n$ qubits has suffered decoherence and is in error by computing the eigenvalues of a group of operators, $\mathbf{M}_i$. We analyze next a simple circuit to measure the eigenvalues of a single-qubit operator $\mathbf{U}$ with two eigenvalues $\lambda = \pm 1$. Such an operator is unitary, thus we can regard $\mathbf{U}$ as a one-qubit gate; $\mathbf{U}$ is also self-adjoint and we can regard it as an observable.

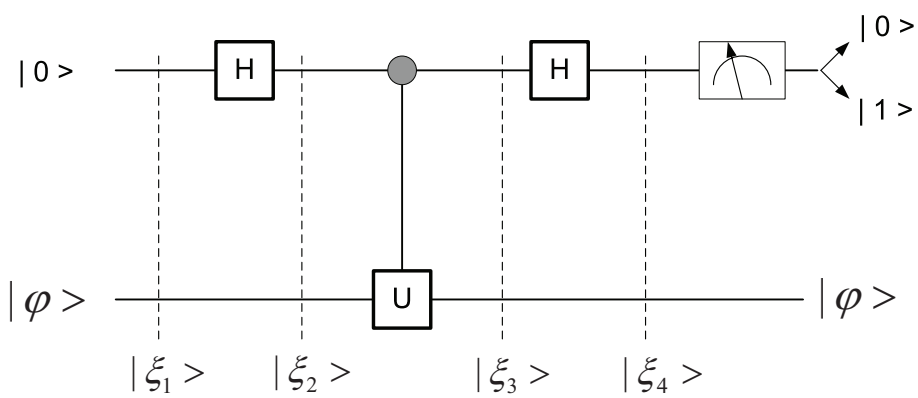



Figure 62: A circuit used to measure the eigenvalues of a single-qubit operator $\mathbf{U}$ with two eigenvalues $\lambda = \pm 1$. The upper output qubit is set to $\mid 0\rangle$ if the eigenvalue of $\mathbf{U}$ is $\lambda = +1$ and to $\mid 1\rangle$ if the eigenvalue $\mathbf{U}$ is $\lambda = -1$. The lower output qubit is $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$, the corresponding eigenvector of $\mathbf{U}$.

**Example.** *A circuit used to measure the eigenvalues of a single-qubit operator.* The circuit in Figure 62 produces an eigenvector of the operator $\mathbf{U}$; the upper qubit is set to $\mid 0\rangle$ if the corresponding eigenvalue is $\lambda = +1$, and to $\mid 1\rangle$ if the eigenvalue is $\lambda = -1$.

We want to determine the unitary transformation

$$\mathbf{U} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}$$

necessary to carry out the computation of the eigenvalues carried out by the circuit. Let $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ be an eigenvector of $\mathbf{U}$. The two eigenvalues, $\lambda_1 = +1$ and $\lambda_2 = -1$ identified after measurement of the first qubit are associated with different expressions for $\mid \xi_4\rangle$ in Figure 62:

$$\lambda_1 = +1 \quad \Longrightarrow \quad \mathbf{U} \mid \varphi\rangle = (+1) \mid \varphi\rangle \quad \Longrightarrow \quad \mid \xi_4\rangle = \mid 0\rangle(\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) = \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle$$

and

$$\lambda_2 = -1 \quad \Longrightarrow \quad \mathbf{U} \mid \varphi\rangle = (-1) \mid \varphi\rangle \quad \Longrightarrow \quad \mid \xi_4\rangle = \mid 1\rangle(\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) = \alpha_0 \mid 10\rangle + \alpha_1 \mid 11\rangle.$$

The states $\mid \xi_1\rangle$, $\mid \xi_2\rangle$, $\mid \xi_3\rangle$, and $\mid \xi_4\rangle$ in Figure 62 are:

$$\mid \xi_1\rangle = \mid 0\rangle \otimes \mid \varphi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ 0 \\ 0 \end{pmatrix},$$

$$\mid \xi_2\rangle = (H \otimes I) \mid \xi_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_0 \\ \alpha_1 \end{pmatrix},$$

$$\mid \xi_3\rangle = G_{controlled-U} \mid \xi_2\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & u_{11} & u_{12} \\ 0 & 0 & u_{21} & u_{22} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_0 \\ \alpha_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_0 u_{11} + \alpha_1 u_{12} \\ \alpha_0 u_{21} + \alpha_1 u_{22} \end{pmatrix},$$

and

$$\mid \xi_4\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_0 u_{11} + \alpha_1 u_{12} \\ \alpha_0 u_{21} + \alpha_1 u_{22} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \alpha_0 + \alpha_0 u_{11} + \alpha_1 u_{12} \\ \alpha_1 + \alpha_0 u_{21} + \alpha_1 u_{22} \\ \alpha_0 - \alpha_0 u_{11} - \alpha_1 u_{12} \\ \alpha_1 - \alpha_0 u_{21} - \alpha_1 u_{22} \end{pmatrix}.$$

The eigenvalues, $\lambda_1 = +1$ and $\lambda_2 = -1$ require different relations among the elements of $\mathbf{U}$; for $\lambda_1 = +1$ we have

$$\lambda_1 = +1 \quad \Longrightarrow \quad \mid \xi_4\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \begin{cases} \alpha_0(1 - u_{11}) = \alpha_1 u_{12} \\ \alpha_1(1 + u_{22}) = \alpha_0 u_{21} \end{cases}$$

This implies that $(1-u_{11})(1+u_{22}) = u_{12}u_{21}$ thus,

$$u_{11} - u_{22} + u_{11}u_{22} + u_{12}u_{21} = 1.$$

For $\lambda_2 = -1$ we have

$$\lambda_2 = -1 \quad \Longrightarrow \quad \mid \xi_4\rangle = \begin{pmatrix} 0 \\ 0 \\ \alpha_0 \\ \alpha_1 \end{pmatrix} \quad \Longrightarrow \quad \begin{cases} \alpha_0(1+u_{11}) = -\alpha_1 u_{12} \\ \alpha_1(1+u_{22}) = -\alpha_0 u_{21} \end{cases}$$

Then $(1+u_{11})(1+u_{22}) = -u_{12}u_{21}$ and

$$u_{11} + u_{22} + u_{11}u_{22} + u_{12}u_{21} = -1.$$

It follows that $u_{22} = -1$, $u_{21} = u_{12} = 0$, and $u_{11} = 1$; thus, the transformation $\mathbf{U}$ in the eigenbasis is

$$\mathbf{U} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = Z.$$

The <u>principle of deferred measurement</u>: we have an algorithm that performs some quantum operations, then measures, and then applies more quantum operations, depending on the outcome of the quantum measurement; *we can simulate this entire algorithm by a completely quantum algorithm with only one final measurement.* We replace classically controlled quantum operations with quantum controlled quantum operations.

**Example.** *Quantum teleportation* The quantum teleportation revisited in Section 2.13, uses the circuit in Figure 63 (a). The topmost two qubits belong to Alice. The first one, whose state is to be teleported, is in state $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ and the second is entangled with Bob's qubit; the pair of qubits is in a maximally entangled state, $\mid \beta_{00}\rangle$. The states at different stages of the quantum circuit in Figure 63 (a) are:

$$\mid \xi_0\rangle = \frac{1}{\sqrt{2}}[\alpha_0 \mid 0\rangle(\mid 00\rangle + \mid 11\rangle) + \alpha_1 \mid 1\rangle(\mid 00\rangle + \mid 11\rangle)]$$

$$\mid \xi_1\rangle = \frac{1}{\sqrt{2}}[\alpha_0 \mid 0\rangle(\mid 00\rangle + \mid 11\rangle) + \alpha_1 \mid 1\rangle(\mid 10\rangle + \mid 01\rangle)]$$

and

$$\begin{aligned} \mid \xi_2\rangle \quad &= \tfrac{1}{2}[\quad \mid 00\rangle(\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) + \mid 01\rangle(\alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle) \\ &+ \quad \mid 10\rangle(\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle) + \mid 11\rangle(\alpha_0 \mid 1\rangle - \alpha_1 \mid 0\rangle)]. \end{aligned}$$

The state $\mid \xi_2\rangle$ is a sum of four terms, each term consisting of the tensor product of the state of two qubits in Alice's possession with the state of the qubit in Bob's possession. If Alice measures her two qubits and finds them in state $\mid 00\rangle$ then Bob's qubit must be in state $\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$, which is the original state $\mid \varphi\rangle$. Thus, when Bob learns that the result of the measurement is 00 he applies the identity transformation $I$ to his qubit. Similarly, if Alice

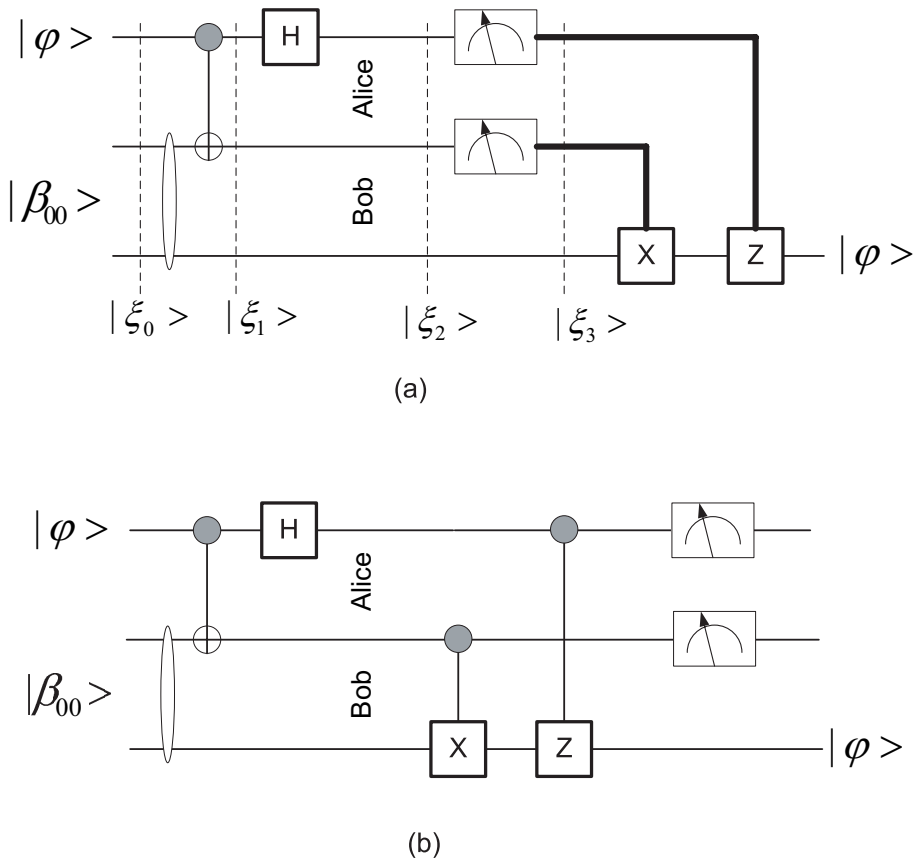



Figure 63: (a) The original circuit for teleportation. Thick lines correspond to a classical communication channel used by Alice to communicate the result of her measurement of the pair in her possession, $00, 01, 10$, or $11$. (b) The circuit with deferred measurements.

measures her two qubits and finds them in state $\mid 01\rangle$, then Bob's qubit must be in state $\alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle = X \mid \varphi\rangle$, thus Bob must apply the $X$ transformation to his qubit; if the result of the measurement is 10 then Bob's qubit must be in state $\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle = Z \mid \varphi\rangle$ and he should apply a $Z$ transform to recover the original state; finally, if the result of the measurement is 11 then Bob's qubit must be in state $\alpha_0 \mid 1\rangle - \alpha_1 \mid 0\rangle = Y \mid \varphi\rangle$ and then he should apply a $Y = XZ$ transform to recover the original state.

According to the principle of deferred measurement, instead of the circuit in Figure 63 (a) we can use the one in Figure 63 (b) where the measurements are moved to the end of the circuit; then, the concept of "teleportation," as discussed in this section is altered, no classical information is sent from Alice to Bob.

Another important concept related to the measurement of quantum circuits is provided by the principle of implicit measurement: *if some qubits in a computation are never used again, one can assume that they have been measured and that the result of the measurement has been ignored.* Formally, this means that the reduced density operator of the remaining qubits, the ones not measured, is the same as if the first group of qubits would have been measured and the result of the measurement ignored.

The important idea here is that a measurement is irreversible if and only if we examine its results and gather information about the quantum state. If the result of a measurement does not reveal any information about the state of the quantum system then the measurement is reversible. It is intuitively obvious that if we totally ignore the results of a measurement the state of the system is not affected.

For example, consider a two-qubit system. Let us perform a measurement in the computational basis of the second qubit. The following proposition is true:

**Proposition.** *Let $\rho$ be the density operator of the original two-qubit system and $\rho'$ the density operator of the system after a measurement by an observer who does not learn the result of the measurement. Then:*

$$\rho' = P_0 \rho P_0 + P_1 \rho P_1.$$

Here $P_0$ and $P_1$ are the projectors on the basis vectors of the states of the second qubit,

$$P_0 = \mid 0\rangle\langle 0 \mid \qquad \text{and} \qquad P_1 = \mid 1\rangle\langle 1 \mid .$$

Moreover, the reduced density operator for the first qubit is not affected by the measurement:

$$\mathrm{tr}_2(\rho) = \mathrm{tr}_2(\rho').$$

**Proof**: We rewrite the expression for $\rho'$

$$\begin{aligned} \rho' &= \mid 0\rangle\langle 0 \mid \rho \mid 0\rangle\langle 0 \mid + \mid 1\rangle\langle 1 \mid \rho \mid 1\rangle\langle 1 \mid \\ &= \mid 0\rangle\langle 0 \mid \rho \mid 0\rangle\langle 0 \mid + (I - \mid 0\rangle\langle 0 \mid)\rho(I - \mid 0\rangle\langle 0 \mid) \\ &= I\rho I = \rho. \end{aligned}$$

A corollary of the principle of implicit measurement is that *measurement commutes with control.* Suppose we use a qubit to control a one-qubit quantum operation $U$, and then we measure the qubit. We could equivalently first measure the control qubit, and then use the measurement result to classically control the application of $U$.

## 2.23 Measurements and Ancilla Qubits

A measurement of a quantum system reveals classical information. For example, if a qubit $Q$ is in state $\mid \psi_Q\rangle = \mid 0\rangle$, then the result of the measurement is one bit of information, $q = 0$ with probability $p = 1$. If a qubit $R$ is in state $\mid \psi_R\rangle = 1/\sqrt{2}(\mid 0\rangle + \mid 1\rangle)$, then the result of the measurement can be $q = 0$ with probability $p_0 = 0.5$, or $q = 1$ with probability $p_1 = 0.5$.

Assume that the two qubits, $Q$ and $R$, are inside a black box and we select one of them at random and perform a measurement. The original states of the two qubits are:

$$\mid \psi_Q\rangle = \mid 0\rangle \qquad \text{and, respectively} \qquad \mid \psi_R\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}.$$

As a result of the measurement we obtain a classical bit of information, $q = 0$ or $q = 1$. When $q = 1$ we are certain that the original qubit was in state $\mid \psi_R\rangle = 1/\sqrt{2}(\mid 0\rangle + \mid 1\rangle)$ thus, we have selected $R$. When the result is $q = 0$ we cannot be certain; there is a $50 - 50$ chance

that we have selected $Q$ or $R$. It follows that we can be absolutely certain that we inferred correctly the state of the qubit selected from the black box only in 25% of the trials. Indeed, we select $R$ in 50% of the total cases and if we have selected $R$ we guess right only in 50% of those selected cases (50% of 50%).

Is there any transformation of the quantum states of the two qubits which allows us to increase our chances of guessing correctly? More precisely, we plan to transform the two qubits inside the black box, and then select and measure one of them. It is easy to see that the same one-qubit transformation applied to both qubits does not increase our probability to guess right. For example, if we apply the Walsh-Hadamard transform (use a Hadamard gate), then the qubits will end up in the following two states:

$$\mid \psi'_Q\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \quad \text{and, respectively,} \quad \mid \psi'_R\rangle = \mid 0\rangle.$$

If we negate the two qubits (use the `X` gate), then the two qubits end up in the states:

$$\mid \psi''_Q\rangle = \mid 1\rangle \quad \text{and, respectively,} \quad \mid \psi''_R\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}.$$

In both instances we gain nothing, we are faced with either the initial problem, or we have to distinguish between states $\mid 1\rangle$ and $1/\sqrt{2}(\mid 0\rangle + \mid 1\rangle)$. Let us now consider a different approach: we prepare an auxiliary (an ancilla) qubit, $A$, in a special state, $\mid \psi_A\rangle = \mid 0\rangle$, then we perform a unitary transformation, $U$, of two systems, each one consisting of the ancilla and one of the original qubits. Finally, we measure the resulting system. Let us assume that the results of the unitary transformations are:

$$\mid \psi_{AQ}\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle \quad \text{and} \quad \mid \psi_{AR}\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \mid 10\rangle$$

with

$$\mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 = 1.$$

A measurement could yield the result 00, 01, or 10. If the result is 01 we are certain that we have selected $Q$. If the result is 10 we are certain that we have selected $R$. If the result is 00 we cannot be sure. So the probability of identifying correctly the qubit we have selected is $\mid \alpha_1 \mid^2$.

Now we have to construct a transformation $\mathbf{U}$ with the following properties:

$$\mathbf{U}[\mid \psi_A\rangle \otimes \mid \psi_Q\rangle] = \mid \psi_{AQ}\rangle \quad \mapsto \quad \mathbf{U} \mid 00\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle,$$

$$\mathbf{U}[\mid \psi_A\rangle \otimes \mid \psi_R\rangle] = \mid \psi_{AR}\rangle \quad \mapsto \quad \mathbf{U}\frac{\mid 00\rangle + \mid 01\rangle}{\sqrt{2}} = \alpha_0 \mid 00\rangle + \alpha_1 \mid 10\rangle,$$

$$\mathbf{U}^\dagger \mathbf{U} = I.$$

Let us consider a transformation $\mathbf{U}$ of the form

$$\mathbf{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{pmatrix}.$$

The first condition can be rewritten as

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ 0 \\ 0 \end{pmatrix}.$$

It follows that

$$u_{11} = \alpha_0, \qquad u_{21} = \alpha_1, \qquad u_{31} = 0, \qquad u_{41} = 0.$$

The second condition can be rewritten as

$$\begin{pmatrix} \alpha_0 & u_{12} & u_{13} & u_{14} \\ \alpha_1 & u_{22} & u_{23} & u_{24} \\ 0 & u_{32} & u_{33} & u_{34} \\ 0 & u_{42} & u_{43} & u_{44} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ \alpha_1 \\ 0 \end{pmatrix}.$$

It follows that

$$u_{12} = \alpha_0(\sqrt{2}-1), \qquad u_{22} = -\alpha_1, \qquad u_{32} = \alpha_1\sqrt{2}, \qquad u_{42} = 0.$$

From the third condition we find that

$$\alpha_0^2 = \frac{1}{\sqrt{2}}, \quad \alpha_1^2 = 1 - \frac{1}{\sqrt{2}},$$

$$u_{13} = -\sqrt{2}\alpha_1^2, \quad u_{43} = u_{14} = u_{24} = u_{34} = 0, \quad u_{44} = 1, \quad u_{23} = u_{33} = \frac{\alpha_1}{\alpha_0}.$$

Thus,

$$\mathbf{U} = \begin{pmatrix} \alpha_0 & \alpha_0(\sqrt{2}-1) & -\sqrt{2}\alpha_1^2 & 0 \\ \alpha_1 & -\alpha_1 & \alpha_1/\alpha_0 & 0 \\ 0 & \sqrt{2}\alpha_1 & \alpha_1/\alpha_0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In follows that when we use an ancilla qubit the probability to guess correctly increases from 0.25 to $\alpha_1^2 = 1 - 1/\sqrt{2} = 0.2929$. An interesting question left as an exercise is whether we can further increase this probability by using more than one ancilla qubits.

In fact we have hinted to a general principle used to measure the state of a quantum system: we let the system of interest interact with an ancilla and then we perform a von Neumann measurement on the ancilla. If $\rho_Q$ is the density operator of the original system and $\rho_A$ is the density operator of the ancilla system then, the composite system has a density operator which is the tensor product of the two density matrices

$$\rho_{QA} = \rho_Q \otimes \rho_A.$$

The interaction of the two systems is unitary

$$\rho_{QA} \mapsto \mathbf{U}\rho_{QA}\mathbf{U}^\dagger.$$

A measurement performed on the ancilla produces a result $q$ with probability

$$\mathrm{Prob}(q) = \mathrm{tr}((I \otimes M_q)\mathbf{U}\rho_{QA}\mathbf{U}^\dagger)$$

with $M_q$ a set of orthogonal projector operators acting on the ancilla qubit.

Next we discuss distinguishability of quantum states and quantum key distribution protocols.

## 2.24 Measurements and Distinguishability of Quantum States

Given a set of measurement operators $\mathbf{M}_i$ we can find a quantum measurement able to identify the state in which the quantum system has been prepared when the states $\{| \psi_i\rangle\}$, are orthonormal; for each possible state (each possible index $i$) we can define a measurement operator

$$\mathbf{M}_i \equiv | \psi_i\rangle\langle\psi_i |$$

such that the probability of outcome $i$ of the measurement when the system prepared in state $| \psi_i\rangle$ is equal to 1

$$p(i) = \langle\psi_i \mid \mathbf{M}_i \mid \psi_i\rangle \equiv 1.$$

If as a result of a measurement of a quantum system prepared in an unknown orthonormal state, the outcome $i$ occurs with certainty, this state is reliably distinguished as being $| \psi_i\rangle$. We cannot reliably distinguish non-orthogonal quantum states; a rigorous proof of this statement can be found in [308]. This property of quantum measurements poses serious challenges for quantum computing, but, at the same time, makes quantum information very attractive for applications such as quantum cryptography and quantum key distribution.

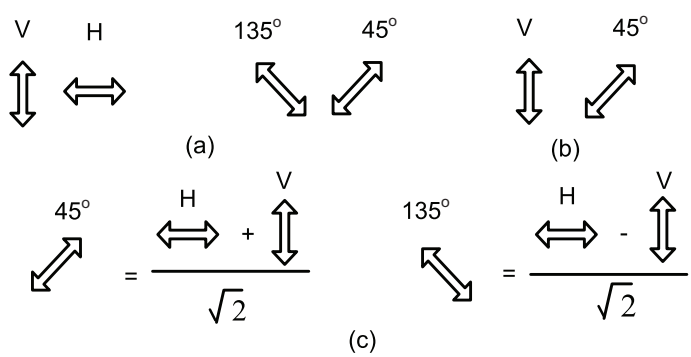



Figure 64: Distinguishability of quantum states. (a) We can reliably distinguish between orthogonal states, e.g., photons with vertical (V) versus horizontal (H) polarization, or photons with 45° from those with 135° polarization. (b) We cannot reliably distinguish between non-orthogonal states, e.g., photons with vertical from those with 45° polarization. (c) We can reliably distinguish between orthogonal superposition states.

Figure 64 illustrates this fundamental property of quantum systems for polarized photons. For example, we cannot reliably distinguish if a qubit is in state $| \psi_1\rangle =| 0\rangle$, or in state $| \psi_2\rangle = (| 0\rangle+ | 1\rangle)/\sqrt{2}$ through a measurement in the computational basis formed by the orthonormal vectors $| 0\rangle$ and $| 1\rangle$. Recall that when measuring the system in state $| \psi_2\rangle$ we get the result $| 0\rangle$ with probability 1/2 or $| 1\rangle$ with probability 1/2. When we measure the system in state $| \psi_1\rangle$ we get the result $| 0\rangle$ with probability 1. Henceforth, when we get the result corresponding to $| 1\rangle$ it is clear that the system must have been in state $| \psi_2\rangle$; yet, when we get the result corresponding to $| 0\rangle$, the system could have been originally either in the state $| \psi_1\rangle$ or in the state $| \psi_2\rangle$. If the states $| \psi_i\rangle$ of a quantum system are not orthonormal

we cannot find a quantum measurement able to distinguish the prepared state from other possible ones. Consider a set of measurement operators $\{\mathbf{M}_i\}$ and assume that we have a qubit that could be in one of two non-orthogonal states, $\mid \psi_1\rangle$ and $\mid \psi_2\rangle$. One of the two states, say $\mid \psi_2\rangle$, can be decomposed into two components, one (non-zero) parallel to $\mid \psi_1\rangle$, and the other, $\mid \psi_1^{\perp}\rangle$, orthogonal to $\mid \psi_1\rangle$

$$\mid \psi_2\rangle = \alpha \mid \psi_1\rangle + \beta \mid \psi_1^{\perp}\rangle.$$

We apply the measurement operator $\mathbf{M}_i$ and based on the result of the measurement decide that the system was in state $\mid \psi_1\rangle$. But, the previous equation shows that even when the system is in state $\mid \psi_2\rangle$ there is a non-zero probability, $\mid \alpha \mid^2$ of deciding that the system was in state $\mid \psi_1\rangle$. Therefore, the measurement does not reveal with certainty the state of the system prior to the measurement. The concepts discussed in this section have practical applications.

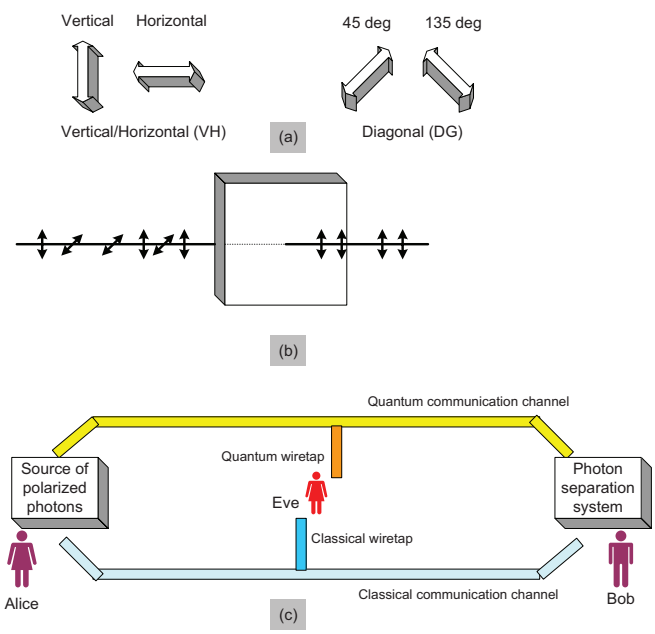



Figure 65: (a) Protocols for quantum key distribution are based on the idea of transmitting non-orthogonal qubit states. e.g., photons polarized in the $X$ and $Z$ bases. (b) A measurement in the wrong base alters the state. (c) The setup for quantum key distribution; Alice and Bob use quantum and classical communication channels and Eve eavesdrops on both.

We conclude the discussion of distinguishability of quantum states an applications, quantum key distribution. To ensure confidentiality, data, or *plaintext*, is encrypted into *ciphertext*. A *cipher* consists of algorithms transforming plaintext into ciphertext and back; a *symmetric cipher* uses the same key for encryption and decryption. Detecting eavesdropping is a major concern for key distribution protocols. Eavesdropping must be distinguished from the noise on the communication channel; to do so, "check" qubits must be interspaced randomly among the "data" qubits used to construct the encryption key. The key distribution protocols for

symmetric cipher cannot guarantee that a classical system is able to detect intrusion, while quantum information theory offers an ingenious solution to the problem.

**Quantum key distribution, QKD.** Protocols for quantum key distribution are based on the idea of transmitting non-orthogonal qubit states and then checking for the disturbance in their transmitted states, as shown in Figure 65(a) and Figure 65(b), respectively. The attempt of an intruder to eavesdrop increases the level of disturbance of the signal on the quantum communication channel, Figure 65(c). The two parties wishing to communicate in a secure manner establish an upper bound for the level of disturbance tolerable. They use a set of check bits to estimate the level of noise, or eavesdropping. Then they reconcile their information and distill a shared secret key.

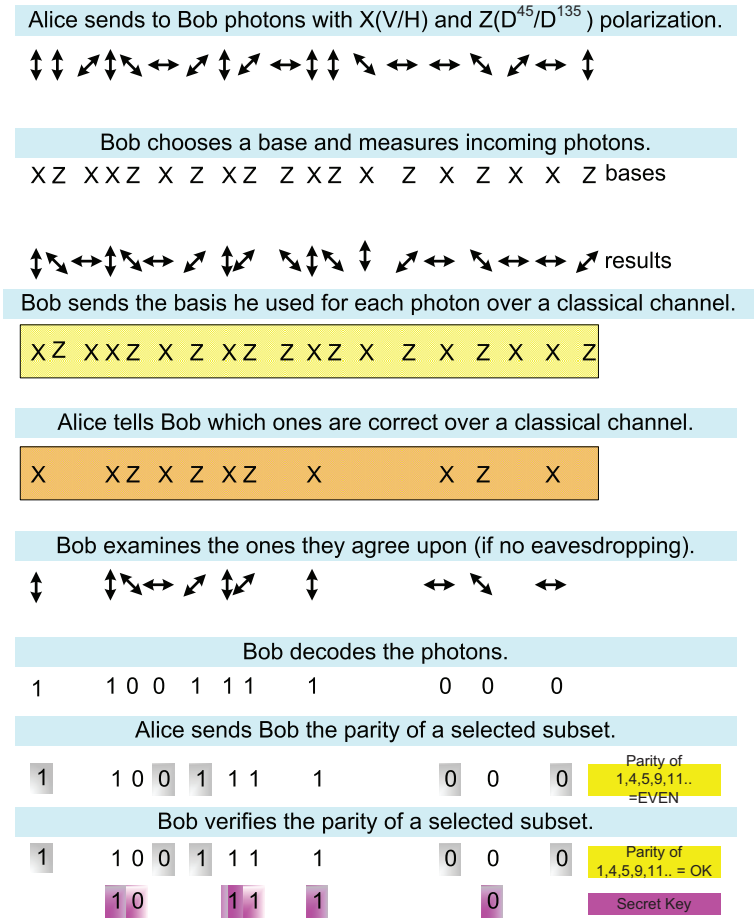



Figure 66: The main steps of the BB84 QKD protocol.

The first QKD protocol, *BB84*, was proposed by Bennett and Brassard in 1984 [32]; a proof of the security of this protocol can be found in [385]. Informally, the protocol works as follows: Alice sends $n$ qubits encoded randomly in one of the two bases she and Bob have initially decided on. Figure 65 illustrates the case when Alice wants to send the encryption key to Bob while Eve attempts to eavesdrop. Bob randomly chooses one of the two bases to

decode each received qubit. Bob ends up guessing correctly the bases for about $n/2$ of the qubits. Then Alice and Bob exchange over the classical communication channel information about the base used by each one of them to encode, respectively, decode each qubit. Alice and Bob know precisely on which qubits they have agreed and use them as the encryption key. When Eve eavesdrops and intercepts each of the $n$ qubits she randomly picks up one of the two bases to measure each qubit and then she re-sends the qubit to Bob. Approximately $n/2$ of the qubits received by Bob have their state altered by Eve. Bob guesses the correct bases for only about half of the qubits whose state has not been altered by Eve. Alice and Bob exchange information over the classical communication channel and detect the presence of an intruder when they agree on considerably fewer than $n/2$ qubits.

A formal description of the protocol can be found in [32] and [308]; its main steps are:

1. Alice selects $n$, the approximate length of the desired encryption key. Then she generates two random strings of bits $a$ and $b$ of length $(4n + \delta)$.

2. Alice encodes the bits in string $a$ using the bits in string $b$ to chose the basis ( either $X$ or $Z$) for each qubit in $a$. She generates $\mid \psi\rangle$, a block of $(4n + \delta)$ qubits

$$\mid \psi\rangle = \bigotimes_{k=1}^{4n+\delta} \mid \psi_{a_k b_k}\rangle$$

where $a_k$ and $b_k$ are the $k$-th bit of strings $a$ and $b$, respectively. Each qubit is in one of four pure states in two bases, $[\mid 0\rangle, \mid 1\rangle]$ and $[(1/\sqrt{2})(\mid 0\rangle + \mid 1\rangle),\ (1/\sqrt{2})(\mid 0\rangle - \mid 1\rangle)]$.
3. If $\mathcal{E}$ describes the combined effect of the channel noise and Eve's interference, then the block of qubits received by Bob is $\mathcal{E}(\mid \psi\rangle\langle\psi \mid)$.

4. Bob constructs a random string of bits, $b'$, of length $(4n + \delta)$. Bob then measures every qubit he receives either in basis $X$, or in basis $Z$, depending on the value of the corresponding bit of $b'$. As a result of his measurement Bob constructs the binary string $a'$.

5. From this instance on, all communications between Alice and Bob occur over the classical communication channel. First, Bob informs Alice that he now expects information about $b$. Alice discloses $b$.

6. Alice and Bob keep only the bits in the set $\{a, a'\}$ for which the corresponding bits of the strings $b$ and $b'$ are equal. Let us assume that Alice and Bob keep only $2n$ bits. By choosing $\delta$ sufficiently large Alice and Bob can ensure that the number of bits kept is close to $2n$ with a very high probability.

7. Alice and Bob perform several tests to determine the level of noise and eavesdropping on the channel. The set of $2n$ bits is split into two subsets of $n$ bits each. One subset will be the *check* bits used to estimate the level of noise and eavesdropping, and the other consists of the *data* bits used for the quantum key. Alice selects $n$ check bits at random and sends the position of the selected bits to Bob. Then Alice and Bob compare the values of the check bits. If more than, say, $t$ bits disagree, then they abort and retry the protocol.

## 2.25 Measurements and an Axiomatic Quantum Theory

We conclude this chapter with a discussion of the role of the measurements in an axiomatic reformulation of quantum theory. Quantum theory is based on the five postulates discussed in Sections 1.4 - 1.7 and we could ask if the postulates are logical and consistent with one another. For example, how do we reconcile the fact that, according to these postulates, the state of a quantum system is subject to linear and reversible evolution transformations, while measurements are non-linear and irreversible transformations that collapse the quantum state? We now consider the question whether one can construct the quantum theory from a set of more logical axioms and what would be the role of measurements in the new context.

George Birkhoff and John von Neumann examined this question in their 1936 paper on the logic of Quantum Mechanics [57]. They start their analysis by reminding us that quantum theory shares the concept of "phase-space" with classical mechanics and with classical electrodynamics; each system $\mathcal{S}$ is at each instance associated with a point $p(t)$ in the phase-space $\Sigma$ and this point represents mathematically the state of the system at time $t$. The principle of "mathematical causation" is reflected by a "law of propagation" which allows us to determine the state $p(t)$ knowing the initial state $p(0)$. In classical mechanics a point in the phase-space $\Sigma$ corresponds to a choice of $n$ position and $n$ conjugate momentum coordinates and the law of propagation is Newton law of attraction. In electrodynamics $\Sigma$ is a function-space of infinitely many dimensions, e.g., specified by electromagnetic and electrostatic potentials, and the law of propagation is given by Maxwell equations; in quantum theory $\Sigma$ is also a function-space, the points correspond to wave-functions and the law of propagation is given by Schrödinger equation. The law of propagation can be viewed as causing a steady fluid motion in the phase-space; in classical dynamics this flow conserves volumes while in Quantum Mechanics it conserves distances because the equations are unitary.

Any mathematical model must be anchored in the physical reality and this requires a mechanism to obtain information about the physical objects. We assume that the outcomes $\sigma_1, \sigma_2, \ldots, \sigma_n$ are obtained after performing a series of measurements $\mu_1, \mu_2, \ldots, \mu_n$ on a quantum system $\mathcal{S}$ and the point $(\sigma_1, \sigma_2, \ldots \sigma_n)$ is located inside the "observation space". The formal connection between physical reality and the mathematical model in the theory proposed by Birkhoff and von Neumann is ensured by a plausible postulate, a propositional calculus of Quantum Mechanics which establishes a connection between subsets of the observation space and subsets of the phase-space of a system $\mathcal{S}$.

More recently Lucien Hardy proposed a set of "more reasonable axioms" [201]; he argues that a better axiomatic formulation of quantum theory is not only more aesthetically appealing, but also allows a deeper conceptual understanding of the theory.

The setup for this axiomatic derivation of quantum theory consists of three types of devices: those used to prepare a quantum state, those used to transform a state, and those used to measure the state and transform quantum to classical information. The quantum theory is in fact a probability theory different from classical probability; measurements allow us to characterize a quantum state by a set of probabilities. To perform a measurement we need a large number of systems that have undergone an identical preparation. The definitions of state, degrees of freedom, and dimension are:

*The state* associated with a particular preparation is a mathematical object that can be used to determine the outcomes of any measurement of the system subject to that specific

preparation.

*The number of degrees of freedom,* $K$, is the minimum number of probability measurements needed to determine the state (or the cardinality of the set of real numbers required to specify the state).

*The dimension,* $N$, is the maximum number of states that can be reliably distinguished from one another in a single-shot measurement.

The five axioms introduced in [201] are:

*1. Probabilities.* Relative frequencies (measured by taking the proportion of occurrences a particular outcome is observed) tend to the same value (called probability) for any case a measurement is performed on an ensemble of $n$ systems prepared by some given preparation in the limit when $n$ becomes infinite.

*2. Simplicity.* $K$ is determined by a function of $N$, $K = f(N)$ where, for a given $N$, $K$ takes the minimum value consistent with the axioms.

*3. Subspaces.* A system whose state is constrained to belong to an M-dimensional subspace (i.e. have support on only $M$ of a set of $N$ possible distinguishable states) behaves like a system of dimension M.

*4. Composite systems.* A composite system consisting of subspaces $A$ and $B$ satisfies:

$$N = N_A N_B \quad \text{and} \quad K = K_A K_B.$$

*5. Continuity.* There exists a continuous reversible transformation on a system between any two pure states of that system. The transformation is continuous in the sense that its probability is a continuous function.

Classical probability theory assumes that there are $N$ distinguishable states called "basis or pure states." The $N$ pure states are described by the vectors:

$$p_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots p_{N-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad p_N = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

The state of the system is described by a vector:

$$p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_N \end{pmatrix}$$

with $p_i$ the probability of the system of being in the pure state $i$. The state $p$ belongs to a convex set[24], in other words the general state can be described as the convex sum of the pure states and the *null state*.

From this brief discussion it follows that the state in a classical probability theory can be determined by measuring $N$ probabilities; thus, $K = N$, *the number of degrees of freedom is equal to the dimension of the space.*

In quantum theory a state is represented by the density matrix, a self-adjoint operator $\rho$ with the property that $0 \leq \text{tr}(\rho) \leq 1$. Measurements are represented also by self-adjoint operators, the projectors $\mathbf{P}_k$ with the property that

$$\sum_k \mathbf{P}_k = I.$$

The probability of the outcome $k$ when a measurement $P_k$ is performed on a system in state $\rho$ is

$$p(k) = \text{tr}(\mathbf{P}_k \rho).$$

One can recast quantum theory in a similar mold; if we consider an $N$-dimensional Hilbert space, there are $N^2$ independent projectors, $P_k,\ 1 \leq k \leq N^2$ that span the space. Indeed, a general self-adjoint operator can be represented as a matrix with $N$ real numbers along the diagonal and $\binom{N}{2} = (N^2 - N)/2$ complex numbers above the diagonal; thus a self-adjoint operator is specified by a total on $N^2$ real numbers. In this case $K = N^2$, *the number of degrees of freedom is equal to the square of the dimension of the space.* Each projector corresponds to one degree of freedom.

While the first four axioms are consistent with classical probability the fifth is not. For example, when $N = 2$ the four projectors we choose to span the space of self-adjoint operators could be

$$\begin{array}{ll} P_0 = \mid 0\rangle\langle 0 \mid & P_2 = (\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle)(\alpha_0^*\langle 0 \mid + \alpha_1^*\langle 1 \mid) \\ P_1 = \mid 1\rangle\langle 1 \mid & P_3 = (\beta_0 \mid 0\rangle + \beta_1 \mid 1\rangle)(\beta_0^*\langle 0 \mid + \beta_1^*\langle 1 \mid) \end{array}$$

with $\mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 = 1$ and $\mid \beta_0 \mid^2 + \mid \beta_1 \mid^2 = 1$.

## 2.26 History Notes

Philosophers have been concerned with the relationship between human perception and "reality" for centuries. Rationalism, is a theory of knowledge based on the idea that the criterion of truth is not sensory but intellectual and deductive. The origins of this philosophy, whose main credo is self-sufficiency of reason, can be traced back to Pythagoras and Plato and was later embraced by Descartes, Spinoza, and Leibnitz who advocated the introduction of mathematical methods in philosophy. Spinoza and Leibnitz believed that in principle, all knowledge, including scientific knowledge, could be gained through the use of reason alone.

Rationalism is often contrasted to Empiricism, a theory of knowledge whose main tenet is that scientific knowledge is related to experience and that experiments are essential for

[24] A convex set $S$ in a vector space over the field $\mathbb{R}$ is one when the line segment joining any two points in $S$ lies entirely in $S$.

the development of scientific concepts, all hypotheses and theories must be tested against observations of the natural world. The Sophists in the $V$-th century BC, Aristotle (384 BC - 322 BC), the Stoics and the Epicureans, a generation later, followed by Thomas Aquinas and Roger Bacon in the XIII century, and John Locke, George Berkeley, David Hume, in the XVII and XVIII century, attribute the sensory perception and the a posteriori observation a major role in the formation of human knowledge. They believed that human knowledge of the natural world is grounded in senses' experience. The work of Immanuel Kant bridges the gap between Rationalism and Empiricism. The author of the "Critique of Pure Reason" expressed the belief that we are never able to transcend the bounds of our own mind, we cannot access the "Ding an sich," in German, the "thing-in-itself."

As expected from this brief discussion of Rationalism and Empiricism the concept of "measurement" proved to be intensely controversial from the early days of Quantum Mechanics. The question whether a numerical value of a physical quantity has any meaning until an observation has been performed was answered differently by the representatives of the metaphysical and realist schools of thought. The argument of Niels Bohr was "If a quantity $Q$ is measured in system $S$ at time $t$, then $Q$ has a particular value in $S$ at time $t$," while Werner Heisenberg expressed his positivistic views as: "It is meaningless to assign $Q$ a value $q$ for $S$ at time $t$, unless $Q$ is measured to value $q$ at time $t$."

Bohr argued that the later interpretation reduced the question of definability to measurability, but he and Heisenberg agreed in broad terms that the experimental conditions play an important role on the description of quantum systems, an important tenet of the Copenhagen Interpretation. There are some similarities between the position taken by Bohr and that of Kant in Transcendental Aesthetic: "..not only are drops of rain mere appearances, but...even the space in which they fall, are nothing in themselves but modifications of our sensible intuition...the transcendental objects remain unknown to us..."

Einstein, Podolsky, and Rosen define the concept of physical reality (see Section 2.16) as objective, independent of any measurement; a measurement reflects passively physical reality. Einstein empiricist position has its roots in the Cartesian and Lockean notion of perception.

John von Neumann provided an axiomatic definition of a quantum measurement process consisting of a linear phase when the behavior of the system is described by the Schrödinger equation followed by a non-linear stage when the state of the system is collapsed. The so called "measurement problem" centers on the apparent conflict between two of the postulates of Quantum Mechanics, the measurements postulate and the dynamics postulate, "...the postulate of collapse seems to be right about what happens when we make a measurement, and the dynamics postulate seems to be bizarrely wrong about what happens when we make measurements, and yet the dynamics postulate seems to be right about what happens when we aren't making measurements..." [10]. Sometimes the measurements postulate is called the postulate of collapse; the dynamics of a quantum system is specified by a self-adjoint operator $\mathbf{H}$, called the Hamiltonian, and the time-evolution of the system is described with the Schrödinger equation, according to the dynamics postulate discussed in Section 1.6.

## 2.27 Summary and Further Readings

Any scientific theory of physical phenomena must be anchored in reality and this requires the means to gather information about the physical objects. Measurements are necessary to assess the consistency of the mathematical model with the physical reality. A mathematical model allows us to define the "state" of a physical system as a point in an abstract "phase-space," while measurements allow us to define the state of the physical system in an "observation space;" the consistency of the mathematical model with the physical reality requires a logical connection between corresponding subspaces of the two spaces.

Asher Peres writes [325]: "It should be clear that the interpretation of raw experimental data always necessitates the use of some theory. Concepts such as *angular momentum* are part of the theory, not of the experiment. Moreover *correspondence rules* are needed to relate the abstract notions of the theory - together with its correspondence rules - which tell us what can, or cannot, be answered. What does not exist in theory cannot be observed in any experiment to be described by that theory. Conversely, anything described by the theory is deemed to be observable, unless the theory itself prohibits to observe it. To be acceptable a theory must have predictive power about the outcomes of the experiments that it describes, so that the theory can actually *fail*. A *good* theory is one that does not fail in its domain of applicability. Today, in our present state of knowledge, quantum theory is the best available one for describing atomic, nuclear, and many other phenomena."

The concept of "measurement" and the relationship between the outcome of the measurement and our knowledge of the state of the quantum system prior, during, and after the measurement proved to be intensely controversial. The Copenhagen Interpretation of Quantum Mechanics expounded by Niels Bohr and Werner Heisenberg was contested by many physicists including Albert Einstein.

According to the measurements postulate, a quantum measurement *collapses the state* of a quantum system. A quantum measurement is an irreversible process which destroys quantum information and replaces it with classical one. To be reversible, a quantum measurement should not reveal any information about the quantum state.

In Section 1.14 we discussed the requirements for building a quantum computer formulated by David DiVincenzo [131]. One of the requirements states that, when we have a register of $n$ qubits and if a the density operator of a qubit is

$$\rho = p \mid 0\rangle\langle 0 \mid + (1-p) \mid 1\rangle\langle 1 \mid + \alpha \mid 0\rangle\langle 1 \mid + \alpha^* \mid 1\rangle\langle 0 \mid,$$

with $\alpha \in \mathbb{C}$, then an "ideal measurement" should provide an outcome of "0" with probability $p$ and an outcome of "1" with probability $(1-p)$, regardless of the value of $\alpha$, regardless of the state of the neighboring qubits, and without changing the state of the quantum computer. A "non-demolition" measurement also leaves the qubit in state $\mid 0\rangle\langle 0 \mid$ after reporting the outcome 0 and leaves it in state $\mid 1\rangle\langle 1 \mid$ after reporting the outcome 1.

John von Neumann treats the measurement apparatus as a quantum object and argues that a measurement consists both of (i) a linear stage, when the joint state of the quantum system and the measuring instrument is a tensor product of the states of the two systems; and (ii) a non-linear stage, when this joint state collapses to one of the tensor products of various eigenstates of the measurement operator and the state of the measuring instrument corresponding to the instantiated eigenvalue.

The outcome of a measurement of an observable $\mathcal{A}$ of a quantum system is an eigenvalue of the self-adjoint operator $\mathbf{A}$ associated with the measurement of the observable. The quantum state immediately after the measurement is an eigenvector of the operator $\mathbf{A}$. Operator $\mathbf{A}$ is a *projector* for observable $\mathcal{A}$ of a quantum system.

Projective measurements are carried by means of self-adjoint operators $\mathbf{M}_i$ such that $\mathbf{M}_i\mathbf{M}_j = \delta_{i,j}\mathbf{M}_i$. Projective measurements are restrictive and are not always possible. Sometimes we destroy a quantum particle in the process of measurement thus, the repeatability of a projective measurement is violated. Orthogonal measurement operators commute and correspond to simultaneous observables. Measurement operators corresponding to non-orthogonal states do not commute and are therefore not simultaneously observable. The POVM measurements are concerned only with the statistics of the measurement. POVM allows us to distinguish the cases when the outcome of a measurement identifies with certainty the original state from the ones where the identification is not possible.

The characterization of the state of a quantum system by means of the density matrix allows us to distinguish between "pure" and "mixed" states and to conclude that we can acquire maximal information only about pure states. When the state of a quantum system is known exactly, the system is in a *pure state*, $\mid \psi\rangle$. Then

$$\rho \equiv \mid \psi\rangle\langle\psi \mid .$$

The trace of the density operator of a pure state satisfies the equation:

$$\mathrm{tr}(\rho^2) = 1.$$

If the state of the system is not known exactly, the system is in a *mixed state*. The set $\mathcal{E} = \{p_i, \mid \psi_i\rangle\}$ is called an *ensemble of pure states*. Given the state $\mid \psi_i\rangle$ and the probability $p_i$ of the quantum system to find itself in state $\mid \psi_i\rangle$ the density operator is defined as

$$\rho \equiv \sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid = \sum_i p_i\rho_i.$$

The trace of the density operator of a mixed state satisfies the equation

$$\mathrm{tr}(\rho^2) < 1.$$

If a unitary operator $U$ is applied to a system which is in state $\mid \psi_i\rangle$ with probability $p_i$, then the resulting state is $U \mid \psi_i\rangle$ with probability $p_i$ and the density operator evolves as follows

$$\rho = \sum_i p_i \mid \psi_i\rangle\langle\psi_i \mid \qquad \Longrightarrow \qquad \rho^{afterU} = \sum_i p_i U \mid \psi_i\rangle\langle\psi_i \mid U^\dagger = U\rho U^\dagger.$$

Mixed states are used to describe statistical mixtures of pure states as well as composite systems. Let $\mathcal{C}$ be a bipartite system consisting of systems $\mathcal{A}$ and $\mathcal{B}$; the state vectors and the density operators of the three systems are respectively: $\mid \xi_\mathcal{C}\rangle$ and $\rho_\mathcal{C}$, $\mid \xi_\mathcal{A}\rangle$ and $\rho_\mathcal{A}$, $\mid \xi_\mathcal{B}\rangle$ and $\rho_\mathcal{B}$. The bipartite system can be either in a product state, or in an entangled state. $\mathcal{C}$ is in a product state if $\mid \xi_\mathcal{C}\rangle = \mid \xi_\mathcal{A}\rangle \otimes \mid \xi_\mathcal{B}\rangle$; in this case the reduced density operators satisfy the equalities $\rho^\mathcal{A} = \rho_\mathcal{A}$ and $\rho^\mathcal{B} = \rho_\mathcal{B}$.

One of the most intriguing properties of quantum information is the shareability of quantum correlations. While classical correlations can be shared among many parties, quantum

correlations cannot be shared freely, quantum correlations of pure states are monogamous. If two quantum systems $\mathcal{A}$ and $\mathcal{B}$ are in a maximally entangled pure state, then neither of them can be correlated with any other system in the universe. The monogamy of entanglement is the deeper root of our inability to clone quantum states.

The EPR gedanken experiment is concerned with the measurement of entangled particles and led some to believe in the existence of hidden variables. The hidden variable theory assumes that in an EPR experiment there is a more complete specification of the state of the two entangled particles than the one given by the wave function. John Bell proved the nonexistence of local hidden variables based on the observation that any hidden variable account of quantum mechanics must have the separability property and this is not possible. Bell inequality is based on two common sense assumptions:

1. *Locality* — measurements of different physical properties of different objects, carried out by different individuals, at distinct locations, cannot influence each other.

2. *Realism* — physical properties are independent of observations; the existence of physical properties is independent of them being observed.

Quantum systems in pure entangled states and quantum systems in some mixed entangled states violate Bell inequality. Quantum theory is consistent with the experimental evidence acquired along the years thus one of these two assumptions must be incorrect.

Bell showed that it is possible to mimic the statistical properties of quantum theory by a deterministic model based on hidden variables. He constructed such a deterministic model generating results whose averages are identical to those predicted by quantum mechanics. When the quantum system is made up of disjoint subsystems entangled like those in the EPR experiment, then the hidden variables cannot be separated as disjoint subsets, local to each subsystem. Bell inequality quantifies how measurements on quantum systems in entangled states can invalidate local classical models of reality.

John von Neumann discusses simultaneous measurability, measurements and reversibility, and the macroscopic measurements in Chapters 3 and 5 of the "Mathematical Foundations of Quantum Mechanics [442].

Part two, "Cryptodeterminism and Quantum Inseparability" of the wonderful book by Asher Peres [326] provides insights in quantum correlations, composite systems, and Bell theorem.

"Speakable and Unspeakable in Quantum Mechanics" is a collection of John Bell papers with an introduction by Alain Aspect [26] is recommended for its clear exposition of the EPR paradox, hidden variables, as well as Bell theorem and other issues related to the philosophical aspects of measurements.

Roger Penrose addresses the question "What is quantum reality?" and discusses the holistic nature of the wave function in Chapter 21 of his book, "The Road to Reality: A Complete Guide to the Laws of the Universe," and then reflects on quantum measurements in the chapter dedicated to quantum algebra [324].

The book "Science and Ultimate Reality" dedicated to the 90th birthday of John Archibald Wheeler [23] consists of a collection of articles written by Paul Davies, David Deutsch, Freeman Dyson, Lucien Hardy, Hideo Mabuchi, Anton Zeilinger, Wojcieck Zurek and other researchers in the field of quantum computing and quantum information theory. Some of the ideas from the paper by Lucien Hardy, "Why is Nature Described by Quantum Theory" [201] are presented in Section 2.25.

## 2.28 Exercises and Problems

**Problem 1.** The following statement appears on page 187 of reference [308]: "To be reversible a quantum measurement should not reveal any information about the quantum state." A measurement is expected to provide information about the system being measured; why should anyone carry out a measurement which does not reveal any information? Discuss how this statement should be understood. How is it possible to detect errors under these conditions?

**Problem 2.** Show that the trace of the product of the square matrix $A[a_{ij}]$ $1 \leq i, j \leq n$ and its complex conjugate $A^* = A[a^*_{ij}]$ $1 \leq i, j \leq n$ is nonnegative:

$$\mathrm{tr}(A^* A) \geq 0.$$

**Problem 3.** Prove that the set of $n \times n$ square matrices with elements from either $\mathbb{R}$ or $\mathbb{C}$ form a vector space with the inner product defined as

$$\langle A, B \rangle = \mathrm{tr}(A^* B).$$

**Problem 4.** Consider the Bell states $\beta_{01}, \beta_{10}$ and $\beta_{11}$.

(4.1) Compute the density operators $\rho_{01}, \rho_{10}$, and $\rho_{11}$ for each of the four states and show that each state is a pure state.

(4.2) Compute the reduced density operators, $\rho^{(1)}_{01}, \rho^{(1)}_{10}$, and $\rho^{(1)}_{11}$ and show that one qubit of the pair is in a mixed state.

(4.3) Consider two systems, $\mathcal{A}$ in a pure state $\mid \psi_{\mathcal{A}} \rangle$ and $\mathcal{B}$ in a pure state $\mid \psi_{\mathcal{B}} \rangle$. Let $\mathcal{C} = \mathcal{AB}$ be in state $\mid \psi_{\mathcal{C}} \rangle = \mid \psi_{\mathcal{C}} \rangle \otimes \mid \psi_{\mathcal{B}} \rangle$. Show that $\mid \psi_{\mathcal{C}} \rangle$ is a pure state.

**Problem 5.** Show that the density operator of an equal mixture of two orthogonal single qubit pure states is

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

**Problem 6.** Show that if the density operator of a qubit has only one eigenvalue then the state is a pure state; if the density operator has more than one eigenvalue the state is a mixed state.

**Problem 7.** Let $\mid a \rangle = \sum_i \alpha_i \mid i \rangle$ and $\mid b \rangle = \sum_i \beta_i \mid i \rangle$ be state vectors in $\mathcal{H}_n$. Show that

$$\mathrm{tr}\left( \mid a \rangle \langle b \mid \right) = \langle a \mid b \rangle.$$

**Problem 8.** Show that the states at different stages of the quantum circuit in Figure 63 (a) are:

$$\mid \xi_0 \rangle = \frac{1}{\sqrt{2}} [\alpha_0 \mid 0 \rangle (\mid 00 \rangle + \mid 11 \rangle) + \alpha_1 \mid 1 \rangle (\mid 00 \rangle + \mid 11 \rangle)]$$

$$| \xi_1\rangle = \frac{1}{\sqrt{2}}[\alpha_0 \mid 0\rangle(\mid 00\rangle + \mid 11\rangle) + \alpha_1 \mid 1\rangle(\mid 10\rangle + \mid 01\rangle)]$$

and

$$\begin{aligned} | \xi_2\rangle \quad &= \tfrac{1}{2}[ \quad \mid 00\rangle(\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) + \mid 01\rangle(\alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle) \\ &+ \quad \mid 10\rangle(\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle) + \mid 11\rangle(\alpha_0 \mid 1\rangle - \alpha_1 \mid 0\rangle)]. \end{aligned}$$

**Problem 9.** Show that using an ancilla system consisting of more than one qubit does not increase the probability of guessing right for the measurements discussed in Section 2.23.

"God exists, since mathematics is consistent, and the Devil exists, since we cannot prove it."
Attributed to Andre Weil by Paul Rosenbloom.

# 3 Classical and Quantum Information Theory

In the late 1940s Claude Shannon stated that: "the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point" [375]. The precise nature of the agents transmitting information to one another, as well as the signalling process, can be very different; this motivated Shannon to develop abstract models of communication and establish the foundation of information theory. Signalling is an overloaded term frequently used in communication; in this context it means the physical perturbation of the transmission media used to carry the information through the communication channel.

During the decades following the publication of Shannon's work came the realization that information theory has a broader scope, it plays an important role in understanding critical phenomena related not only to the transmission, but also to the storage and the transformation of information. Indeed, information can be transmitted from one place to another; information can also be transmitted from one moment in time to another, when it is either stored or transformed with a computing device.

Today we study not only classical information, but also quantum and biological information. The three flavors of information are anchored in physical reality, the information is embodied by some property of a physical system. Classical information is transmitted,

stored, and transformed using physical processes governed by the laws of Classical Physics, while quantum information is carried by systems governed by the laws of Quantum Mechanics; little is known about the properties of biological information, information manipulated at cell level in living organisms.

The properties of classical and quantum information are very different; classical information is carried by systems with a definite state, it can be replicated and measured without being altered. Quantum information is encoded as a property of quantum systems e.g., photon polarization or particle spin, and has special properties such as superposition and entanglement with no classical counterpart; quantum information cannot be cloned and it is altered as a result of a measurement.

Quantum information theory studies: (a) transmission of classical information over quantum channels; (b) transmission of quantum information over quantum channels; (c) the effect of quantum entanglement on information transmission; and (d) the informational aspect of the quantum measurement process, the tradeoffs between the disturbance of the quantum state and the accuracy of the measurement.

The roots of quantum information theory can be traced back to John von Neumann who in the early 1930s introduced critical concepts such as the von Neumann entropy [442]. In 1973 Alexander Holevo found an upper bound of the accessible information in a quantum measurement [211].

Quantum communication involves a source that supplies quantum systems in a given state, a quantum channel that "transports" the quantum system, and the recipient that receives and decodes the quantum information. A quantum channel, like a classical one, can transmit quantum information *from one place to another and from one time to another.* For example, a photon transmitted through an optical fiber could carry information from site $A$ to site $B$, while a storage device implemented as an ion trap could be regarded as a channel transmitting information in time. The information transmitted through a quantum channel can be affected by errors and the means to correct such errors are critical for any practical applications of quantum information. This view of a quantum channel give us an indication that quantum information theory and quantum error correcting codes are important not only for the transmission, but also for processing and storage of quantum information.

Several physical embodiments of quantum information communication model are possible: the source could be a laser producing monochromatic photons, the channel could be an optical fiber and the recipient a photocell; the source could also be an ion trap controlled by laser pulses, the channel an ion trap, and the receiver a photo detector reading out the state of the ions via laser-induced fluorescence [276].

The properties of quantum information analyzed in Chapters 1 and 2 have profound implications for the discussion in this chapter. We expect to reproduce with fidelity the input of a classical noiseless channel to its output; this is no longer true for quantum channels. A quantum channel cannot reproduce with outmost fidelity its input when the input consists of an ensemble of mixed non-orthogonal states; we can only distinguish orthogonal states thus, we expect that any measurement performed at the output of the channel will alter the states.

Rather than presenting classical and quantum information separately we interleave the discussion of them. Concepts we are familiar with in classical information theory, such as joint and conditional entropy, and mutual information, have a quantum counterpart; Fano's inequality and the data processing inequality have a quantum analog; von Neumann as well

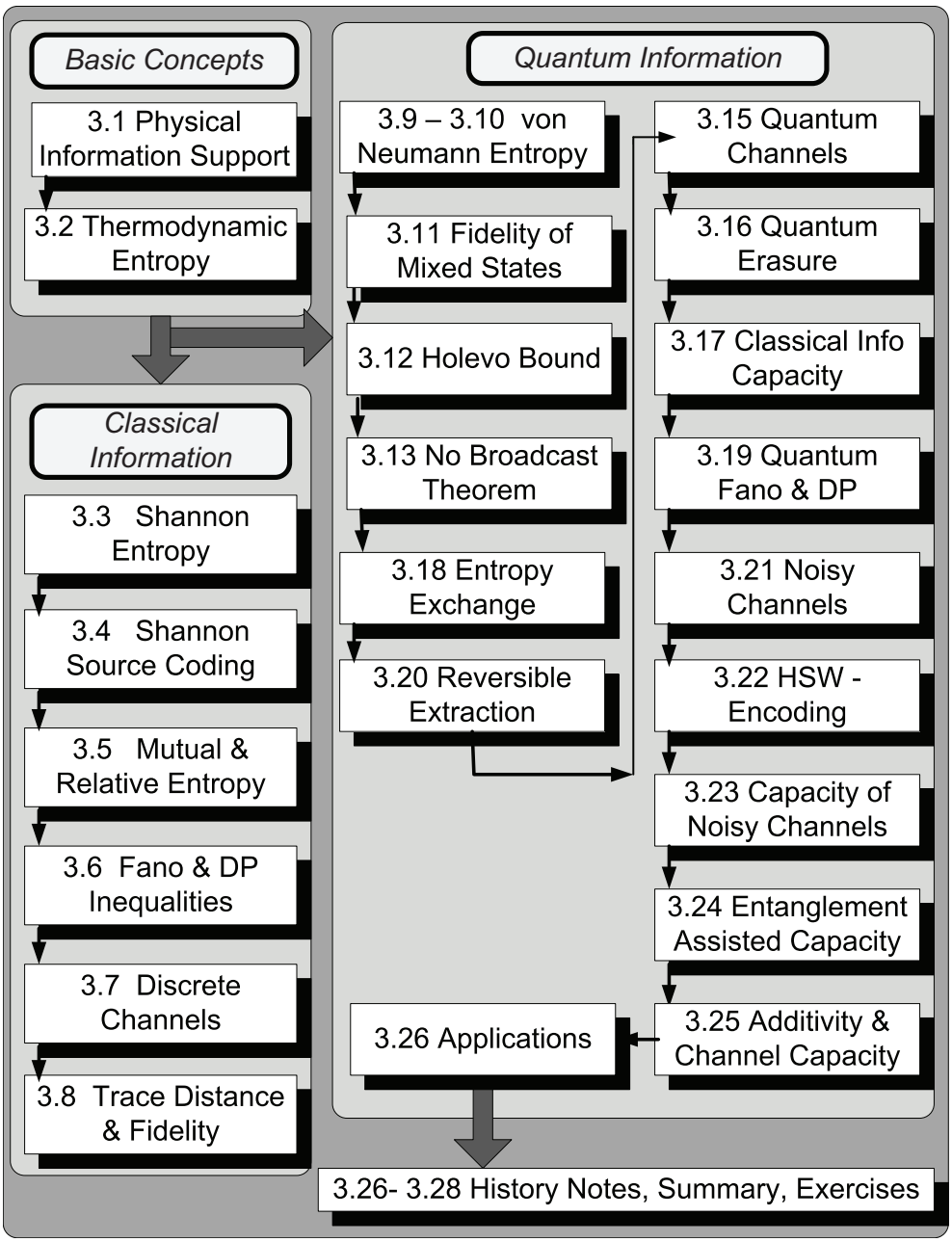



Figure 67: The organization of Chapter 3 at a glance.

as Shannon entropy enjoy strong subadditivity, and the list could go on and on.

The organization of the chapter is summarized in Figure 67. We first discuss the physical support of information then we overview several properties of classical information based on Shannon's theory; then we concentrate on the properties of quantum information and discuss the sources and the quantum channels.

## 3.1 The Physical Support of Information

Information must have a physical support. Indeed, to transmit, store, and process information we must act on a property of the physical medium embodying the information; to extract information at the output of a channel we have to perform a measurement of the physical media. It seems unlikely that one would dispute the statement that information is physical; nevertheless, the profound implications of this truth, the connection between the fundamental laws of physics and information have escaped the scientific community for a very long time [330, 335].

In 1929, Leó Szilárd stipulated that information is physical while trying to explain Maxwell's daemon paradox [415]. In 1961, Rolf Landauer uncovered the connection between information and the second law of thermodynamics [265]. Landauer followed a very simple argument to show that classical information can be copied reversibly, but the erasure of information is an irreversible process. We accept easily, and possibly painfully, the fact that erasure of information is irreversible; think about the information you lost last time you accidentally erased an important file from your storage media!

To quantify the amount of energy required to erase one bit of information we have to remember a basic attribute of classical information, its independence of the physical support; a bit can be stored as the presence or absence of a single pebble in a square drawn on the sand, a flower pot or the absence of it on the balcony of Juliet, or a molecule of gas in a cylinder.

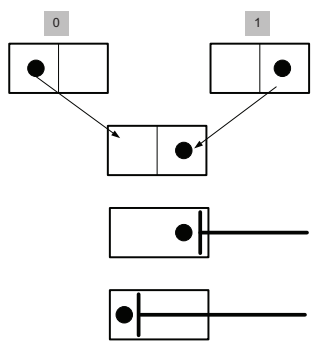



Figure 68: A gedanken experiment to quantify the amount of energy required to erase a bit of information. We encode one bit of information as the presence of one molecule of gas in a cylinder with two compartments. A "0" corresponds to the molecule in the left compartment and a "1" to the molecule in the right compartment (top). To erase the information we remove the wall separating the two compartments and insert a piston on the right side of the cylinder, push the piston and compress the gas. The amount of energy required to erase the information is equal to $k_B T \ln 2$.

We encode one bit of information as the presence of one molecule of gas in a cylinder with two compartments. A "0" corresponds to the molecule on the left compartment and a "1" to the molecule on the right compartment, Figure 68. To erase the information we proceed as follows: (a) we remove the wall separating the two compartments and insert a piston on the right side of the cylinder; (b) we push the piston and compress the one-molecule gas. At the

end of step (b) we have "erased" the information, we no longer know in which compartment the molecule was initially.

If the compression is isothermal (the temperature does not change) and quasi-static (the state changes very slowly), then the laws of thermodynamics tell us that the energy required to compress $m$ molecules of gas is equal to $mk_BT\ln 2$. In our experiment $m = 1$, thus the total amount of energy dissipated by us in erasing the information is $k_BT\ln 2$, equal to the amount of heat dumped into the environment.

Landauer's principle can be formulated as follows: *to erase one bit of information we have to dissipate an amount of energy of at least* $k_BT\ln 2$ with $k_B$ the Boltzmann's constant, $k_B = 1.3807 \times 10^{-23}$ Joules per degree Kelvin, and $T$ the temperature of the system.

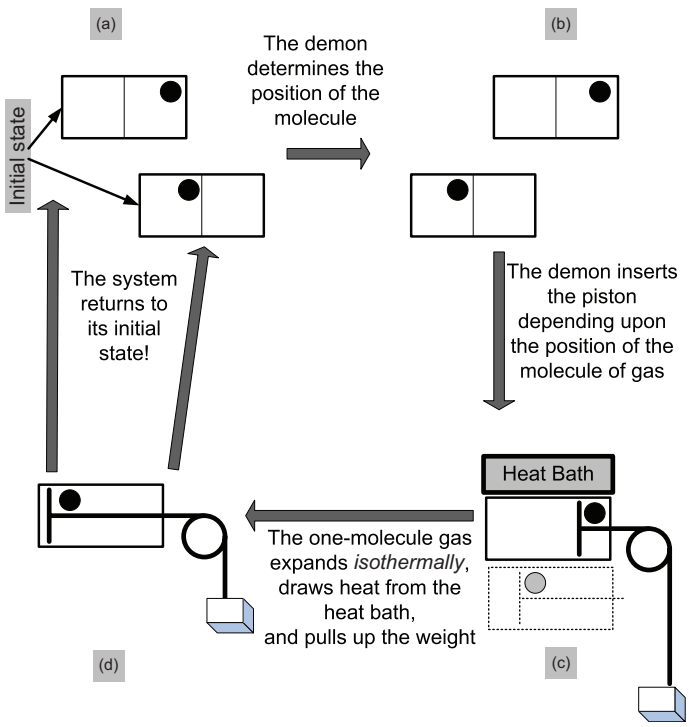



Figure 69: A cycle of Szilárd's engine. (a) One molecule of gas is located either in the left half, or in the right half of a cylinder with a partition in the middle. (b) The demon measures the position of the molecule. (c) The demon removes the partition and, based on the information regarding the position of the particle, inserts a piston in the cylinder to the left of the molecule. (d) The gas expands isothermally and lifts the weight. During the expansion the one-molecule gas is put in contact with a reservoir of heat and draws heat from the reservoir. Finally, the system returns to its initial state and a new cycle begins. The cylinder with the one-molecule gas could be returned to the initial state if and only if the demon provides the information regarding the initial position of the molecule.

**Maxwell's demon.** Maxwell imagined a gedanken experiment involving the famous demon as a challenge to the second law of thermodynamics discussed in Section 3.2. Maxwell's demon is a mischievous fictional character who controls a trapdoor between two containers, $A$ and $B$, filled with the same gas at equal temperatures; the containers are placed next to each

other, $B$ on the left of $A$. When a faster-than-average molecule from container $A$ approaches the trapdoor, the demon opens it, and allows the molecule to fly from container $A$ to $B$; when a slower-than-average molecule from $B$ approaches the trapdoor, the demon opens it, and allows the molecule to fly from $B$ to $A$. The average molecular speed corresponds to temperature, therefore the temperature will decrease in $A$ and increase in $B$, in violation of the second law of thermodynamics. If you are about to invent a perpetuum mobile based on this gedanken experiment read further before trying to patent you invention; if you already know what is wrong with this argument, you may skip the remaining of this section.

**The thermodynamic engine.** One of the most famous demonstrations that the second law of thermodynamics is not violated by Maxwell's demon was suggested by Leó Szilárd who considered the total entropy of the gas and demon combined. Szilárd imagined an engine powered by this demon. As before, we have one molecule of gas located either in the left half, or in the right half of a cylinder with a partition in the middle. The demon measures the position of the molecule, then removes the partition and, based on the information regarding the position of the particle, the demon inserts a piston in the cylinder to the left of the molecule. On the right of the piston, an arm is connected to a weight hanging from a pulley. As the gas expands, the piston moves to the left and pulls up the weight. During the expansion the one-molecule gas is put in contact with a reservoir of heat, draws heat from the reservoir, expands isothermally (the temperature $T$ is maintained throughout the experiment), and lifts the weight. The work extracted by the engine in the isothermal expansion is

$$W_{isothermalExpansion} = +k_B T \ln 2.$$

Then the cycle repeats itself and the Szilárd engine keeps on converting the heat generated when the gas expands isothermally into work, a process forbidden by the second law of thermodynamics.

The point one could easily miss is that at the beginning of each cycle the ensemble consisting of the cylinder with its auxiliary components, *as well as the demon,* should be returned to the initial state. Figure 69(d) shows that the gas cylinder with the one-molecule gas could be returned to the initial state *if and only if the demon provides the information regarding the initial position of the molecule.*

The demon is an essential component of the engine, the contraption could not function without the demon. The demon's memory stores one bit of information acquired during the first step of the process: "0" if the molecule is located on the left partition and "1" if the molecule is on the right partition. This information should be erased at the beginning of each cycle; according to Landauer's principle the energy dissipated in erasing one bit of information is

$$W_{erasure} = -k_B T \ln 2.$$

Thus, the work gained by the engine is required to erase the demon's memory; the heat generated by the erasure of demon's memory is transferred back to the heat bath to compensate for the heat absorbed during the expansion of the one-molecule gas. No net gain is registered during a cycle of this hypothetical engine since

$$W_{cycle} = W_{isothermalExpansion} + W_{erasure} = +k_B T \ln 2 - k_B T \ln 2 = 0$$

and the second law of thermodynamics is not violated.

We introduce in the next section the concept of thermodynamic entropy, a measure of the amount of chaos in a microscopic system. The information available to the demon is a part of the system state and it is related to the entropy. One could consider a generalized definition of entropy $\mathcal{I}$ (in bits), as the difference between the thermodynamic entropy $\Delta S$ and the information $I$ about the system available to an external observer (in our case the demon) [335]

$$\mathcal{I} = \Delta S - I.$$

The energy per cycle can be expressed in terms of $\Delta S$, the difference between the entropy at the beginning and the end of one cycle, and the change in the information available to an external observer at the same instances, at temperature $T$ as

$$W_{cycle} = W_{isothermalExpansion} + W_{erasure} = T(\Delta S - I) = T\mathcal{I}.$$

It follows that

$$W_{cycle} = 0 \quad \Leftrightarrow \quad \mathcal{I} = 0.$$

Indeed, information is physical and it contributes to define the state of a physical system.

## 3.2 Entropy; Thermodynamic Entropy

Entropy is one of the most important concepts in Physics and in information theory. Informally, the entropy is a measure of the amount of disorder in a physical, or a biological system. The higher the entropy of a system, the less information we have about the system. *Henceforth, information is a form of negative entropy.*

We discuss first the thermodynamic entropy introduced by Ludwig Boltzmann in the 1870s; the thermodynamic entropy is proportional to the logarithm of the number of microstates of the system. After half a century, in the late 1920s, John von Neumann wanted to quantify the entropy of a quantum system in a mixed state and introduced his own definition of entropy. The von Neumann entropy is related to the number of states of a quantum system. Several decades later, in the late 1940s, Claude Shannon formulated the entropy in the context of classical information theory.

According to Alfred Wehrl [446] "entropy relates macroscopic and microscopic aspects of nature and determines the behavior of macroscopic systems, i.e., real matter, in equilibrium (or close to equilibrium)". The entropy is a measure of the amount of chaos in a microscopic system. The concept of entropy was first introduced in thermodynamics. Thermodynamics is the study of energy, its ability to carry out work, and the conversion between various forms of energy, such as the internal energy of a system, heat, and work. The laws of thermodynamic are derived from statistical mechanics.

There are several equivalent formulations of each of the three laws of thermodynamics. The *First Law of Thermodynamics* states that energy can be neither created, nor destroyed, it can only be transformed from one form to another. An equivalent formulation is that the

heat flowing into a system equals the sum of the change of the internal energy and the work done by the system.

The *Second Law of Thermodynamics* states that it is impossible to create a process that has the unique effect of subtracting positive heat from a reservoir and converting it into positive work. An equivalent formulation is that the entropy of a closed system never decreases, whatever the processes that occur in the system: $\Delta S \geq 0$, where $\Delta S = 0$ refers to reversible processes and $\Delta S > 0$ refers to irreversible ones. A consequence of this law is that no heat engine can have 100% efficiency.

The *Third Law of Thermodynamics* states that the entropy of a system at zero absolute temperature is a well-defined constant. This is due to the fact that many systems at zero temperature are in their ground states and the entropy is determined by the degeneracy of the ground state. For example, a crystal lattice has a unique ground state and has zero entropy when the temperature reaches 0° Kelvin.

Clausius defined *entropy* as a measure of energy unavailable for doing useful work. He discovered the fact that entropy can never decrease in a physical process and can only remain constant in a reversible process. This result became known as the Second Law of Thermodynamics. There is a strong belief among astrophysicists and cosmologists that our Universe started in a state of perfect order and its entropy is steadily increasing, leading some to believe in the possibility of a "heat death, " a state when there is no free thermodynamic energy to sustain life.

The thermodynamic entropy of a gas, $S$, is also defined statistically but it does not reflect a macroscopic property. *The entropy quantifies the notion that a gas is a statistical ensemble and it measures the randomness, or the degree of disorder, of the ensemble.* The entropy is larger when the vectors describing the individual movements of the molecules of gas are in a higher state of disorder than when all molecules are well organized and moving in the same direction with the same speed. Ludwig Boltzmann postulated that

$$S \;=\; k_B \ln(\Omega)$$

where $k_B$ is Boltzmann constant and $\Omega$ is the number of dynamical microstates, all of equal probability, consistent with a given macrostate; the microstates are specified by the position and the momentum of each molecule of gas[25].

The Second Law of Thermodynamics tells us that the entropy of an isolated system never decreases. Indeed, differentiating the previous equation we get

$$\delta S \;=\; k_B \frac{\delta\Omega}{\Omega} \;\geq 0.$$

We wish now to gain insights into Boltzmann's definition of entropy. He wanted to provide a macroscopic characterization on an ensemble of microscopic systems; he considered $N$ particles such that each particle is at one of the $m$ energy levels, $E_1 \leq E_2 \leq \ldots \leq E_m$. The number of particles at energy levels $E_1, E_2, \ldots, E_m$ are $N_1, N_2, \ldots, N_m$, respectively, with $N = \sum_{i=1}^{m} N_i$. The probability that a given particle is at energy level $E_i$ is $p_i = N_i/N, \;\; 1 \leq i \leq m$, Figure 70.

[25] A version of the equation $S \;=\; k_B \ln(\Omega)$ is engraved on Boltzmann's tombstone.

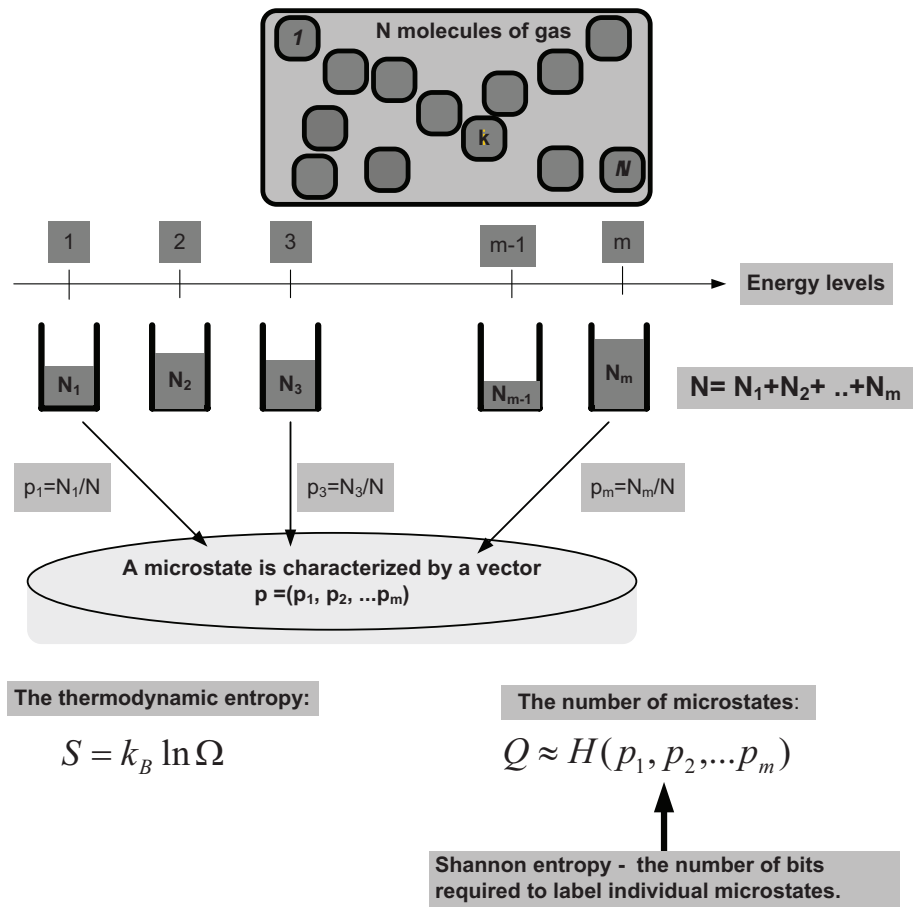



Figure 70: The relationship between the thermodynamic entropy and Shannon entropy.

A microstate $\sigma_i$ is characterized by an $m$-dimensional vector $(p_1, p_2, \ldots, p_m)$ when $N$ tends to infinity. The question posed by Boltzmann is how many microstates exist. To answer this question let us assume that we have $m$ boxes and we wish to count the number of different ways the given quantitative distribution of the $N$ particles among these boxes can be accomplished irrespective of which box a given particle is in. The answer to this question is

$$\binom{N}{N_1, N_2, \ldots, N_m} = \frac{N!}{N_1! N_2! \ldots N_m!} = \Omega.$$

Call $\mathcal{Q}$ the average number of bits required to characterize the state of the system; $\mathcal{Q}$ is proportional with the logarithm of $\Omega$, the number of microstates, more precisely,

$$\mathcal{Q} = \frac{1}{N} \log \Omega.$$

We will show that the number of bits, $\mathcal{Q}$, is given by

$$\mathcal{Q} = \frac{1}{N} \log \frac{N!}{N_1!, N_2!, \ldots, N_m!} = H(p_1, p_2, \ldots, p_m) + \mathcal{O}(N^{-1} \log N).$$

In this expression, $H(p_1, p_2, \ldots, p_m) = -\sum_i p_i \log p_i$ denotes the informational Shannon entropy, a quantity we shall have much to say about later in this chapter. To prove this equality we express $\mathcal{Q}$ as

$$\mathcal{Q} = \frac{1}{N}\left\{\log(N!) - [\log(N_1!) + \log(N_2!) + \ldots + \log(N_m!)]\right\}.$$

We use Stirling's approximation: $n! \approx \sqrt{2\pi n}(n^n/e^n)$ and $\mathcal{Q}$ becomes

$$\begin{aligned}
\mathcal{Q} \;\; = \tfrac{1}{N}[ \;\; & (\log\sqrt{2\pi N} + N\log N - N\log e) \\
- \;\; & (\log\sqrt{2\pi N_1} + N_1\log N_1 - N_1\log e) \\
- \;\; & (\log\sqrt{2\pi N_2} + N_2\log N_2 - N_2\log e) \\
- \;\; & \vdots \\
- \;\; & (\log\sqrt{2\pi N_m} + N_m\log N_m - N_m\log e)].
\end{aligned}$$

Now $\mathcal{Q}$ can be expressed as a sum of two terms

$$\mathcal{Q} = \mathcal{Q}_1 + \mathcal{Q}_2$$

with

$$\mathcal{Q}_1 = \log N - \left[\frac{N_1}{N}\log N_1 + \frac{N_2}{N}\log N_2 + \ldots + \frac{N_m}{N}\log N_m\right]$$

and

$$\mathcal{Q}_2 = \frac{1}{N}\log\frac{\sqrt{2\pi N}}{\sqrt{(2\pi N_1)(2\pi N_2)\ldots(2\pi N_m)}} = \frac{1}{N}\log\left[\frac{1}{(2\pi)^{(m-1)/2}}\sqrt{\frac{N}{N_1 N_2 \ldots N_m}}\right].$$

It is easy to see that

$$\log N = N\frac{1}{N}\log N = \frac{N_1 + N_2 + \ldots + N_m}{N}\log N = p_1\log N + p_2\log N + \ldots + p_m\log N.$$

Thus,

$$\mathcal{Q}_1 = p_1\log N + p_2\log N + \ldots + p_m\log N - [p_1\log N_1 + p_2\log N_2 + \ldots + p_m\log N_m].$$

But $p_i\log N - p_i\log N_i = -p_i\log(N_i/N) = -p_i\log p_i$ Thus,

$$\mathcal{Q}_1 = -\sum_i p_i\log p_i = H(p_1, p_2, \ldots, p_m).$$

We see that

$$Q_2 = \mathcal{O}(N^{-1}\log N).$$

We have thus shown that

$$\mathcal{Q} = H(p_1, p_2, \ldots, p_m) + \mathcal{O}(N^{-1}\log N).$$

We conclude that

$$\mathcal{Q} = H(p_1, p_2, \ldots, p_m) \quad \text{when} \quad N \to \text{very large.}$$

$\square$

This expression confirms our intuition; if all $N$ particles have the same energy, then the entropy of such a perfectly ordered system is zero. If a system has an infinite number of particles, all states are equally probable and the second law of thermodynamics is inapplicable to it.

To find the most probable microstate we have to maximize $H(p_1, p_2, \ldots, p_N)$ subject to some constraint. For example, if the constraint is $\sum_i p_i E_i = \kappa$ (the average energy is constant), then

$$p_i = \frac{e^{-\lambda E_i}}{\sum_j e^{\lambda E_j}}$$

with $\lambda$ the solution of the equation

$$\sum_i E_i \frac{e^{-\lambda E_i}}{\sum_j e^{\lambda E_j}} = \kappa.$$

This equation has a unique solution if $E_1 \leq \kappa \leq E_m$; this solution is known as the Maxwell-Boltzmann distribution.

Now we switch our attention from thermodynamics to messages generated by a source and label each message by a binary string of length $N$; let $p_0 = N_0/N$ be the probability of a 0 in a label and $p_1 = N_1/N$ be the probability of a 1 in a label. $N_0$ is the typical number of zeros and $N_1$ the typical number of one's in a label; then the number of bits required to label a message will be

$$\nu = \log \binom{N}{N_0, N_1} = -N \left[ p_0 \log p_0 + p_1 \log p_1 \right].$$

Thus, if we have an alphabet with $n$ letters occurring with probability $p_i$, $1 \leq i \leq n$, the average information in bits per letter transmitted when $N$ is large is

$$\frac{\nu}{N} = H(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log p_i$$

with $H(p_1, p_2, \ldots, p_n)$ the Shannon entropy. The connection between the thermodynamic entropy of Boltzmann and the informational entropy introduced by Shannon and discussed in depth in the next section is inescapable; Boltzmann's entropy can be written as

$$S = -k_B \ln 2 \sum_i p_i \log p_i \qquad \text{or} \qquad S = k_B \ln 2 \times H(p).$$

To convert one bit of classical information into units of thermodynamic entropy we multiply Shannon's entropy by the constant $(k_B \ln 2)$; the quantity $k_B \ln 2$ is the amount of entropy associated with the erasure of one bit of information according to Landauer's principle.

Next we discuss basic concepts of the classical information theory of Shannon.

## 3.3 Shannon Entropy

The statistical theory of communication introduced by Claude Shannon answers fundamental questions about a system consisting of agents exchanging classical information through a classical communication channel. Some of these questions are: How much information can be generated by an agent acting as a source of information? How much information can be squeezed through a classical channel? Is one agent capable to reproduce exactly, or only approximately a message generated by the source and transmitted though a classical channel? Should the information be encoded to improve the chance of reproducing exactly the message? Recall that encoding is the process of transforming information from one representation to another or "etching" it on a physical carrier and decoding is the process of restoring encoded information to its original form, or extracting information from a carrier.

In his endeavor to construct mathematically tractable models of communication Shannon concentrated on stationary and ergodic[26] sources of classical information. A stationary source of information emits symbols with a probability that does not change over time; an ergodic source emits information symbols with a probability equal to the frequency of their occurrence in a long sequence. Stationary ergodic sources of information have a finite but arbitrary and potentially long correlation time.

In the late 1940s, Shannon introduced a measure of the quantity of information a source could generate [376]. Earlier, in 1927, another scientist from Bell Labs, Ralph Hartley had proposed to take the logarithm of the total number of possible messages as a measure of the amount of information in a message generated by a source of information arguing that the logarithm tells us how many digits or characters are required to convey the message. Shannon recognized the relationship between thermodynamic entropy and informational entropy and, on von Neumann's advice, he called the negative logarithm of probability of an event, entropy[27].

Consider an event which happens with probability $p$; we wish to quantify the information content of a message communicating the occurrence of this event and we impose the condition that the measure should reflect the "surprise" brought by the occurrence of this event. An initial guess for a measure of this *surprise* would be $1/p$, the lower the probability of the event the larger the surprise. But this simplistic approach does not resist scrutiny; the surprise should be additive. If an event is composed of two independent events which occur with probabilities $q$ and $r$ then the probability of the event should be $p = qr$, but we see that

$$\frac{1}{p} \neq \frac{1}{q} + \frac{1}{r}.$$

On the other hand, if the *surprise* is measured by the logarithm of $1/p$, then the additivity property is obeyed

$$\log \frac{1}{p} = \log \frac{1}{q} + \log \frac{1}{r}.$$

[26] A stochastic process is said to be ergodic if time averages are equal to ensemble averages, in other words if its statistical properties such as its mean and variance can be deduced from a single, sufficiently long sample (realization) of the process

[27] It is rumored that von Neumann told Shannon "It is already in use under that name and, besides, it will give you a great edge in debates because nobody really knows what entropy is anyway" [78].

All logarithms are in base 2 unless stated otherwise. Given a probability distribution $\sum_i p_i = 1$ we see that the uncertainty is in fact equal to the average *surprise*

$$\sum_i p_i \log \frac{1}{p_i}.$$

The entropy is a measure of the uncertainty of a single random variable $X$ before it is observed, or the average uncertainty removed by observing it. This quantity is called entropy due to its similarity to the thermodynamic entropy.

The entropy of a random variable $X$ with a probability density function $p_X(x)$ is

$$H(X) = -\sum_x p_X(x) \log p_X(x).$$

The entropy of a random variable is a positive number. Indeed, the probability $p_X(x)$ is a positive real number between 0 and 1 therefore, $\log p_X(x) \leq 0$ and $H(X) \geq 0$.

Let $X$ be a binary random variable and $p = p_X(x = 1)$ be the probability that the $X$ takes the value 1; then the entropy of $X$, is

$$H(p) = -p \log p - (1-p) \log(1-p).$$

If the logarithm is in base 2 then the binary entropy is measured in bits. The entropy, depicted in Figure 71, has a maximum of 1 bit when $p = 1/2$ and goes to zero when $p = 0$ or $p = 1$; intuitively, we expect the entropy to be zero when the outcome is certain and reach its maximum when both outcomes are equally likely. It is easy to see that

(i) $H(X) > 0$ for $0 < p < 1$;

(ii) $H(X)$ is symmetric about $p = 0.5$;

(iii) $lim_{p \to 0} H(X) = lim_{p \to 1} H(X) = 0$;

(iv) $H(X)$ is increasing for $0 < p < 0.5$, decreasing for $0.5 < p < 1$ and has a maximum for $p = 0.5$.

(V) The binary entropy is a concave function of $p$, the probability of an outcome.

Before discussing this property of binary entropy we review a few properties of convex and concave functions. A function $f(x)$ is convex over an interval $(a, b)$ if

$$f(\gamma x_1 + (1-\gamma)x_2) \leq \gamma f(x_1) + (1-\gamma) f(x_2) \qquad \forall\, x_1, x_2 \in (a, b) \quad \text{and} \quad 0 \leq \gamma \leq 1.$$

The function is strictly convex iff this inequality holds only for $\gamma = 0$ or $\gamma = 1$. A function $f(x)$ is concave only if $(-f(x))$ is convex over the same interval.

It is easy to prove that if the second derivative of the function $f(x)$ is non-negative, then the function is convex. Call $x_0 = \gamma x_1 + (1-\gamma)x_2$. From the Taylor series expansion

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2}(x - x_0)^2, \quad x_0 \leq \xi \leq x$$

and from the fact that the second derivative is non-negative, it follows that $f''(\xi)(x-x_0)^2 \geq 0$. When $x = x_1$ then $x_1 - x_0 = (1-\gamma)(x_1 - x_2)$ and

$$f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0) \quad \text{or} \quad f(x_1) \geq f(x_0) + f'(x_0)(1-\gamma)(x_1 - x_2).$$

When $x = x_2$ then $x_2 - x_0 = \gamma(x_2 - x_1)$ and

$$f(x_2) \geq f(x_0) + f'(x_0)(x_2 - x_0) \quad \text{or} \quad f(x_2) \geq f(x_0) + f'(x_0)\gamma(x_2 - x_1).$$

It follows that

$$\gamma f(x_1) + (1-\gamma) f(x_2) \geq f(\gamma x_1 + (1-\gamma)x_2). \qquad \square$$

Convex functions enjoy a number of useful properties; for example, if $X$ is a discrete random variable with the probability density function $p_X(x_i)$ and $f(x)$ is a convex function, then $f(x)$ satisfies Jensen inequality

$$\sum_i p_X(x_i) f(x_i) \geq f\left(\sum_i p_X(x_i) x_i\right).$$

**Example.** *Given two probability density functions of the random variable $X$, $p_i = p_X(x_i)$ and $q_i = q_X(x_i)$ then*

$$\sum_i p_i \log p_i \geq \sum_i p_i \log q_i.$$

*Proof:* This is a consequence of the log sum inequality: given two sets of non-negative real numbers $\{a_1, a_2, \ldots, a_n\}$ and $\{b_1, b_2, \ldots, b_n\}$ then

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq \sum_{i=1}^n a_i \log \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}.$$

First, we observe that $f(x) = x \log x$ is a convex function as $f''(x) > 0$ when $x > 0$. Indeed, $f(x) = x \log e \ln x$, $f'(x) = \log e(1 + \ln x)$, and $f''(x) = (1/x) \log e$.

Call $a = \sum_{i=1}^n a_i$ and $b = \sum_{i=1}^n b_i$. Then according to Jensen inequality applied to the function $f(x) = x \log x$ we have

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} = \sum_{i=1}^n b_i \frac{a_i}{b_i} \log \frac{a_i}{b_i} = \sum_{i=1}^n b_i f\left(\frac{a_i}{b_i}\right) \leq b f\left(\sum_{i=1}^n \frac{b_i}{b} \frac{a_i}{b_i}\right) = b f\left(\frac{a}{b}\right) = b \frac{a}{b} \log \frac{a}{b}.$$

$\square$

**Example.** *Given a non-negative integer-valued random variable $X$ find the probability density function $p_i = p_X(x_i)$ which maximizes the entropy $H(X)$ subject to the constraint*

$$E(X) = \sum_i i p_i = A.$$

We apply the inequality $\sum_i p_i \log p_i \geq \sum_i p_i \log q_i$ with $q_i = \alpha\beta^i$ and $0 \leq \alpha, \beta \leq 1$:

$$H(X) = -\sum_i p_i \log p_i \leq -\sum_i p_i \log p_i = -\sum_i p_i(\log \alpha + i \log \beta) = -\log \alpha - A \log \beta.$$

To derive this inequality we noticed that $\sum_i p_i = 1$ and $\sum_i i p_i = A$; we also observe that the inequality holds independently of $p_i$. To determine the parameters $\alpha \leq 1$ and $\beta \leq 1$ we recall that $\sum_{i=0}^{\infty} \beta^i = 1/(1-\beta)$ and $\sum_{i=0}^{\infty} i\beta^i = \beta/(1-\beta)^2$ and write:

$$\sum_{i=0}^{\infty} q_i = \sum_{i=0}^{\infty} \alpha\beta^i = 1 \;\Rightarrow\; \frac{\alpha}{1-\beta} = 1 \quad \text{and} \quad A = \sum_{i=0}^{\infty} i q_i = \sum_{i=0}^{\infty} i\alpha\beta^i \;\Rightarrow\; A = \frac{\beta}{1-\beta}.$$

The two parameters $\alpha$ and $\beta$, the probability distribution $p_i$, and the maximum entropy are:

$$\alpha = \frac{1}{1+A}, \;\; \beta = \frac{A}{1+A}, \;\; p_i = \frac{1}{1+A}\left(\frac{A}{1+A}\right)^n, \quad \text{and} \quad H_{max}(X) = -\log \alpha - A \log \beta.$$

It is easy to prove that $f(x)$ is concave if and only if

$$f\left(\frac{x_1 + x_2}{2}\right) \geq \frac{f(x_1) + f(x_2)}{2}.$$

Figure 71 illustrates the fact that the binary entropy $H(p)$ is a concave function of $p$; the function lies above any chord, in particular above the cord connecting the points $(p_1, H(p_1))$ and $(p_2, H(p_2))$

$$H\left(\frac{p_1 + p_2}{2}\right) \geq \frac{H(p_1) + H(p_2)}{2}.$$

Table 2 shows some values for $H(X)$ for $0.0001 \leq p \leq 0.5$.

Table 2: The entropy of a binary random variable for $0.0001 \leq p \leq 0.5$.

| p | H(X) | p | H(X) | p | H(X) | p | H(X) |
|---|---|---|---|---|---|---|---|
| 0.0001 | 0.001 | 0.01 | 0.081 | 0.2 | 0.722 | 0.4 | 0.971 |
| 0.001 | 0.011 | 0.1 | 0.469 | 0.3 | 0.881 | 0.5 | 1.000 |

Given two random variables $X$ and $Y$ with the probability density functions $p_X(x)$ and $q_Y(y)$ let $p_{XY}(x, y)$ be the joint probability density function of $X$ and $Y$. To quantify the uncertainty about the pair $(x, y)$ we introduce the joint entropy of the two random variables.

The <u>joint entropy</u> of two random variables $X$ and $Y$ with $p_{XY}(x, y)$ the joint probability distribution function is defined as

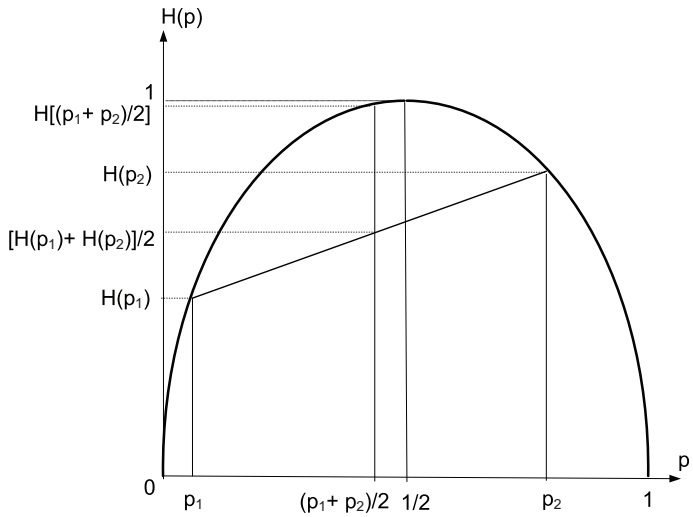



Figure 71: The entropy of a binary random variable function of the probability of an outcome.

$$H(X,Y) = -\sum_{x,y} p_{XY}(x,y) \log p_{XY}(x,y).$$

If we have acquired all the information about the random variable $X$ we may ask how much uncertainty is still there about the pair of the two random variables, $(X,Y)$. To answer this question we introduce the conditional entropy.

The conditional entropy of random variable $Y$ given $X$ with $p_{XY}(x,y)$ the joint probability distribution function and $p_{Y|X}(y|x)$ the conditional probability distribution function is defined as

$$H(Y|X) = -\sum_{x}\sum_{y} p_{XY}(x,y) \log p_{Y|X}(y|x).$$

**Example.** *Entropy, joint entropy and conditional entropy of $X$ and $Y$.* Consider two random variables $X$ and $Y$. Each of them takes values over a five-letter alphabet consisting of the symbols $a, b, c, d, e$. The joint distribution of the two random variables is given in Table 3.

Table 3: The joint probability distribution matrix of random variables X and Y.

| $p_{X,Y}(x,y)$ | a | b | c | d | e |
|---|---|---|---|---|---|
| a | 1/10 | 1/20 | 1/40 | 1/80 | 1/80 |
| b | 1/20 | 1/40 | 1/80 | 1/80 | 1/10 |
| c | 1/40 | 1/80 | 1/80 | 1/10 | 1/20 |
| d | 1/80 | 1/80 | 1/10 | 1/20 | 1/40 |
| e | 1/80 | 1/10 | 1/20 | 1/40 | 1/80 |

The marginal distribution of $X$ and $Y$ can be computed from the relations

$$p_X(x) = \sum_y p(x,y), \quad \text{and} \quad p_Y(y) = \sum_x p(x,y)$$

as follows

$$p_X(a) = \sum_y p_{XY}(x=a,y) = \frac{1}{10} + \frac{1}{20} + \frac{1}{40} + \frac{1}{80} + \frac{1}{80} = \frac{1}{5}.$$

Similarly, we obtain

$$p_X(x=b) = p_X(x=c) = p_X(x=d) = p_X(x=e) = \frac{1}{5},$$

and

$$p_Y(y=a) = p_Y(y=b) = p_Y(y=c) = p_Y(y=d) = p_Y(y=e) = \frac{1}{5}.$$

The entropy of $X$ is thus,

$$H(X) = -\sum_x p_X(x) \log p_X(x) = 5\left[\frac{1}{5}\log 5\right] = \log 5 \text{ bits.}$$

Similarly, the entropy of $Y$ is

$$H(Y) = -\sum_y p_Y(y) \log p_Y(y) = 5\left[\frac{1}{5}\log 5\right] = \log 5 \text{ bits.}$$

The joint probability of $X$ and $Y$ is

$$\begin{aligned} H(X,Y) &= -\textstyle\sum_x \sum_y p_{XY}(x,y) \log p_{XY}(x,y) \\ &= -5\left[\tfrac{1}{10}\log\tfrac{1}{10} + \tfrac{1}{20}\log\tfrac{1}{20} + \tfrac{1}{40}\log\tfrac{1}{40} + \tfrac{1}{80}\log\tfrac{1}{80} + \tfrac{1}{80}\log\tfrac{1}{80}\right] \\ &= \log 5\left[\tfrac{1}{2}\log\tfrac{1}{2} + \tfrac{1}{4}\log\tfrac{1}{4} + \tfrac{1}{8}\log\tfrac{1}{8} + \tfrac{1}{16}\log\tfrac{1}{16} + \tfrac{1}{16}\log\tfrac{1}{16}\right] \\ &= \tfrac{15}{8}\log 5 \text{ bits.} \end{aligned}$$

It is easy to see that

$$H(Y,X) = \frac{15}{8}\log 5 \text{ bits.}$$

To compute the matrix of conditional probabilities shown in Table 4 we use the known relation between the joint probability density and the conditional probability, $p_{X|Y}(x|y) = p_{XY}(x,y)/p_Y(y)$. For example:

$$p_{X|Y}(x=a|y=b) = \frac{p_{XY}(x=a,y=b)}{p_Y(y=b)} \quad \Longrightarrow \quad p_{X|Y}(x=a|y=b) = \frac{1/20}{1/5} = \frac{1}{4}.$$

The actual value of $H(X|Y)$ is

Table 4: The conditional probability distribution matrix of random variables X and Y.

| $p_{X\|Y}(x\|y)$ $p_{Y\|X}(y\|x)$ | a | b | c | d | e |
|---|---|---|---|---|---|
| a | 1/2 | 1/4 | 1/8 | 1/16 | 1/16 |
| b | 1/4 | 1/8 | 1/16 | 1/16 | 1/2 |
| c | 1/8 | 1/16 | 1/16 | 1/2 | 1/4 |
| d | 1/16 | 1/16 | 1/2 | 1/4 | 1/8 |
| e | 1/16 | 1/2 | 1/4 | 1/8 | 1/16 |

$$H(X|Y) = \sum_y p_Y(y) H(X|Y = a, b, c, d, e) = 5\frac{1}{5} H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\right) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}\right)$$

or

$$H(X|Y) = \left[\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{8}\log\frac{1}{8} + \frac{1}{16}\log\frac{1}{16} + \frac{1}{16}\log\frac{1}{16}\right] = \frac{15}{8} \text{ bits.}$$

In this particular example $H(Y|X)$ has same value

$$H(Y|X) = \frac{15}{8} \text{ bits.}$$

It is easy to see that $H(X), H(Y), H(X, Y),$ and $H(X|Y) \geq 0$. Indeed,

$$H(X, Y) = H(X|Y)H(Y) \quad \Longrightarrow \quad \frac{15}{8}\log 5 = \left[\frac{15}{8}\right]\log 5,$$

$$H(X, Y) = \frac{15}{8}\log 5 \quad \leq \quad H(X) + H(Y) = 2\log 5$$

$$H(X, Y) = H(Y, X),$$

$$H(X|Y) = \frac{15}{8} \quad \leq \quad H(X) = \log 5.$$

**Properties of joint and conditional entropy.** The joint and conditional entropy of random variables $X$, $Y$, and $Z$ have several properties summarized in Table 5. The fact that the joint entropy is symmetric and non-negative follows immediately from its definition. The conditional entropy is non-negative because $0 \leq p_{Y|X}(y \mid x) \leq 1$ and $0 \leq p_{X|Y}(x \mid y) \leq 1$; the equality occurs only when $Y$ is a deterministic function of $X$ or, respectively, $X$ is a deterministic function of $Y$.

From the definition of the conditional entropy and the expression of conditional probability, $p_{X|Y}(x \mid y) = p_{XY}(x, y)/p_Y(y)$ it follows that

Table 5: Properties of joint and conditional entropy.

| | |
|---|---|
| $H(X,Y) = H(Y,X)$ | symmetry of joint entropy |
| $H(X,Y) \geq 0$ | nonnegativity of joint entropy |
| $H(X \mid Y) \geq 0, \quad H(Y \mid X) \geq 0$ | nonnegativity of conditional entropy |
| $H(X \mid Y) = H(X,Y) - H(Y)$ | conditional and joint entropy relation |
| $H(X,Y) \geq H(Y)$ | joint entropy versus the entropy of a single rv |
| $H(X,Y) \leq H(X) + H(Y)$ | subadditivity |
| $H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z)$ | strong subadditivity |
| $H(X \mid Y) \leq H(X)$ | reduction of uncertainty by conditioning |
| $H(X,Y,Z) = H(X) + H(Y \mid X) + H(Z \mid X,Y)$ | chain rule for joint entropy |
| $H(X,Y \mid Z) = H(Y \mid X,Z) + H(X \mid Z)$ | chain rule for conditional entropy. |

$$\begin{aligned} H(X \mid Y) &= \textstyle\sum_x \sum_y p_{XY}(x,y) \log p_Y(y) - \sum_x \sum_y p_{XY}(x,y) \log p_{XY}(x,y) \\ &= -H(Y) + H(X,Y). \end{aligned}$$

But $H(X \mid Y) \geq 0$, thus $H(X,Y) \geq H(Y)$.

To prove <u>subadditivity</u> we first show that

$$\ln a \leq a - 1, \quad \text{if} \quad a \geq 1$$

with equality if and only if $a = 1$. Indeed, if we use the substitution $a = e^b$ the inequality becomes $b \leq e^b - 1$ for $b \geq 0$. From the Taylor series expansion of $e^b$ we have:

$$e^b - 1 = b(1 + \frac{b}{2!} + \frac{b^2}{3!} + \ldots \frac{b^k}{(k+1)!} + \ldots).$$

The parenthesis on the right hand side of this equation is greater than 1 when $b > 0$ and is equal to 1 only when $b = 0$. Thus, $e^b - 1 \geq y$, or $a - 1 \geq \ln a$; the last inequality can also be written as

$$\log a \leq \frac{1}{\ln 2}(a-1) \quad \text{if } a \geq 1.$$

It is easy to see that $\ln a = \log a \cdot \ln 2$; indeed, any $a \geq 1$ can be expressed as $a = e^{\ln a}$ but also as $a = 2^{\log a}$, thus $e^{\ln a} = 2^{\log a}$. If we apply on both sides the natural logarithm we get the desired expression.

Now we return to the subadditivity and express $H(X)$ and $H(Y)$ function of the joint probability density $p_{XY}(x,y)$

$$H(X) = -\sum_x p_X(x) \log p_X(x) = -\sum_x \left[ \sum_y p_{XY}(x,y) \right] \log p_X(x),$$

and

$$H(Y) = -\sum_x p_Y(y) \log p_Y(y) = -\sum_y \left[\sum_x p_{XY}(x,y)\right] \log p_Y(y).$$

We wish to show that $H(X,Y) - H(X) - H(Y) \leq 0$ and compute

$$H(X,Y) - H(X) - H(Y) = \sum_x \sum_y p_{XY}(x,y) \log \frac{p_X(x)p_Y(y)}{p_{XY}(x,y)}.$$

Here, $p_X(x)p_Y(y) \geq p_{XY}(x,y)$ with equality only when $X$ and $Y$ are independent random variables thus, we use the inequality $\log a \leq \frac{1}{\ln 2}(a-1)$ with $a = \frac{p_X(x)p_Y(y)}{p_{XY}(x,y)} \geq 1$ and rewrite the expression above as

$$\sum_x \sum_y p_{XY}(x,y) \log \frac{p_X(x)p_Y(y)}{p_{XY}(x,y)} \leq \frac{1}{\ln 2} \sum_x \sum_y p_{XY}(x,y) \left[\frac{p_X(x)p_Y(y)}{p_{XY}(x,y)} - 1\right].$$

Thus,

$$H(X,Y) - H(X) - H(Y) \leq \frac{1}{\ln 2} \sum_x \sum_y [p_X(x)p_Y(y) - p_{XY}(x,y)] \leq 0$$

or,

$$H(X,Y) - H(X) - H(Y) \leq \frac{1}{\ln 2} \left[\sum_x \sum_y p_X(x)p_Y(y) - \sum_x \sum_y p_{XY}(x,y)\right] = 0.$$

The *strong subadditivity* can be proved in a similar manner.

Finally, we show that $H(X) \leq H(X,Y)$:

$$\begin{aligned} H(X) - H(X,Y) &= \textstyle\sum_x \sum_y p_{XY}(x,y) \log \frac{p_X(x)}{p_{XY}(x,y)} \\ &\leq \textstyle\sum_x \sum_y p_{XY}(x,y) \left[\frac{p_X(x)}{p_{XY}(x,y)} - 1\right]. \end{aligned}$$

But

$$\sum_x \sum_y [p_X(x) - p_{XY}(x,y)] = \sum_x \sum_y [p_X(x)] - \sum_x \sum_y [p_{XY}(x,y)] = 1 - 1 = 0.$$

We conclude that

$$H(X) - H(X,Y) \leq 0, \qquad \text{or} \qquad H(X) \leq H(X,Y)$$

with equality if and only if $Y$ is a function of $X$.

The chain rule for joint entropy can be proved knowing the conditional and joint entropy relation $H(X,Y) = H(X) + H(Y \mid X)$. Then

$$H(X,Y,Z) = H(X) + H(Y,Z \mid X) = H(X) + H(Y \mid X) + H(Z \mid X,Y)$$

This rule can be generalized for $n$ random variables $X_1, X_2, \ldots, X_n$ as

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_{i-1}, X_{i-2}, \ldots, X_2, X_1).$$

To prove the chain rule for conditional entropy we start from the definition

$$H(X, Y \mid Z) = H(X, Y, Z) - H(Z).$$

We know that $H(X \mid Y) = H(X, Y) - H(Y)$ thus the right hand side of this equation can be written as

$$[H(X, Y, Z) - H(X, Z)] + [H(X, Z) - H(Z)] = H(Y \mid X, Z) - H(X \mid Z).$$

Thus,

$$H(X, Y \mid Z) = H(Y \mid X, Z) + H(X \mid Z).$$

The chain rule for the conditional entropy can be generalized for $n+1$ random variables:

$$H(X_1, X_2, \ldots, X_n \mid Y) = \sum_{i=1}^{n} H(X_i \mid Y, X_1, \ldots X_{i-1}).$$

Last, but not least, we mention without proof the grouping property of Shannon entropy [15].

$$\begin{aligned} H(p_1, p_2, ..., p_n) = \; & H(p_1 + p_2 + \ldots + p_\sigma + p_{\sigma+1} + \ldots + p_n) \\ & + \textstyle\sum_{k=1}^{\sigma} p_k H\left(p_1/(\sum_{k=1}^{\sigma} p_k), p_2/(\sum_{k=1}^{\sigma} p_k), \ldots, p_\sigma/(\sum_{k=1}^{\sigma} p_k)\right) \\ & + \textstyle\sum_{k=\sigma+1}^{n} p_k H\left(p_{\sigma+1}/(\sum_{k=\sigma+1}^{n} p_k), p_{\sigma+2}/(\sum_{k=\sigma+1}^{n} p_k), \ldots p_n/(\sum_{k=\sigma+1}^{n} p_k)\right). \end{aligned}$$

**Example.** *The entropy of a function of a random variable.* If $X$ is a discrete random variable and $Y$ is a function of $X$, $Y = f(X)$ then its entropy cannot be larger than the entropy of $X$, $H(X) \geq H(f(X))$; when $f$ is an one-to-one function we have the equality $H(X) = H(f(X))$.

*Proof:* From the chain rule of joint entropy it follows that

$$H(X, f(X)) = H(X) + H(f(X) \mid X) = H(X).$$

Indeed, $H(f(X) \mid X) = 0$ as $f$ is a deterministic function, for any given value $x$ of $X$, $f(x)$ is fixed. The chain rule for mutual information allow us to express $H(X, f(X))$ as $H(X, f(X)) = H(f(X)) + H(X \mid f(X))$. But $H(X \mid f(X)) \geq 0$ with equality if and only if $f$ is one-to one. We conclude that $H(f(X)) \leq H(X)$. For example, when $Y = X^{2n+1}$ then $H(Y) = H(X)$ as $f$ is invertible; when $Z = \tan(X)$ then $H(Z) < H(X)$ because the tangent is not a one-to-one function.

**Example.** *The entropy of a sum of random variables,* $Z = X + Y$. If $Z = X + Y$ then $H(Z \mid X) = H(Y \mid Y)$. When $X$ and $Y$ are independent $H(X) \leq H(Z)$ and $H(Y) \leq H(Z)$, addition or independent random variables adds uncertainty. Moreover, when $X$ and $Y$ are independent and the function $Z = f(X, Y) = X+Y$ is one-to-one then $H(Z) = H(X)+H(Y)$.

*Proof:* First we observe that $p_{Z|X}(z \mid x) = p_{Y|X}(y \mid x)$; Indeed,
$\text{Prob}[Z = z \mid X = x] = \text{Prob}[X + Y = x \mid X = x] = \text{Prob}[Y = z - x \mid X = x]$.
Then:

$$\begin{aligned}
H(Z \mid X) &= -\sum_x \sum_z p_{Z|X}(z \mid x) p_X(x) \log p_{Z|X}(z \mid x) \\
&= -\sum_x p_X(x) \sum_z p_{Z|X}(z \mid x) \log p_{Z|X}(z \mid x) \\
&= -\sum_x p_X(x) \sum_z p_{Y|X}(z - x \mid x) \log p_{Y|X}(z - x \mid x) \\
&= -\sum_x p_X(x) \sum_y p_{Y|X}(y \mid x) \log p_{Y|X}(y \mid x) \\
&= -\sum_x \sum_y p_{Y|X}(y \mid x) p_X(x) \log p_{Y|X}(y \mid x) \\
&= H(Y \mid X).
\end{aligned}$$

We now show that when $X$ and $Y$ are independent $H(Z) \geq H(Y)$; indeed, in this case $H(Y \mid X) = H(Y)$. Then $H(Z \mid Z) = H(Y \mid X) = H(Y)$. But $H(Z) \geq H(Z \mid Y)$ so $H(Z) \geq H(Y)$; a similar argument leads to $H(Z) \geq H(X)$. Lastly, when the function $f$ in $Z = f(X, Y) = X+Y$ is one-to-one, then according to the previous example $H(Z) = H(X, Y)$; when $X$ and $Y$ are independent $H(X, Y) = H(X) + HY$, therefore $H(Z) = H(X) + H(Y)$.

We conclude the discussion of entropy with the observation that in his original paper Shannon proved that a logarithmic measure for the entropy can be derived from several axioms [375]. Given a sequence of symmetric functions $H_1(p_1), H_2(p_1, p_2), \ldots, H_n(p_1, p_2, p_3, \ldots, p_n)$ the function $H_n$ must be of the form $H_n(p_1, p_2, p_3, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log p_i$ when the functions satisfy three axioms:

(i) *Normalization:* $H_2(1/2, 1/2) = 1, \ldots, H_n(1/n, 1/n, \ldots, 1/n) = 1$;

(ii) *Continuity:* $H_2(p, 1 - p)$ is a continuous function of $p$;

(iii) *Grouping:*
$H_n(p_1, p_2, p_3, \ldots, p_n) = H_{n-1}(p_1 + p_2, p_3, \ldots, p_n) + (p_1 + p_2) H_2(p_1/(p_1 + p_2), p_2/(p_1 + p_2))$.

## 3.4 Shannon Source Coding

We know that a source of classical information with probability distribution $p_X(x)$ is characterized by Shannon entropy $H(p_X(x))$. Source coding addresses two questions: (i) Is there an upper bound for the data compression level and, if so what is this upper bound; (ii) Is it always possible to reach this upper bound? The answer to both questions is "yes." Classical, as well as quantum source coding discussed in Section 3.14, are based on the idea of *typical sequence of symbols* from an alphabet and *typical subspaces of a Hilbert space,* respectively.

Informally, Shannon source encoding theorem states that a message containing $n$ independent, identically distributed samples of a random variable $X$ with entropy $H(X)$ can be compressed to a length

$$l_X(n) = nH(X) + \mathcal{O}(n).$$

The justification of this theorem is based on the *weak law of large numbers* which states that $\bar{x} = \sum_i x_i p_{x_i}$, the mean of a large number of independent, identically distributed random variables, $x_i$, approaches $1/n \sum_i x_i$, the average, with a high probability when $n$ is large:

$$\text{Prob}\left[ |\frac{1}{n}\sum_{i=1}^{n} x_i - \bar{x}| > \delta \right] < \epsilon,$$

with $\delta$ and $\epsilon$ two arbitrarily small positive real numbers.

Given a source of classical information $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ with an alphabet with $m$ symbols, let $\{p(a_1), p(a_2), \ldots, p(a_n)\}$ be the probabilities of individual symbols; then the Shannon entropy of this source is

$$H(\mathcal{A}) = -\sum_{i=1}^{m} p(a_i) \log p(a_i).$$

A message $a = (a_{k_1}, a_{k_2}, \ldots, a_{k_n})$ consisting of a sequence of $n$ symbols independently selected from the input alphabet has a probability

$$P(a) = p(a_{k_1})p(a_{k_2}) \ldots p(a_{k_n}).$$

If we define a random variable $x_i = -\log p(a_i)$ then, we can establish the following correspondence with the quantities in the expression of the weak law of large numbers:

$$\sum_{i=1}^{n} x_i = -\sum_{i=1}^{n} \log p(a_{k_i}) = -\log P(a),$$

$$\bar{x} = \sum_{i=1}^{m} x_i p(x_i) = -\sum_{i=1}^{m} p(a_i) \log p(a_i) = H(\mathcal{A}).$$

It follows that given two arbitrarily small real numbers $(\delta, \epsilon) \geq 0$ then, for sufficiently large $n$, the following inequality holds [369, 375]

$$\text{Prob}\left[ | -\frac{1}{n} \log P(a) - H(\mathcal{A})| > \delta \right] < \epsilon.$$

This inequality partitions $\Lambda$, the set of stings of length $n$, into two subsets, see Figure 72:

(a) The subset of <u>Typical</u> strings

$$\Lambda_{Typical} = \left\{ a_{Typical} : | -\frac{1}{n} \log P(a_{Typical}) - H(\mathcal{A})| \leq \delta \right\}$$

that occur with high probability, $\text{Prob}\,(a_{Typical}) \geq (1 - \epsilon)$; and

(b) The subset of <u>Atypical</u> strings

$$\Lambda_{Atypical} = \left\{ a_{Atypical} : | -\frac{1}{n} \log P(a_{Atipical}) - H(\mathcal{A})| > \delta \right\}.$$

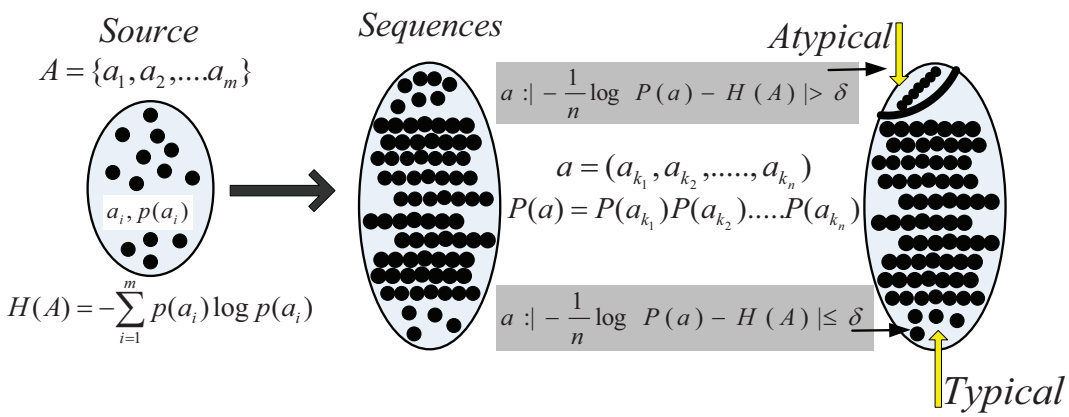



Figure 72: Source coding. The inequality Prob $\left[| -\frac{1}{n}\log P(a) - H(\mathcal{A})| > \delta\right] < \epsilon$ partitions the set of strings of length $n$, into two classes: (i) typical strings that occur with high probability, and (ii) atypical strings that occur with a vanishing probability

that occur with a vanishing probability, Prob $(a_{Atypical}) < \epsilon$. The two subsets are disjoint and complementary:

$$\Lambda_{Typical} \bigcap \Lambda_{Atypical} = \emptyset \quad \text{and} \quad \Lambda = \Lambda_{Typical} \bigcup \Lambda_{Atypical}.$$

We concentrate on typical strings generated by the source as atypical strings occur with a vanishing probability and can be ignored. Now we shall determine $\mid \Lambda_{Typical} \mid$, the cardinality of the set of typical strings, $a \in \Lambda_{Typical}$. The inequality defining the strings in this subset can be rewritten as

$$-\delta \leq -\frac{1}{n}\log P(a_{Typical}) - H(\mathcal{A}) \leq +\delta \quad \Longrightarrow \quad 2^{-n(H(\mathcal{A})+\delta)} \leq P(a_{Typical}) \leq 2^{-n(H(\mathcal{A})-\delta)}.$$

The first inequality can be expressed in terms of the cardinality set:

$$\sum_{a \in \Lambda_{Typical}} P(a_{typical}) \geq \sum_{a \in \Lambda_{Typical}} 2^{-n(H(\mathcal{A})+\delta)} = \mid \Lambda_{Typical} \mid 2^{-n(H(\mathcal{A})+\delta)}.$$

It follows that

$$\mid \Lambda_{Typical} \mid \quad \leq \quad 2^{n(H(\mathcal{A})+\delta)}.$$

Similarly, the second inequality can bee expressed as

$$\sum_{a \in \Lambda_{Typical}} P(a_{typical}) \leq \sum_{a \in \Lambda_{Typical}} 2^{-n(H(\mathcal{A})-\delta)} = \mid \Lambda_{Typical} \mid 2^{-n(H(\mathcal{A})-\delta)}.$$

This implies that

$$\mid \Lambda_{Typical} \mid \quad \geq \quad 2^{n(H(\mathcal{A})-\delta)}.$$

When $\delta \mapsto 0$ then $\mid \Lambda_{Typical} \mid \ \mapsto \ 2^{nH(\mathcal{A})}$, the cardinality of the set of typical string converges to $2^{n(H(\mathcal{A})-\delta)}$. *There are* $2^{nH(\mathcal{A})}$ *typical strings, therefore we need* $\log 2^{nH(\mathcal{A})} = nH(\mathcal{A})$ *bits to encode all possible typical strings;* this is the upper bound for the data compression provided by Shannon's classical source encoding theorem.

**Example.** *Message length and entropy.* Eight cars labelled $\mathbb{C}_1, \mathbb{C}_2, \ldots, \mathbb{C}_8$ compete in several Formula I races. The probability of winning calculated based on the past race history for the eight cares, $p_1, p_2, \ldots, p_8$, are:

$$p_1 = \frac{1}{2},\ p_2 = \frac{1}{4},\ p_3 = \frac{1}{8},\ p_4 = \frac{1}{16},\ p_5 = \frac{1}{64},\ p_6 = \frac{1}{64},\ p_7 = \frac{1}{64},\ \text{ and }\ p_8 = \frac{1}{64}.$$

To send a binary message revealing the winner of a particular race we could encode the identities of the winning car in several ways. For example, we can use an "obvious" encoding scheme; the identities of $\mathbb{C}_1, \mathbb{C}_2, \mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_5, \mathbb{C}_6, \mathbb{C}_7, \mathbb{C}_8$ could be encoded using three bits, the binary representation of integers 0 to 7, namely $000, 001, 010, 011, 100, 101, 110, 111$, respectively. Obviously, in this case we need three bits, the average length of the string used to communicate the winner of any race is $\bar{l} = 3$.

Let us now consider an encoding which reduces the average number of bits transmitted. The cars have different probability to win a race and it makes sense to assign a shorter string to a car which has a higher probability to win. Thus, a better encoding of the identities of $\mathbb{C}_1, \mathbb{C}_2, \mathbb{C}_3, \mathbb{C}_4, \mathbb{C}_5, \mathbb{C}_6, \mathbb{C}_7, \mathbb{C}_8$ is

$$0, 10, 110, 1110, 111100, 111101, 111110, 111111.$$

The corresponding lengths of the strings encoding the identity of each car are:

$$l_1 = 1,\ \ l_2 = 2,\ \ l_3 = 3,\ \ l_4 = \ 4,\ \ l_5 = l_6 = l_7 = l_8 = 6.$$

In this case, $\bar{l}$, the average length of the string used to communicate the winner is

$$\bar{l} \ = \ \sum_{i=1}^{8} l_i \times p_i \ = \ 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + 4 \times (4 \times \frac{1}{64}) = 2 \text{ bits.}$$

Using this encoding scheme the expected length of the string designating the winner over a large number of car races is less than the one for "obvious" encoding presented above. Now we calculate the entropy of the random variable $W$ indicating the winner

$$H(W) = -\frac{1}{2} \times \log \frac{1}{2} - \frac{1}{4} \times \log \frac{1}{4} - \frac{1}{8} \times \log \frac{1}{8} - \frac{1}{16} \times \log \frac{1}{16} - \frac{4}{64} \times \log \frac{1}{64} = 2 \text{ bits.}$$

We observe that the average length of the string identifying the outcome of a race for this particular encoding scheme is equal with the entropy, which means that this is an optimal encoding scheme. Indeed, the entropy provides the average information, or, equivalently, the average uncertainty removed by receiving a message.

While the entropy is a quantitative measure for an information source, the mutual and the conditional entropy together with the mutual information introduced in the next section provide a characterization of a communication channel, another element of Shannon's communication model.

## 3.5 Mutual Information; Relative Entropy

We introduce two important concepts from Shannon's Information Theory: the *mutual information* which measures the reduction in uncertainty of a random variable $X$ due to another random variable $Y$, and the *conditional mutual information* of $X$, $Y$, and $Z$ which quantifies the reduction in the uncertainty of random variables $X$ and $Y$ given $Z$.

The <u>mutual information</u> of two random variables $X$ and $Y$ with the joint probability distribution function $p_{XY}(x, y)$ is defined as

$$I(X;Y) = \sum_x \sum_y p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}.$$

The symmetry, $I(X;Y) = I(Y;X)$, follows immediately from the definition of mutual information. To establish the relation of mutual information with entropy and conditional entropy we use the fact that $p_{XY}(x,y) = p_{X|Y}(x \mid y)p_Y(y)$

$$\begin{aligned} I(X;Y) &= -\textstyle\sum_x \sum_y p_{XY}(x,y) \log p_X(x) + \sum_x \sum_y p_{XY}(x,y) \log p_{X|Y}(x \mid y) \\ &= H(X) - H(X \mid Y). \end{aligned}$$

We also note that $H(X \mid X) = 0$ thus $I(X;X) = H(X)$. We have proved earlier that $H(X) \geq H(X \mid Y)$ thus $I(X,Y) \geq 0$ with equality only if $X$ and $Y$ are independent.

The equality $I(X;Y) = H(X) + H(Y) - H(X,Y)$ follows immediately from the definition of $I(X;Y)$

$$\begin{aligned} I(X;Y) &= -\textstyle\sum_x \left[\sum_y p_{XY}(x,y)\right] \log p_X(x) - \sum_y \left[\sum_x p_{XY}(x,y)\right] \log p_Y(y) \\ &\quad + \textstyle\sum_x \sum_y p_{XY}(x,y) \log p_{XY}(x,y) \\ &= H(X) + H(Y) - H(X,Y). \end{aligned}$$

The <u>conditional mutual information</u> of random variables $X$, $Y$, and $Z$ is defined as

$$I(X;Y \mid Z) = \sum_x \sum_y \sum_z p_{XYZ}(x,y,z) \log \frac{p_{XY|Z}(x,y \mid z)}{p_{X|Z}(x \mid z)p_{Y|Z}(y \mid z)}.$$

It follows immediately that

$$I(X;Y \mid Z) = H(X \mid Z) - H(X \mid Y,Z).$$

The mutual information of random variables $X$, $Y$, and $Z$ has several properties summarized in Table 6. To prove the chain rule for mutual information we start from the definition of mutual information of three random variables:

Table 6: Properties of mutual information.

| | |
|---|---|
| $I(X;Y) = I(Y;X)$ | symmetry of mutual entropy |
| $I(X;Y) = H(X) - H(X \mid Y)$ | mutual information, entropy, and conditional entropy |
| $I(X;Y) = H(Y) - H(Y \mid X)$ | mutual information, entropy, and conditional entropy |
| $I(X;X) = H(X)$ | mutual self information and entropy |
| $I(X;X) \geq 0,$ | non-negativity of mutual self information |
| $I(X;Y) = H(X) + H(Y) - H(X,Y)$ | mutual information, entropy, and joint entropy |
| $I(X;Y \mid Z) = H(X \mid Z) - H(X \mid Y,Z)$ | conditional mutual information and conditional entropy |
| $I(X,Y;Z) = I(X;Z \mid Y) + I(Y;Z)$ | chain rule for mutual information |
| $I(X;Y) \leq I(X;Z)$ if $X \mapsto Y \mapsto Z$ | data processing inequality |

$$I(X,Y;Z) = H(X,Y) - H(X,Y \mid Z) = [H(X \mid Y) + H(Y)] - [H(X \mid Y,Z) + H(Y \mid Z)]$$

But,

$$H(X \mid Y) - H(X \mid Y,Z) = I(X;Z \mid Y) \quad \text{and} \quad H(Y) - H(Y \mid Z) = I(Y;Z).$$

Thus,

$$I(X,Y;Z) = I(X;Z \mid Y) + I(Y;Z).$$

This chain rule can be generalized to $n+1$ random variables, $X_1, X_2, \ldots X_n, Y$ as

$$I(X_1, X_2, \ldots X_n; Y) = \sum_i I(X_i; Y \mid X_1, X_2, \ldots X_{i-1}).$$

The diagrams in Figure 73 illustrates the relations among the mutual entropy, joint entropy, and conditional entropies of random variables $X$ and $Y$ we have discussed earlier:

$$I(X;Y) = H(X) + H(Y) - H(X,Y),$$

$$H(X,Y) = H(X \mid Y) + H(Y \mid X) + I(X;Y),$$

$$H(X \mid Y) = H(X) - I(X;Y), \quad H(Y \mid X) = H(Y) - I(X;Y).$$

In practice, we rarely have complete information about an event, there are discrepancies between what we expect and the real probability distribution of a random variable. For example, to schedule its gate availability an airport assumes that a particular flight arrives on time 75% of the time, while in reality the flight is on time only 61% of the time.

To measure how close two distributions are from each other we introduce the concept of *relative entropy.* Given a random variable and two probability distributions $p_X(x)$ and $q_X(x)$

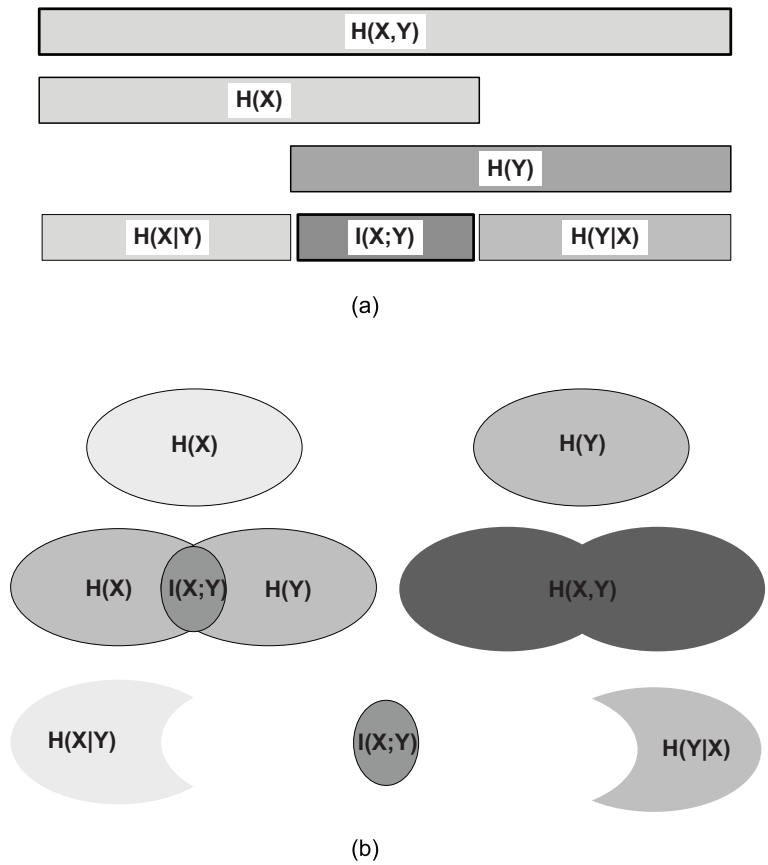



Figure 73: Graphical representation of the relationship between $H(X)$ and $H(Y)$, the entropy of random variables $X$ and $Y$, the mutual information $I(X;Y)$, the joint entropy $H(X,Y)$, and the conditional entropy, $H(X \mid Y)$. (a) An intuitive representation; (b) Traditional representation using Venn diagrams.

we wish to develop an entropy-like measure of how close the two distributions are. If the real probability distribution of events is $p_X(x)$ but we erroneously assume that it is $q_X(x)$ then the surprise when an event occurs is $\log(1/q_X(x))$ and the average surprise is

$$\sum_x p_X(x) \log \frac{1}{q_X(x)}.$$

In this expression we use $p_X(x)$, the correct probabilities for averaging. Yet, the correct amount of information is given by Shannon's entropy, $H(X) = -\sum p_X(x) \log p_X(x)$. The relative entropy is defined as the difference between the average surprise and Shannon's entropy.

The relative entropy between two distributions $p_X(x)$ and $q_X(x)$, also called the Kullbach-Leibler distance, is defined as

$$H(p_X(x) \mid\mid q_X(x)) = \sum_x p_X(x) \log \frac{p_X(x)}{q_X(x)}.$$

The relative entropy is not a metric in mathematical sense, it is not symmetric:

$$H(p_X(x) \mid\mid q_X(x)) \neq H(q_X(x) \mid\mid p_X(x)).$$

The expression of the relative entropy can also be written as

$$H(p_X(x) \mid\mid q_X(x)) = \sum_x p_X(x) \log p_X(x) + \sum_x p_X(x) \log \frac{1}{q_X(x)} = -H(X) - \sum_x p_X(x) \log q_X(x).$$

This expression justifies our assertion that the relative entropy is equal to the difference between the average surprise and Shannon's entropy. We show now that the relative entropy is non-negative thus, as pointed out in [437], we notice an "uncertainty deficit" caused by our inaccurate assumptions; our average surprise is larger than Shannon's entropy.

**Proposition.** *The relative entropy between two distributions $p_X(x)$ and $q_X(x)$ of the discrete random variable $X$ is non-negative; it is zero only when $p_X(x) = q_X(x)$.*

*Proof:* By virtue of the inequality $\ln x \leq x - 1$ for $x \geq 0$ and of the fact that $\log x = \ln x / \ln 2$ it follows that

$$H(p_X(x) \mid\mid q_X(x)) = \sum_x p_X(x) \log \frac{p_X(x)}{q_X(x)} \geq \frac{1}{\ln 2} \sum_x p_X(x) \left[1 - \frac{q_X(x)}{p_X(x)}\right].$$

But

$$\sum_x p_X(x) \left[1 - \frac{q_X(x)}{p_X(x)}\right] = \sum_x p_X(x) - \sum_x q_X(x) = 1 - 1 = 0.$$

Thus

$$H(p_X(x) \mid\mid q_X(x)) \geq 0$$

with equality if and only if $p_X(x) = q_X(x)$. □

This shows that indeed relative entropy is a measure of the similarity of the two distributions. For example, if $X$ is a binary random variable and the two distributions are:

$$p_X(0) = 1 - a,\ p_X(1) = a \quad \text{and} \quad q_X(0) = 1 - b, \quad q_X(1) = b, \quad 0 \leq (a, b) \leq 1$$

then,

$$D(p_X(x) \mid\mid q_X(x)) = (1 - a) \log \frac{1 - a}{1 - b} + a \log \frac{a}{b}$$

and

$$D(q_X(x) \mid\mid p_X(x)) = (1 - b) \log \frac{1 - b}{1 - a} + b \log \frac{b}{a}.$$

It is easy to see that $D(p_X(x) \mid\mid q_X(x)) = D(q_X(x) \mid\mid p_X(x))$ if and only if $a = b$.

As a final observation, the mutual information $I(X; Y)$ is the relative entropy between the joint density function $p_{XY}(x, y)$ and the product of the probability density functions $p_X(x)$ and $p_Y(y)$,

$$I(X,Y) = \sum_x \sum_y p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)} = D(p_{XY}(x,y) \mid\mid p_X(x)p_Y(y)).$$

When $X$ and $Y$ are uncorrelated, $p_{XY}(x,y) = p_X(x)p_Y(y)$, and then both the mutual information and the relative entropy are equal to zero.

The next question we address is to estimate the error when we use a random variable $Y$ to guess the value of a correlated random variable $X$.

## 3.6 Fano Inequality; Data Processing Inequality

Fano's inequality is used to find a lower bound on the error probability of any decoder; it relates the average information lost in a noisy channel to the probability of the categorization error. We first introduce the concept of a Markov chain, then state Fano's inequality, and finally prove the data processing inequality, a result with many applications in classical information theory.

Random variables $X$, $Y$, and $Z$ form a <u>Markov chain</u> denoted as $X \mapsto Y \mapsto Z$, if the conditional probability distribution of $Z$ depends only on $Y$ (it is independent of $X$)

$$p_{XYZ}(x,y,z) = p_X(x)p_{Y|X}(y \mid x)p_{Z|Y}(z \mid y) \quad \Longrightarrow \quad X \mapsto Y \mapsto Z.$$

**Example.** *Mutual information and Markov chains.* The random variables $X_1, X_2, X_3, \ldots, X_n$ form a Markov chain in the order $X_1 \mapsto X_2 \mapsto X_3 \mapsto \ldots \mapsto X_n$; this means that

$$p_{X_1X_2\ldots X_n} = p_{X_1}(x)p_{X_2|X_1}(x_2 \mid x_1)p_{X_3|X_2}(x_3 \mid x_2)\ldots p_{X_n|X_{n-1}}(x_n \mid x_{n-1}).$$

It is easy to see that

$$I(X_1 : X_2, X_3, \ldots, X_N) = I(X_1; X_2).$$

Indeed, we apply the chain rule for mutual information and

$$I(X_1 : X_2, X_3 \ldots, X_n) = I(X_1; X_2) + I(X_1; X_3 \mid X_2) + I(X_1; X_n \mid X_2, X_3, \ldots, X_{n-1}).$$

The random variables $X_1, X_2, X_3, \ldots, X_n$ could represent the states of a system and, as we know, the Markov property means that the past and the future state are conditionally independent given the present state; this implies that all terms in this sum except the first are equal to zero.

Consider now a random variable $X$ with the probability density function $p_X(x)$ and let $\mid \mathcal{X} \mid$ denote the number of elements in the range of $X$. Let $Y$ be another random variable related to $X$, with the conditional probability $p_{Y|X}(y \mid x)$. Intuitively, we expect that the error when we use $Y$ to estimate $X$ to be small when the conditional entropy $H(X \mid Y)$ is small. Our intuition is quantified by Fano Inequality.

**Proposition.** *Fano's Inequality: when we estimate $X$ based on the observation of $Y$ as $\hat{X} = f(X)$ we make an error $p_{err} = Prob(\tilde{X} \neq X)$ and*

$$H(p_{err}) + p_{err} \log(\mid \mathcal{X} \mid -1) \geq H(X \mid Y).$$

Weaker forms of Fano's inequality are :

$$1 + p_{err} \log(\mid \mathcal{X} \mid) \geq H(X \mid Y) \quad \text{and} \quad p_{err} \geq \frac{H(X \mid Y) - 1}{\log(\mid \mathcal{X} \mid)}.$$

*Proof:* We assume that the discrete random variable $X$ has a probability distribution $p_X(x)$ and call $\mathcal{X}$ the range of values of $X$. We base our estimation of $X$ on the random variable $Y$ which is related to $X$ by the conditional density function $p_{Y|X}(y \mid x)$; once we know $Y$ we estimate $X$ as a function of $Y$, $\tilde{X} = g(Y)$. Clearly $X, Y, \tilde{X}$ form a Markov chain, $X \mapsto Y \mapsto \tilde{X}$. Now we introduce a binary random variable $A$ defined as

$$A = \begin{cases} 1 & \text{if} \quad \tilde{X} \neq X \\ 0 & \text{if} \quad \tilde{X} = X \end{cases}$$

We apply the chain rule for entropy and write

$$H(A, X \mid Y) = H(X \mid Y) + H(A \mid X, Y) = H(X \mid Y).$$

Indeed, $H(A \mid X, Y) = 0$ as $A$ is completely determined once we know $X$ and $Y$. We apply again the chain rule for entropy to express $H(A, X \mid Y)$ differently

$$H(A, X \mid Y) = H(A \mid Y) + H(X \mid A, Y).$$

$A$ is a binary random variable and $\text{Prob}(A = 0) = 1 - p_{err}$ and $\text{Prob}(A = 1) = p_{err}$ thus, $H(A) = H(p_{err})$; also $H(A \mid Y) \leq H(A)$ as conditioning reduces the entropy. Now,

$$H(X \mid A, Y) = \text{Prob}(A = 0) \cdot H(X \mid Y, A = 0) + \text{Prob}(A = 1) \cdot H(X \mid Y, A = 1).$$

We recognize that $H(X \mid Y, A = 0) = 0$ because $X = \tilde{X}$ when $A = 0$; $\mid \mathcal{X} \mid$ is the cardinality of the target set of $X$ thus, $H(X \mid Y, A = 1) = \log(\mid \mathcal{X} \mid -1)$ and conclude that

$$H(X \mid A, Y) \geq H(p_{err}) + (1 - p_{err}) \cdot 0 + p_{error} \log(\mid \mathcal{X} \mid -1) = H(p_{err}) + p_{error} \log(\mid \mathcal{X} \mid -1).$$

Combing the two expressions of $H(X \mid A, Y)$ we obtain the Fano inequality

$$H(p_{err}) + p_{error} \log(\mid \mathcal{X} \mid -1) \geq H(X \mid Y)$$

But $H(p_{err}) \leq 1$ and $\mid \mathcal{X} \mid$, the cardinality of the set of sample values of the random variable $X$, is expected to be large, $\mid \mathcal{X} \mid -1 \approx \mid \mathcal{X} \mid$; therefore, weaker forms of Fano inequality are:

$$1 + p_{error} \log(\mid \mathcal{X} \mid) \geq H(X \mid Y) \quad \text{or} \quad p_{err} \geq \frac{H(X \mid Y) - 1}{\log(\mid \mathcal{X} \mid)}. \qquad \square$$

**Proposition.** *Data processing inequality: if random variables $X, Y, Z$ form a Markov chain, $X \to Y \to Z$, then*

$$I(X;Z) \leq I(X;Y).$$

*Proof:* The inequality follows from the chain rule which allows us to expand the mutual information in two different ways:

$$\begin{aligned} I(X;Y,Z) &= I(X;Z) + I(X;Y \mid Z) \\ &= I(X;Y) + I(X;Z \mid Y). \end{aligned}$$

$X$, $Y$, and $Z$ form a Markov chain $X \mapsto Y \mapsto Z$, and this implies that $X$ and $Z$ are independent given $Y$, and $I(X;Z \mid Y) = 0$. Then,

$$I(X;Z) + I(X;Y \mid Z) = I(X;Y) \quad \Longrightarrow \quad I(X;Z) \leq I(X;Y)$$

as $I(X;Y \mid Z)$ is non-negative, $I(X;Y \mid Z) \geq 0$.

Informally, the data processing inequality states that one cannot gather more information, $Z$, by processing a set of data, $Y$, than the information provided by the data to begin with, $X$; e.g., no amount of signal processing could increase the information we receive from a space probe. The quantum correspondents of Fano and data processing inequalities will be discussed in Section 3.19.

## 3.7 Classical Information Transmission through Discrete Channels

When two agents communicate, each one can influence the physical state of the other through some physical process. The precise nature of the agents and of the signalling process can be very different, thus it is necessary to consider an abstract model of communication. In this model a source has the ability to select among a number of distinct physical signals, pass the selected signal to a communication channel which then affects the physical state of the destination; finally, the destination attempts to identify precisely the specific signal that caused the change of its physical state.

We are interested in two fundamental questions regarding communication over noisy channels: (i) Is it possible to encode the information transmitted over a noisy channel to minimize the probability of errors? (ii) How does the noise affect the capacity of a channel?

Intuitively, we know that whenever the level of noise in a room or on a phone line is high, we have to repeat words and sentences several times before the other party understands us. Thus, the actual rate we are able to transmit information through a communication channel is lowered by the presence of the noise. Rigorous answers to both questions, consistent with our intuition, are provided by two theorems, the *source coding theorem* discussed in Section 3.4, which provides an upper bound for the compression of a message and the *channel coding theorem* which states that the two agents can communicate at a rate close to the channel capacity $C$ and with a probability of error $\epsilon$ arbitrarily small [376, 377, 378].

**Encoding and decoding.** *Source encoding* is the process of transforming the information produced by the source into messages. The source may produce a continuous stream of

symbols from the source alphabet. Then the source encoder cuts this stream into blocks of fixed size. The source decoder performs an inverse mapping and delivers symbols from the output alphabet.

*Channel encoding* allows transmission of the message generated by the source through the channel. The channel encoder accepts as input a set of messages of fixed length and maps the source alphabet into a channel alphabet, then adds a set of redundancy symbols, and finally sends the message through the channel. The *channel decoder* first determines if the message is in error and takes corrective actions; then it removes the redundancy symbols, maps the channel alphabet into the source alphabet, and hands each message to the source decoder. The *source decoder* processes the message and passes it to the receiver, Figure 74.

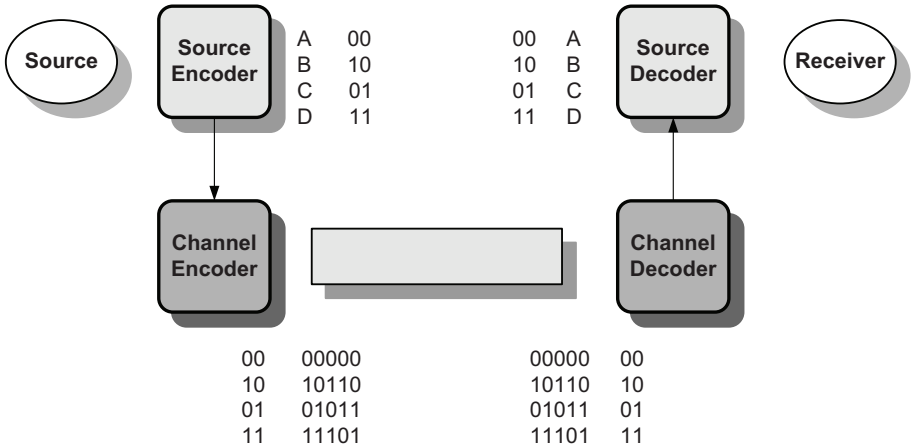



Figure 74: Source and channel encoding and decoding. The source alphabet consists of four symbols, A,B,C, and D; the source encoder maps each input symbol into a two-bit code, $A \mapsto 00$, $B \mapsto 10$, $C \mapsto 01$, and $D \mapsto 11$. The channel encoder maps a two-bit string into a five-bit string; if a one-bit or a two-bit error occur, the channel decoder receives a string that does not map to any of the valid five-bit strings and detects an error. The source decoder performs an inverse mapping and delivers a string consisting of the four output alphabet symbols.

Information encoding allows us to accomplish critical tasks for processing, transmission, and storing of information, such as:

(i) *Error control.* An error occurs when an input symbol is distorted during transmission and interpreted by the destination as another symbol from the alphabet. The error control mechanisms transform a noisy channel into a noiseless one; they are built into communication protocols to eliminate the effect of transmission errors.

(ii) *Data compression.* We wish to reduce the amount of data transmitted through a channel with either no or minimal effect on the ability to recognize the information produced by the source. In this case encoding and decoding are called *compression* and *decompression*, respectively.

(iii) *Support confidentiality.* We want to restrict access to information to only those who have the proper authorization. The discipline covering this facet of encoding is called *cryptography* and the processes of encoding/decoding are called *encryption/decryption*.

**Channel capacity.** In this section we consider an abstraction, the discrete memoryless channel, defined as the triplet $(X, Y, p_{Y|X}(y \mid x))$ with:

- The input channel alphabet $X$; the selection of an input symbol $x \in X$ is modelled by the random variable $X$ with the probability density function $p_X(x)$.
- The output channel alphabet $Y$; the selection of an output symbol $y \in Y$ is modelled by the random variable $Y$ with the probability density function $p_Y(y)$.
- The conditional probability density function $p(y \mid x), x \in X, y \in Y$, the probability of observing output symbol $y$ when the input symbol is $x$. The channel is memoryless if the probability distribution of the output depends only on the input at that time and not on the past history.

The capacity of a discrete memoryless channel $(X, Y, p_{Y|X}(y \mid x))$ with $p_X(x)$ the input distribution is defined as $C = \max_{p_X(x)} I(X;Y)$. In real life, various sources of noise distort transmission and lower the channel capacity; the capacity $C$ of a noisy channel is

$$C \;=\; B \times \log\left(1 + \frac{Signal}{Noise}\right)$$

where $B$ is the bandwidth, the maximum transmission rate, $Signal$ is the average power of the signal, and $Noise$ is the average noise power. The signal-to-noise ratio is usually expressed in decibels (dB) given by the formula: $10 \times \log_{10}(\frac{Signal}{Noise})$. A signal-to-noise ratio of $10^3$ corresponds to 30 dB, and one of $10^6$ corresponds to 60 dB.

Interestingly enough, we can reach the same conclusion regarding the capacity of a noisy communication channel based on simple physical arguments and the mathematical tools used in Section 3.2 for expressing Boltzmann's entropy. Consider the following gedanken experiment [335]: Alice wants to send $m$ bits of information and she knows that the probability of a bit being in error is $p$. Alice decides to pack the $m$ bits into a message of $n > m$ bits. When $n$ is large $pn$ bits in the string received by Bob could be in error and $(1-p)n$ bits may not be affected by errors. Alice is "clairvoyant" and knows ahead of time what bits will be affected by error and she uses $k$ bits to specify how the $pn$ errors are distributed among the $n = m + k$ bits of the message.

To specify $k$, the number of ways the errors are distributed we have to count the number of ways $pn$ bits in error and $(1-p)n$ bits that are not in error can be packed in a message of $n$ bits. This number is $k = \log \binom{n}{pn}$; we have encountered this expression in Section 3.2 and we know that when $n$ is large then the number of bits required to specify the number of ways the errors are distributed is

$$k \approx nH(p, 1-p) = nH(p) = -n[p \log p + (1-p)\log(1-p)].$$

Recall that in Section 3.3 we used the notation $H(p)$ rather than $H(p, 1-p)$ for binary information.

We conclude that in order to specify how the $pn$ errors are distributed among the $n$ bits of the encoded message, Alice needs $k$ bits. Never mind that this approach is unfeasible because Alice cannot possibly know the position of the random errors when she transmits the message,

but there is a lesson to be learned from this experiment. When Bob receives the message of $n$ bits and erases it, then, according to Landauer's principle the amount of heat/energy he generates is $n \times (k_B T \ln 2)$. A fraction of this heat/energy, namely $k \times (k_B T \ln 2)$, is generated while Bob attempts to extract the $k$ bits of the additional information revealing were the errors are placed in the garbled string of $n$ bits.

This additional energy is due to the fact that the communication channel is noisy and Alice had to add to the $m$ bits she wanted to transmit the information about the errors. In other words, *a noisy channel prevents us from utilizing the full channel capacity, we have to spend an additional amount of energy to erase the additional information used to encode the message.* To locate the errors we need at least $k = nH(p)$ bits thus, the maximum capacity of the noisy channel is the fraction of the channel capacity used to transmit useful information $m = n - k = n(1 - H(p))$.

**Example.** *Maximum data rate over a phone line.* A phone line allows transmission of frequencies in the range 500 Hz to 4000 Hz and has a signal-to-noise ratio of 30 dB. The maximum data rate through the phone line is

$$C \;=\; (4000 - 500) \;\times\; \log(1 + 1000) \;\approx 3500 \times \log 2^{10} = \;35 \text{ kbps}.$$

If the signal-to-noise ratio improves to 60 dB, the maximum data rate doubles to about 70 kbps. However, improving the signal-to-noise ratio of a phone line by three orders of magnitude, from $10^3$ to $10^6$, is extremely difficult.

The channel coding theorem: given a noisy channel with capacity $C$ and given $0 < \epsilon < 1$, there is a coding scheme that allows us to transmit information through the channel at a rate arbitrarily close to channel capacity, $C$, with a probability of error less than $\epsilon$. This result constitutes what mathematicians call a theorem of existence; it only states that a solution for transforming a noisy communication channel into a noiseless one exists, without giving a hint of how to achieve this result.

**Simple models of classical memoryless channels.** A binary channel is one where the input and the output alphabet consists of binary symbols: $X = \{0, 1\}$ and $Y = \{0, 1\}$. A *unidirectional binary communication channel* is one where the information propagates in one direction only, from the source to the destination.

A noiseless binary channel transmits each symbol in the input alphabet without errors, as shown in Figure 75(a); The noiseless channel model is suitable in some cases for performance analysis, but it is not useful for reliability analysis when transmission errors have to be accounted for.

The noisy binary symmetric channelis a model based on the assumption that an input symbol is affected by error with probability $p > 0$; a 1 at the input becomes a 0 at the output and a 0 at the input becomes a 1 at the output, as shown in Figure 75(b).

Assume that the two input symbols occur with probabilities $\text{Prob}(X = 0) = q$ and $\text{Prob}(X = 1) = 1 - q$. In this case,

$$H(Y \mid X) \;=\; \sum_{x \in X} \text{Prob}(X = x) \times H(Y \mid X = x)$$

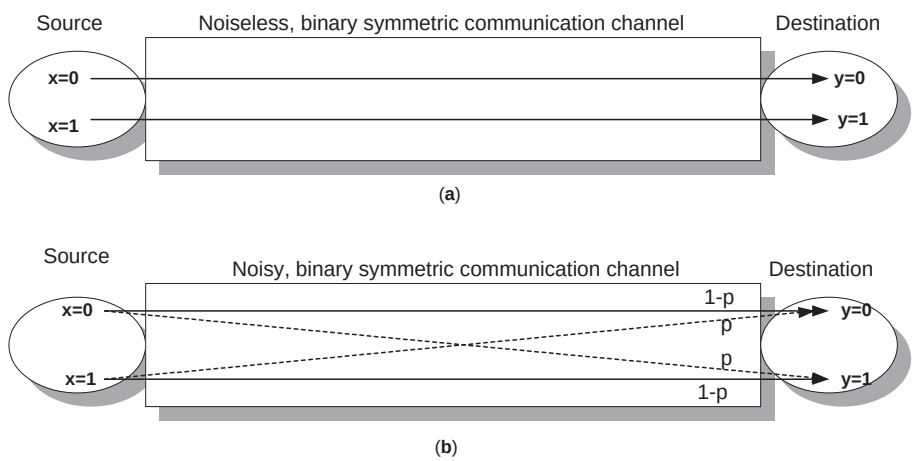



Figure 75: (a) A noiseless binary symmetric channel maps a 0 at the input into a 0 at the output and a 1 into a 1. (b) A noisy symmetric channel maps a 0 into a 1, and a 1 into a 0 with probability $p$. An input symbol is mapped into itself with probability $1-p$.

$$H(Y \mid X) \;=\; -\{q\;[p\;\log p + (1-p)\log(1-p)] + (1-q)\;[p\log p + (1-p)\log(1-p)]\}$$

$$H(Y \mid X) \;=\; -\,[p\;\log p + (1-p)\;\log(1-p)].$$

Then the mutual information is

$$I(X;Y) = H(Y) - H(Y \mid X) = H(Y) + [p\;\log p + (1-p)\;\log(1-p)].$$

We can maximize $I(X;Y)$ over $q$, by making $H(Y) = 1$ and get the channel capacity per symbol in the input alphabet

$$C_s \;=\; 1 + [p\;\log p + (1-p)\;\log(1-p)].$$

When $p = \frac{1}{2}$ the channel capacity is 0 because the output is independent of the input. When $p = 0$ or $p = 1$, the capacity is 1, and we have in fact a noiseless channel.

The next model, the <u>binary erasure (BER) channel</u> captures the following behavior: the transmitter sends a bit and the receiver receives with probability $1-p_e$ the bit unaffected by error, or it receives with probability $p_e$ a message that the bit was "erased", where $p_e$ is the "erasure probability." A binary erasure channel with input $X$ and output $Y$ is characterized by the following conditional probabilities:

$$\begin{array}{rclcl}
\text{Prob}\,(Y=0 \mid X=0) & = & \text{Prob}\,(Y=1 \mid X=1) & = & 1-p_e \\
\text{Prob}\,(Y=e \mid X=0) & = & \text{Prob}\,(Y=e \mid X=1) & = & p_e \\
\text{Prob}\,(Y=0 \mid X=1) & = & \text{Prob}\,(Y=1 \mid X=0) & = & 0.
\end{array}$$

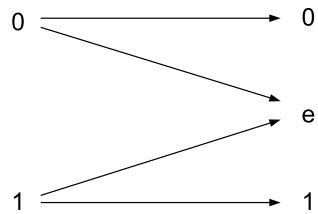



Figure 76: A binary erasure channel. The transmitter sends a bit and the receiver receives with probability $1-p_e$ the bit unaffected by error, or it receives with probability $p_e$ a message $e$ that the bit was "erased." $p_e$ is the "erasure probability."

The binary erasure channel is in some sense error-free, the sender does not have to encode the information and the receiver does not have to invest any effort to decode it, Figure 76. $C_{BEC}$, the capacity of a binary erasure channel is

$$C_{BEC} = 1 - p_e.$$

Indeed,

$$C_{BEC} = \max I(X;Y) = \max_{p_X(x)} (H(Y) - H(Y \mid X)) = \max_{p_X(x)} (H(Y)) - H(p_e).$$

If $E$ denotes the erasure event, $Y = e$, then

$$H(Y) = H(Y, E) = H(E) + H(Y \mid E).$$

If the probability $\text{Prob}\,(X = 1) = p_1$ then:

$$H(Y) = H\left((1-p_1)(1-p_e), p_e, p_1(1-p_e)\right) = H(p_e) + (1-p_e)H(p_1).$$

It follows that

$$C_{BEC} = \max_{p_X(x)} (H(Y) - H(p_e)) = \max_{p_1}(1-p_e)H(p_1) = (1-p_e).$$

The following argument provides an intuitive justification for the expression of the binary erasure channel capacity: assume that the sender receives instantaneous feedback when the bit was erased and then retransmits the bit in error. This would result in a average transmission rate of $(1-p_e)$ bits/sec. Of course this ideal scenario is not feasible thus $(1-p_e)$ is an upper bound on the channel rate.

This brings us to the role of the feedback; it can be proved that the rate $(1-p_e)$ is the best that can be achieved with, or without feedback [107]. It is surprising that the feedback does not increase the capacity of discrete memoryless channels.

Another model, the q-ary erasure channel assumes a channel with an alphabet of $q$ symbols; the transmitter sends a symbol and the receiver receives with probability $1-p_e$ the symbol unaffected by error, or it receives with probability $p_e$ a message that the bit was "erased." It is easy to show that the channel capacity in this case is $C_{QEC} = (1-p_e)\log q$.

If the input symbols have the probabilities $p_X(a_i) = p_i,\ 1 \le i \le q$ then the realizations of $Y$ could be any of the $q$ input symbols which occur with probabilities $p_i(1-p_e)$, respectively, or the erasure $E$ which occurs with probability $p_e$. Then

$$H(Y) = -\left[\sum_{i=1}^{q} p_i(1-q)\log p_i(1-q)\right] - p_e \log p_e$$

The output entropy is maximized when $p_i = 1/q$; then

$$\max H(Y) = -(1-p_e)\log(1-p_e) + (1-p_e)\log q - p_e \log p_e.$$

As before, $H(Y|X) = H(p_e) = -p_e \log p_e - (1-p_e)\log(1-p_e)$ thus, $C_{QEC} = (1-p_e)\log q$.

Lastly, the q-ary symmetric channel model assumes that the probability that a symbol is correctly transmitted is equal to $1-p$ and when an error occurs, the symbol is replaced by one of the other $q-1$ symbols with equal probability. For the following analysis we make the assumption that *errors are introduced by the channel at random; the probability of an error in one coordinate is independent of errors in adjacent coordinates*; the probability that the received $n$-tuple $r$ is decoded as codeword $c$ at the output of a q-ary symmetric channel is

$$\text{Prob}(r,c) = (1-p)^{n-d}\left(\frac{p}{q-1}\right)^d.$$

where $d(r,c) = d$, $d$ is the distance of the code and $n$ is the length of a codeword. Indeed, $n-d$ coordinate positions in $c$ are not altered by the channel; the probability of this event is $(1-p)^{n-d}$. In each of the remaining $d$ coordinate positions, a symbol in $c$ is altered and transformed into the corresponding symbol in $r$; the probability of this is $(p/(q-1))^d$.

The next section allows us to quantify the degree of similarity between two probability density functions, a subject important for classical, as well as, quantum information theory.

## 3.8 Trace Distance and Fidelity

Information theory is concerned with statistical properties of random variables and often demands an answer to the question how similar two probability density functions are. For example, we have two information sources $S_1$ and $S_2$ which share the same alphabet, but have different input and output probability density functions, $p_X(x)$ and $p_Y(x)$, and we ask ourselves how similar are the two sources. Another question we may pose is how well a communication channel preserves information over time. Let $I$ be the random variable describing the input of a noisy channel and $O$ its output. Let $\tilde{I}$ by a copy of $I$. Now we are interested in the relationships of the joint probability distributions $p_{I,\tilde{I}}(x_i, x_i)$ and $p_{I,O}(x_i, y_i)$.

We introduce two analytical measures of the similiarity/dissimilarity of two probability density functions, $p_X(x)$ and $p_Y(x)$. Whenever there is no ambiguity we omit the argument of the density function and write $p_X$ instead of $p_X(x)$.

The trace distance, $D(p_X(x), p_Y(x))$, (also called Kolmogorov, or L1 distance) of two probability density functions, $p_X(x)$ and $p_Y(x)$, is defined as

$$D(p_X(x), p_Y(x)) = \frac{1}{2}\sum_x \mid p_X(x) - p_Y(x) \mid .$$

The fidelity, $F(p_X(x), p_Y(x))$ of two probability density functions, $p_X(x)$ and $p_Y(x)$ is defined as

$$F(p_X(x), p_Y(x)) = \sum_x \sqrt{p_X(x)p_Y(x)}.$$

It is easy to see that the fidelity is the inner product of unitary vectors with components equal to $\sqrt{p_X(x)}$ and $\sqrt{p_Y(x)}$. These vectors connect the center with points on a sphere of radius one; indeed, $\sum_x \sqrt{p_X(x)}^2 = 1$ and $\sum_x \sqrt{p_Y(x)}^2 = 1$. A geometric interpretation of fidelity is, thus, the cosine of $\theta$, the angle between the two vectors,

$$F(p_X(x), p_Y(x)) = \cos\theta.$$

**Example.** *Two noisy binary symmetric channels* with input and output alphabets $\{0, 1\}$ have probability of error $p$ and $q$, respectively. This means that when we transmit one of the input symbols the probability of recovering it correctly is $1 - p$ for the first channel and $1 - q$ for the second. The two measures of similarity between the two distributions are:

$$D(p_X(x), p_Y(x)) = \frac{1}{2}[\mid p - q \mid + \mid (1 - p) - (1 - q) \mid] = \mid p - q \mid$$

and

$$F(p_X(x), p_Y(x)) = \sqrt{pq} + \sqrt{(1 - p)(1 - q)} = \sqrt{pq} + \sqrt{pq + 1 - (p + q)}.$$

If $p = 0.4$ and $q = 0.9$ then

$$D(p_X(x), p_Y(x)) = \mid p - q \mid = 0.5, \quad F(p_X(x), p_Y(x)) = \sqrt{pq} + \sqrt{pq + 1 - (p + q)} = 0.6245.$$

*The trace distance is a metric while the fidelity is not.* Indeed, the trace distance satisfies the properties of a metric:

| | |
|---|---|
| $D(p_X, p_Y) \geq 0$ | non-negativity |
| $D(p_X, p_Y) = 0$ iff $p_X = p_Y$ | identity of indiscernibles |
| $D(p_X, p_Y) = D(p_Y, p_X)$ | symmetry |
| $D(p_X, p_Z) \leq D(p_X, p_Y) + D(p_Y, p_Z)$ | triangle inequality. |

The fidelity fails to satisfy one of the required properties:

$$F(p_X(x), p_X(x)) = \sum_x \sqrt{p_X(x)p_X(x)} = \sum_x p_X(x) = 1 \neq 0.$$

Next we turn our attention to quantum information and discuss the characterization of a source of quantum information using the von Neumann entropy.

## 3.9 von Neumann Entropy

Pure states of quantum systems correspond to classical states we are familiar with, they are the really physical states. In the standard Quantum Mechanics the pure states of a purely quantum physical system with separable complex Hilbert space $\mathcal{H}$ are represented by the rays of this Hilbert space or by rank one operators. Normalized linear superpositions of pure states are also pure states. When the preparation process cannot be rigourously controlled the resulting mixed state can be considered a superposition of pure states with a given probability distribution and is represented by a rank one statistical operator, the density matrix. Mixed states do not have a classical correspondent.

John von Neumann realized that Quantum Mechanics requires a definition of entropy [441] covering pure, as well as, mixed states; he associated the entropy to a statistical operator, the density matrix[28], that he used to characterize quantum systems in a mixed state, as well as composite systems. He presented the mathematical formalism for Quantum Mechanics in a book published in 1932 [442].

A system is in a *pure state* when its density matrix contains only one term; all entries of the diagonalized density matrix of a pure state are zero with the exception of a single term equal to one on the main diagonal. Moreover, if the state is pure, then the trace of the square of the density matrix is equal to one. In this case there is no lack of knowledge on the preparation of the system, the preparation generates the desired state with certainty.

The von Neumann entropy is a function of the density matrix $\rho$ of the quantum system

$$S(\rho) = -\mathrm{tr}\left[\rho \log \rho\right].$$

In this expression $(\log \rho)$ is the Hermitian operator that has precisely the same eigenvectors as $\rho$; we simply take the logarithm base two of the corresponding eigenvalues. If we consider $0 \log 0 = 0$ then the expression $\rho \log \rho$ may be defined for all positive semidefinite $\rho$, though technically $(\rho \log \rho)$ is only defined for positive definite $\rho$.

The expression of von Neumann entropy requires the evaluation of a function of a matrix[29]. The function can be evaluated as an ordinary function of the eigenvalues; the eigenvalues, $p_i$, are invariant under a change of basis, thus, the function itself is invariant. The eigenvalues of the density operator are probabilities, therefore we use the notation $p_i$ rather than $\lambda_i$ for the eigenvalues of the density matrix; $\mathrm{tr}\left[\rho\right] = 1$ reflects the fact that $\sum_i p_i = 1$.

It is easy to see that when the states $\mid e_i\rangle \in \mathcal{H}^{\otimes n}$ are orthonormal and the spectral decomposition of $\rho$ is

$$\rho = \sum_{i=1}^{2^n} p_i \mid e_i\rangle\langle e_i \mid$$

then,

[28]Some attribute the introduction of the density matrix to John von Neumann and believe that Lev Landau and Felix Bloch have introduced it independently in 1927.

[29]The logarithm of a matrix is another matrix, such that if $B$ is the logarithm of a given matrix $A$, then $e^B = A$. Not all matrices have a logarithm, but those that do, may have more than one logarithm. A complex matrix has a logarithm if and only if it is invertible. If the matrix has only nonnegative real eigenvalues, then it has a unique logarithm. The eigenvalues of this unique logarithm lie all in the domain $\{z \in \mathbb{C} \mid -\pi < \Im(z) < \pi\}$. For any invertible matrix $A$ $A^{-1} = e^{-\ln A}$.

$$S(\rho) = -\mathrm{tr}\left[\sum_{i=1}^{2^n} p_i \mid e_i\rangle\langle e_i \mid \sum_{j=1}^{2^n} \log p_j \mid e_j\rangle\langle e_j \mid\right] = -\sum_{i=1}^{2^n} p_i \log p_i = H(p_1, p_2, \ldots, p_n).$$

In this case von Neumann entropy is equal to $H(p_1, p_2, \ldots, p_n)$, Shannon entropy we have encountered earlier. If we prepare the system in a *maximally mixed state*, a state when all $N$ pure states are equally likely, then the entropy is equal to $\log N$ in agreement with Boltzmann and Shannon entropy.

If $\mid \varphi\rangle$ is a pure state then its density matrix $\rho$ has only one eigenvalue equal to 1, its trace is equal to 1, then $\log \rho = 0$ and its von Neumann entropy is zero.

Let us now consider the ensemble of pure states $\mid e_i\rangle$, $1 \leq i \leq N$ and prepare a mixed state where each pure state $\mid e_i\rangle$ has probability $p_i$. To do so we have first to create a list and use $H(p_1, p_2, \ldots, p_N)$ bits/letter to specify the probability of each one of the $N$ states. After the creation of the mixed state we have to erase this information and, according to Landauer principle, at temperature $T$ we will generate $k_B T H(p_1, p_2, \ldots, p_N)$ units of heat per erased letter. This explains why *Shannon entropy is larger or equal to von Neumann entropy.*

Recall that we cannot distinguish among the different preparations leading to the same density matrix $\rho$. *If $\rho$ is the density matrix of a mixed state, then $S(\rho)$, the von Neumann entropy, is the smallest amount of information that has to be invested to create the mixed state characterized by $\rho$, and also the minimum amount of classical information that we can access from the system in that state.*

We have a bipartite system consisting of two quantum registers, $A$ and $B$, owned by Alice and Bob, respectively and let $\mid \varphi\rangle$ be a pure state in the Hilbert space shared by Alice and Bob, $\mid \varphi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$; it is easy to see that the reduced states have the same entropy

$$S\left(\mathrm{tr}_B\left[\mid \varphi\rangle\langle\varphi \mid\right]\right) = S\left(\mathrm{tr}_A\left[\mid \varphi\rangle\langle\varphi \mid\right]\right).$$

When von Neumann associated a statistical operator $\rho$ to entropy [441], he first assumed that $S(\rho)$ is a continuous function of $\rho$. Then he established a mix-in property for orthogonal pure states

$$S\left(\sum_i p_i \mid \varphi_i\rangle\langle\varphi_i \mid\right) = \sum_i p_i S\left(\mid \varphi_i\rangle\langle\varphi_i \mid\right) - k_B \sum_i p_i \log p_i.$$

Finally, he showed that $S(\mid \varphi\rangle\langle\varphi \mid)$ is independent of the state $\mid \varphi\rangle$, thus if we normalize and set the additive constant equal to zero we have:

$$S\left(\sum_i p_i \mid \varphi_i\rangle\langle\varphi_i \mid\right) = -k_B \sum_i p_i \log p_i.$$

**The mix-in property.** von Neumann considered a quantum system in a mixed state

$$\mid \psi\rangle = p \mid \varphi_1\rangle + (1-p) \mid \varphi_2\rangle.$$

He imagined a gedanken experiment [330] and constructed a thermodynamical model of the quantum system in a mixed state, $D$. In this model $N$ molecules of gas are mixed together, $pN$ molecules of gas are in state $\varphi_1$ described by the statistical operator $D_1$ and $(1-p)N$ molecules are in state $\varphi_2$ described by $D_2$. The two states, $D_1$ and $D_2$, are thermodynamically disjoint in the following sense: there is a wall completely permeable for the first gas, the one in state $\varphi_1$, but totally impermeable to the molecules of the second gas, the one in state $\varphi_2$. This condition translates into orthogonality of the eigenvectors corresponding to the nonzero eigenvalues of the density matrices of the quantum systems modelled by the gas.

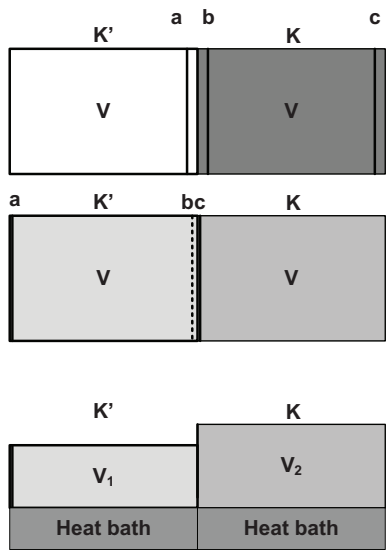



Figure 77: A gedanken experiment to show the mix-in property of von Neumann entropy.

The gedanken experiment imagined by von Neumann is conducted as follows:

1. Initially the gas in the mixed state $D$ is in a cubic recipient $K$ of volume $V$. The entropy of the gas is $S_{gas}^{(1,2)}$ and it is proportional with the number of molecules, $N$; its density is $\frac{N}{V}$ .

2. We attach a new cubic recipient $K'$ of an identical volume $V$ to the left of $K$.

3. We replace the wall separating $K$ and $K'$ by two walls $a$ and $b$ and add a third wall $c$ to the right of $b$, as in Figure 77 (top). Wall $a$ is impermeable for both types of molecules, wall $b$ is permeable for molecules of type $\varphi_1$ and impermeable for molecules of type $\varphi_2$, and wall $c$ is impermeable for molecules of type $\varphi_1$ and permeable for molecules of type $\varphi_2$.

4. We push walls $a$ and $c$ to the left maintaining the distance between them. As a result the $\varphi_1$-molecules are pressed through wall $b$ into $K'$ and $\varphi_2$-molecules are left in $K$, Figure 77 (center).

5. We replace walls $b$ and $c$ with impermeable walls and remove wall $a$. The $\varphi_1$ and $\varphi_2$ molecules have been separated without doing work, changing temperature or producing

heat. Then we compress both the gas in $K'$ from a volume $V$ to a volume $V_1 = pV$ and the gas in $K$ to a volume $V_2 = (1-p)V$. The densities of the two gasses are $pN/V_1$ and $(1-p)N/V_2$, respectively; they are equal to the density of gas $D$. The compression is isothermal; the two volumes absorb heat from two heat reservoirs, as shown in Figure 77 (bottom), and maintain their temperature $T$. The first gas absorbs a quantity of heat equal to $k_B T(pN)\log p$ and the second gas absorbs a quantity of heat equal to $k_B T(1-p)N\log(1-p))$. The entropy of the gasses changes as follows: the entropy of the first gas decreases with the amount $S^{(1)}_{gas}(pN) = k_B(pN)\log p$, while the entropy of the second gas decreases with the amount $S^{(2)}_{gas}((1-p)N) = k_B(1-p)N\log(1-p)$.

6. We mix the two gasses of equal density and find the gas $D$ with $N$ molecules in a volume $V$; we note that

$$S^{(1)}_{gas}(pN) + S^{(2)}_{gas}((1-p)N) = S^{(1,2)}_{gas}(N) + k_B(pN)\log p + k_B(1-p)N\log(1-p).$$

We assume that the entropy of a gas is proportional to the number of molecules in the gas and if we divide the previous expression by $N$ we obtain

$$pS^{(1)}_{gas} + (1-p)S^{(2)}_{gas} = S^{(1,2)}_{gas} + k_B\left[p\log p + (1-p)\log(1-p)\right]$$

or:

$$S^{(1,2)}_{gas} = \left[pS^{(1)}_{gas} + (1-p)S^{(2)}_{gas}\right] - k_B\left[(p)\log p + (1-p)\log(1-p)\right].$$

The result can be extended from a two-component mixture to an infinite mixture:

$$S\left(\sum_i p_i \mid \varphi_i\rangle\langle\varphi_i \mid\right) = \sum_i p_i S\left(\mid \varphi_i\rangle\langle\varphi_i \mid\right) - k_B\sum_i p_i \log p_i.$$

The proof that $S(\mid \varphi\rangle\langle\varphi \mid)$ is independent of the state $\mid \varphi\rangle$ if the state obeys Schrödinger equation, is slightly more complex. Any two states $\mid \varphi_i\rangle$ and $\mid \varphi_j\rangle$ in the evolution of the system are related by a time-independent energy operator $H$:

$$\mid \varphi_j\rangle = e^{iHt} \mid \varphi_i\rangle.$$

The operator $e^{iHt}$ is unitary and the eigenvalues of the density matrix at time $t$, $\rho(t)$ are the same as the eigenvalues of the density matrix at time 0, $\rho(0)$. The density matrix at time $t$ is obtained as

$$\rho(t) = e^{-iHt}\rho e^{iHt}.$$

We observe that $S[\rho(t)] = S[\rho]$ for a closed system with a time-independent Hamiltonian because the expression of entropy involves only the eigenvalues of the density matrix. The reality is closer to the second law of thermodynamics which states that the entropy of a closed system never decreases, it can remain constant or increases.

John von Neumann proved that

$$S\left(\mid \varphi_i\rangle\langle\varphi_i \mid\right) \leq S\left(\mid \varphi_j\rangle\langle\varphi_j \mid\right)$$

by constructing a large number of measurements of the system in state $\mid \varphi_i\rangle$ such that the state after all the measurements becomes an ensemble which is arbitrarily close to $\mid \varphi_j\rangle\langle\varphi_j \mid$. Recall from Section 2.7 that a von Neumann projective measurement of the observable $\mathcal{O} = \sum_i p_i \mid e_i\rangle\langle e_i \mid$ of a system in state $\mid \varphi\rangle$ projects this state on the basis vectors $\mid e_i\rangle$ and produces pure states $\mid e_i\rangle\langle e_i \mid$ with weights $\langle\varphi \mid e_i\rangle$, respectively, when the eigenvalues $p_i$ of the operator $\mathcal{O}$ are distinct.

We now discuss how to erase classical information from quantum states. As we have mentioned earlier, the von Neumann entropy and, thus, the information content of a pure state is equal to zero. Therefore, *to erase information from a mixed state we have to return it to a pure state.*

Let $\rho = \sum_i p_i \mid e_i\rangle\langle e_i \mid$ be the density matrix of a mixed state with $\mid e_i\rangle$ the energy eigenstates of the system. We perform measurements in the energy eigenbasis; after each measurement, the system ends up in a pure state $\mid e_i\rangle$ and the corresponding eigenvalue $p_i$ represents a classical record of the outcome of our measurement. To complete the erasure of the quantum information we have to erase the classical information as well; in this process we generate $k_B T \ln 2$ units of heat per bit. In our example the amount of heat is $k_B T \ln 2 S(\rho) < k_B T \ln 2 H(p)$. *When the quantum measurements are conducted in the basis constructed from the eigenstates of the density matrix $\rho$, then the von Neumann entropy is equal to Shannon entropy and we have an optimal erasure procedure.*

We conclude this section with the statement of two basic properties of von Neumann entropy. The proofs or references to the proofs can be found in the seminal paper of Wehrl [446].

*Additivity.* Let $\rho_1, \rho_2, \ldots \rho_n$ be the density matrices of $n$ quantum system. Then,

$$S\left(\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n\right) = S\left(\rho_1\right) + S\left(\rho_2\right) + \ldots + S\left(\rho_n\right).$$

*Concavity.* Let $\rho_1, \rho_2, \ldots \rho_n$ be the density matrices of $n$ quantum systems and $p_1, p_2, \ldots p_n$ nonnegative integers, $p_i \geq 0$ such that $\sum_i p_i = 1$. Then,

$$S\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i S\left(\rho_i\right).$$

The concavity of the entropy is the result of the concavity of the logarithm function. The concavity of von Neumann entropy indicates that when ensembles are fitted together into a composite system in a mixed state, one loses information about which ensemble a particular state arises from; thus, the entropy increases.

In summary, von Neumann used thermodynamic arguments to define the entropy of a quantum system with the density matrix $\rho$ as $S(\rho) = -\mathrm{tr}\left[\rho \log \rho\right]$; the expression $\rho \log \rho$ is well defined as the density matrix is positive semidefinite. It is easy to see that von Neumann entropy is the Shannon entropy of the eigenvalues $p_i$ of a diagonal density matrix $\rho$. Indeed, when $\rho$ is a diagonal matrix with eigenvalues $p_i$ then $\mathrm{tr}(\rho) = \sum_i p_i = 1$ and $\rho \log \rho$ is also diagonal with eigenvalues $p_i \log p_i$.

## 3.10 Joint, Conditional, and Relative von Neumann Entropy

The joint, conditional, and relative von Neumann entropy are concepts similar to those defined by classical information theory. Consider a pair of quantum systems, $\mathcal{A}$ and $\mathcal{B}$ and the composite system $(\mathcal{A}, \mathcal{B})$; let $\rho^{\mathcal{A}}$, $\rho^{\mathcal{B}}$, and $\rho^{(\mathcal{A},\mathcal{B})}$ be the density operators of $\mathcal{A}$, $\mathcal{B}$, and the composite system $(\mathcal{A}, \mathcal{B})$, respectively. The density operators $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ are given by the partial traces of the joint density matrix:

$$\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}} \left[\rho^{(\mathcal{A},\mathcal{B})}\right] \quad \text{and} \quad \rho^{\mathcal{B}} = \mathrm{tr}_{\mathcal{A}} \left[\rho^{(\mathcal{A},\mathcal{B})}\right].$$

The <u>joint von Neumann entropy</u> of the pair of quantum systems $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$S(\mathcal{A}, \mathcal{B}) = S(\rho^{(\mathcal{A},\mathcal{B})}).$$

The <u>conditional von Neumann entropy</u> of system $\mathcal{A}$ conditioned by system $\mathcal{B}$ is defined as

$$S(\mathcal{A} \mid \mathcal{B}) = S(\mathcal{A}, \mathcal{B}) - S(\mathcal{B}) = S(\rho^{(\mathcal{A},\mathcal{B})}) - S(\rho^{\mathcal{B}}).$$

The conditional von Neumann entropy $S(\mathcal{A} \mid \mathcal{B})$ is concave [446].
The <u>mutual information</u> of the pair of systems ($\mathcal{A}$ and $\mathcal{B}$) is defined as

$$I(\mathcal{A}, \mathcal{B}) = S(\mathcal{A}) + S(\mathcal{B}) - S(\mathcal{A}, \mathcal{B}) = S(\rho^{\mathcal{A}}) + S(\rho^{\mathcal{B}}) - S(\rho^{(\mathcal{A},\mathcal{B})}).$$

The <u>relative entropy</u> $S(\rho_1 \mid\mid \rho_2)$ of two states described by the density matrices $\rho_1$ and $\rho_2$ is defined as

$$S(\rho_1 \mid\mid \rho_2) = \mathrm{tr}[\rho_2(\log \rho_2 - \log \rho_1)].$$

The two density matrices $\rho_1$ and $\rho_2$ have the spectral decompositions

$$\rho_1 = \sum_j q_j \mid u_j\rangle\langle u_j \mid \quad \text{and} \quad \rho_2 = \sum_i p_i \mid e_i\rangle\langle e_i \mid$$

where $\{\mid e_i\rangle\}$ and $\{\mid u_j\rangle\}$ are two orthonormal bases for the same physical system.

It is easy to see that the quantum relative entropy is non-negative, $S(\rho_1 \mid\mid \rho_2 \geq 0$, more precisely,

$$S(\rho_1 || \rho_2) = \begin{cases} > 0 & \text{if} \quad \rho_1 \neq \rho_2 \\ = 0 & \text{if} \quad \rho_1 = \rho_2 \end{cases}$$

Then, the relative entropy can be evaluated in the basis where $\rho_2$ is diagonal

$$\begin{aligned}
\mathrm{tr}\left[\rho_2 \log \rho_2\right] - \mathrm{tr}\left[\rho_2 \log \rho_1\right] &= \textstyle\sum_i \langle e_i \mid \rho_2 \log(\rho_2) \mid e_i\rangle - \sum_i \langle u_i \mid \rho_2 \log(\rho_1) \mid u_i\rangle \\
&= \textstyle\sum_i p_i \left(\log p_i - \sum_j \mid \langle u_j \mid e_i\rangle \mid^2 \log q_j\right) \\
&= -\tfrac{1}{\ln 2} \textstyle\sum_i \sum_j p_i \mid \langle u_j \mid e_i\rangle \mid^2 \ln \tfrac{q_j}{p_i} \\
&\geq -\tfrac{1}{\ln 2} \textstyle\sum_i \sum_j p_i \mid \langle u_j \mid e_i\rangle \mid^2 \left(\tfrac{q_j}{p_i} - 1\right) = 0.
\end{aligned}$$

This inequality is strict unless $q_j = p_i,\ \forall i, j$ satisfying $\mid \langle u_j \mid e_i \rangle \mid^2 > 0$. Under this assumption the equality occurs when:

$$\sum_i \sum_j \mid \langle u_j \mid e_i \rangle \mid^2 (p_i - q_j)^2 = 0.$$

This implies that $\| \rho_2 - \rho_1 \| = 0$ or $\rho_2 = \rho_1$.

The subadditivity of von Neumann entropy requires that

$$S(\rho^{(\mathcal{A},\mathcal{B})}) \leq S(\rho^{\mathcal{A}}) + S(\rho^{\mathcal{B}}).$$

This inequality means that a pair of correlated system contains more information than the two systems considered separately.
Given three quantum systems $\mathcal{A}, \mathcal{B}$, and $\mathcal{C}$ the strong subadditivity of von Neumann entropy, [446] requires that

$$S(\rho^{(\mathcal{A},\mathcal{B},\mathcal{C})}) + S(\rho^{\mathcal{A}}) \leq S(\rho^{(\mathcal{A},\mathcal{B})}) + S(\rho^{(\mathcal{A},\mathcal{C})}).$$

A consequence of the strong subadditivity is that when $\mid \varphi_i^{\mathcal{A}} \rangle$ is a pure state of the system $\mathcal{A}$, and $\rho^{(\mathcal{B},\mathcal{C})}$ describes an arbitrary joint state of the systems $\mathcal{B}$ and $\mathcal{C}$ then,

$$\rho^{(\mathcal{A},\mathcal{B},\mathcal{C})} = \mid \varphi_i^{\mathcal{A}} \rangle \langle \varphi_i^{\mathcal{A}} \mid \otimes \rho^{(\mathcal{B},\mathcal{C})}.$$

This property will be used in Section 3.12 to establish an upper bound on the mutual information between a source of classical information and the recipient of it, when information is transmitted through a quantum channel.

## 3.11 Trace Distance and Fidelity of Mixed Quantum States

The outcome of the measurement of an observable $A$ is a random variable, the eigenvalue $x_i$ of the projector $\mathbf{M}$ applied to the system in state $\mid \psi \rangle$. If we have $N$ copies of the system and measure the same observable, then the outcomes of the measurement satisfy the equation $\mathbf{M} \mid \psi \rangle = x_i \mid \psi \rangle$ for $i = 1, \ldots, N$. The expected value of the outcome of our measurements is

$$\langle \mathbf{M} \rangle = \langle \psi \mid \mathbf{M} \mid \psi \rangle.$$

Let $p_X(x)$ be the discrete probability density function of $x_i$, the outcomes of the measurements. Let us assume that the quantum state $\mid \psi \rangle$ represents the state of a control register; we measure the same observable before and after the transformation performed by a quantum circuit and we are interested how similar are the two probability density functions $p_X(x)$ and $p_Y(x)$, respectively, of the outcomes of the measurements. Obviously, $\sum_i p_X(x_i) = 1$ and $\sum_i p_Y(x_i) = 1$. We can use the two measures, the fidelity and the trace distance, to assess the effect of the transformation. The two measures introduced in this section for probability density functions can be extended to linear operators in a Hilbert space $\mathcal{H}_2^{\otimes n}$.

**Trace distance.** The concept of trace distance discussed in Section 3.8 can be extended to quantum states. Consider two quantum systems $\mathcal{A}$ and $\mathcal{B}$ and let their states be described by density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$, respectively.

The trace distance between quantum systems $\mathcal{A}$ and $\mathcal{B}$ in states with density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ is defined as

$$D(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \frac{1}{2}\text{tr}\left[\mid \rho^{\mathcal{A}} - \rho^{\mathcal{B}} \mid\right].$$

If $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ commute and if $\lambda_i^{\mathcal{A}}$ and $\lambda_i^{\mathcal{B}}$, $1 \leq i \leq 2^n$ are the eigenvalues of $\rho^A$ and $\rho^B$ then,

$$D(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \frac{1}{2}\sum_{i=1}^{2^n} \mid \lambda_i^{\mathcal{A}} - \lambda_i^{\mathcal{B}} \mid .$$

**Example.** *The trace distance of two qubits is equal to the distance on the Bloch sphere between the two states.*

Consider two qubits and the corresponding vectors on a Bloch sphere, $\bar{r}^{\mathcal{A}} = (r_x^{\mathcal{A}}\ r_y^{\mathcal{A}}\ r_z^{\mathcal{A}})$ and $\bar{r}^{\mathcal{B}} = (r_x^{\mathcal{B}}\ r_y^{\mathcal{B}}\ r_z^{\mathcal{B}})$. If $\bar{\sigma} = (\sigma_x\ \sigma_y\ \sigma_z)$ the two density matrices are:

$$\rho^{\mathcal{A}} = \frac{1}{2}(I + \bar{r}^{\mathcal{A}} \times \bar{\sigma}) \qquad \text{and} \qquad \rho^{\mathcal{B}} = \frac{1}{2}(I + \bar{r}^{\mathcal{B}} \times \bar{\sigma}).$$

Thus,

$$D(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \frac{1}{2}\text{tr}\left[\mid \rho^{\mathcal{A}} - \rho^{\mathcal{B}} \mid\right] \ = \frac{1}{4}\text{tr}\left[(\bar{r}^{\mathcal{A}} - \bar{r}^{\mathcal{B}}) \times \bar{\sigma}\right].$$

It is easy to see that the eigenvalues of $(\bar{r}^{\mathcal{A}} - \bar{r}^{\mathcal{A}}) \times \bar{\sigma}$ are $\lambda_{1,2} = \pm \mid \bar{r}^{\mathcal{A}} - \bar{r}^{\mathcal{B}} \mid$, the distance on the Bloch sphere between the two states. Let

$$R = (\bar{r}^{\mathcal{A}} - \bar{r}^{\mathcal{B}}) \times \bar{\sigma} = \begin{pmatrix} r_z^{\mathcal{A}} - r_z^{\mathcal{B}} & r_x^{A} - r_x^{\mathcal{B}} - i(r_y^{\mathcal{A}} - r_y^{\mathcal{B}}) \\ r_x^{\mathcal{A}} - r_x^{\mathcal{B}} + i(r_y^{\mathcal{A}} - r_y^{\mathcal{B}}) & -(r_z^{\mathcal{A}} - r_z^{\mathcal{B}}) \end{pmatrix}.$$

The characteristic equation, $\det(R - \lambda I) = 0$, allows us to determine the eigenvalues

$$\lambda^2 = (r_x^{\mathcal{A}} - r_x^{\mathcal{B}})^2 + (r_y^{\mathcal{A}} - r_y^{\mathcal{B}})^2 + (r_z^{\mathcal{A}} - r_z^{\mathcal{B}})^2.$$

Thus,

$$\lambda_{1,2} = \pm\sqrt{(r_x^{\mathcal{A}} - r_x^{\mathcal{B}})^2 + (r_y^{\mathcal{A}} - r_y^{\mathcal{B}})^2 + (r_z^{\mathcal{A}} - r_z^{\mathcal{B}})^2} = \pm \mid \bar{r}^{\mathcal{A}} - \bar{r}^{\mathcal{B}} \mid .$$

Then

$$D(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \frac{1}{4}(\lambda_1 - \lambda_2) = \frac{1}{2}\sqrt{(r_x^{\mathcal{A}} - r_x^{\mathcal{B}})^2 + (r_y^{\mathcal{A}} - r_y^{\mathcal{B}})^2 + (r_z^{\mathcal{A}} - r_z^{\mathcal{B}})^2}.$$

The trace distance between two quantum systems $\mathcal{A}$ and $\mathcal{B}$ with density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$, respectively, is invariant to a linear transformation $\mathbf{U}$ of the states:

$$D(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = D(\mathbf{U}\rho^{\mathcal{A}}\mathbf{U}^\dagger, \mathbf{U}\rho^{\mathcal{B}}\mathbf{U}^\dagger).$$

The trace distance of two operators $\mathbf{A}$ and $\mathbf{B}$ is defined as

$$D(\mathbf{A}, \mathbf{B}) = \frac{1}{2}\text{tr}\left[\mid A - B \mid\right] = \frac{1}{2}\text{tr}\sqrt{(A - B)^\dagger(A - B)}$$

with $A$ and $B$ the matrices corresponding to $\mathbf{A}$ and $\mathbf{B}$, respectively.

**Proposition.** *If two operators* $\mathbf{A}$ *and* $\mathbf{B}$ *commute, then their trace distance is equal to the trace distance between the eigenvalue distribution of the two operators,* $D(\mathbf{A},\mathbf{B}) = D(a,b)$ *with* $a = \{a_1 a_2 \ldots a_{2^n}\}$ *and* $b = \{b_1 b_2 \ldots b_{2^n}\}$, $a_i$ *the eigenvalues of* $A$ *and* $b_i$ *the eigenvalues of* $B$.

*Proof:* If $[\mathbf{A},\mathbf{B}] = 0$ then the corresponding matrices can be diagonalized in the same orthonormal basis formed by unitary vectors $\mid i\rangle$, $1 \leq i \leq 2^n$:

$$A = \sum_i a_i \mid i\rangle\langle i \mid \qquad \text{and} \qquad B = \sum_i b_i \mid i\rangle\langle i \mid .$$

It follows that

$$D(\mathbf{A},\mathbf{B}) = \frac{1}{2}\text{tr}|\sum_i (a_i - b_i) \mid i\rangle\langle i \mid| = \frac{1}{2}\sum_i \mid a_i - b_i \mid = D(a,b).$$

Recall that a *positive operator* $\mathbf{A}$ is any Hermitian operator with non-negative eigenvalues; $\sqrt{\mathbf{A}}$ denotes its unique positive root. Then the fidelity of positive operators $\mathbf{A}$ and $\mathbf{B}$ is defined as

$$F(\mathbf{A},\mathbf{B}) = \text{tr}\sqrt{\mathbf{A}^{1/2}\mathbf{B}\mathbf{A}^{1/2}}.$$

**Proposition.** *If two operators* $\mathbf{A}$ *and* $\mathbf{B}$ *commute, then the fidelity* $F(\mathbf{A},\mathbf{B})$ *is equal to the fidelity of the eigenvalue distribution,* $F(\mathbf{A},\mathbf{B}) = F(a,b)$ *with* $a = \{a_1 a_2 \ldots a_{2^n}\}$ *and* $b = \{b_1 b_2 \ldots b_{2^n}\}$, $a_i$ *the eigenvalues of* $\mathbf{A}$ *and* $b_i$ *the eigenvalues of* $\mathbf{B}$.

*Proof:* Following the same arguments as before we write

$$A = \sum_i a_i \mid i\rangle\langle i \mid \qquad \text{and} \qquad B = \sum_i b_i \mid i\rangle\langle i \mid .$$

It follows that

$$F(\mathbf{A},\mathbf{B}) = \text{tr}\sqrt{\sum_i a_i b_i \left(\mid i\rangle\langle i \mid\right)^2} = \text{tr}\sum_i \sqrt{a_i b_i} \mid i\rangle\langle i \mid = \sum_i \sqrt{a_i b_i} = F(a,b).$$

Next we introduce another very important concept in quantum information theory the fidelity of quantum states.

**Fidelity of quantum states.** To quantify the properties of a quantum channel or to determine whether a quantum circuit performs its prescribed function we correlate the information at the input with the information at the output of the quantum device. The fidelity of quantum states is a measure of the distance between quantum states.

Fidelity allows us to answer the question "how close are two quantum states to each other?" The quantum fidelity is a measure of the distinguishability of two quantum states, $\mathcal{A}$ and $\mathcal{B}$, described by their density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$, respectively [221]. The fidelity is a real number between 0 and 1; it takes the value 0 if and only if the two states are orthogonal and it takes the value 1 if and only if $\rho^{\mathcal{A}} = \rho^{\mathcal{B}}$.

The mathematical apparatus required for the discussion of fidelity was introduced in Section 1.3 where we defined concepts such as positive operators, the modulus and the square root of an operator. The definition of the fidelity of two quantum states follows from the definition of the fidelity of general operators in Section 3.8; it uses the density operators of the two states rather than general operators. Recall that the density $\rho$ is a Hermitian operator with nonnegative eigenvalues; $\sqrt{\rho}$ denotes the unique positive square root of $\rho$.

Jozsa [221] defines the fidelity of two quantum states with density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ as

$$F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \left\{ \text{tr} \left[ \sqrt{(\rho^{\mathcal{A}})^{1/2} \rho^{\mathcal{B}} (\rho^{\mathcal{A}})^{1/2}} \right] \right\}^2 .$$

This quantity can be interpreted as a generalization of the transition probability for pure states. A simpler form is more convenient for some applications [308]:

$$F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \text{tr} \left[ \sqrt{(\rho^{\mathcal{A}})^{1/2} \rho^{\mathcal{B}} (\rho^{\mathcal{A}})^{1/2}} \right] .$$

The fidelity of two pure states $\mathcal{A}$ and $\mathcal{B}$ is *the norm of their inner product*; indeed, when

$$\rho^{\mathcal{A}} = \mid \varphi^{\mathcal{A}} \rangle \langle \varphi^{\mathcal{A}} \mid \quad \text{and} \quad \rho^{\mathcal{B}} = \mid \varphi^{\mathcal{B}} \rangle \langle \varphi^{\mathcal{B}} \mid$$

then,

$$F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \left\{ \text{tr} \left[ \sqrt{((\mid \varphi^{\mathcal{A}} \rangle \langle \varphi^{\mathcal{A}} \mid)^{1/2} (\mid \varphi^{\mathcal{B}} \rangle \langle \varphi^{\mathcal{B}} \mid)^{1/2}) \, ((\mid \varphi^{\mathcal{B}} \rangle \langle \varphi^{\mathcal{B}} \mid)^{1/2} (\mid \varphi^{\mathcal{A}} \rangle \langle \varphi^{\mathcal{A}} \mid)^{1/2})} \right] \right\}^2 .$$

It follows that $F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \mid \langle \varphi^{\mathcal{A}} \mid \varphi^{\mathcal{B}} \rangle \mid^2$. Defined as such, the fidelity of two pure states represents the usual transition probability between pure states.

The fidelity of a pure state $\mathcal{A}$ and a mixed state $\mathcal{B}$ is *the average of the previous expression over the ensemble of pure states with density matrix $\rho^{\mathcal{B}}$;* indeed, in this case $\rho^{\mathcal{A}}$, the pure state, has the same expression as before but now $\rho^{\mathcal{B}} = \sum_i p_i \mid \varphi_i^{\mathcal{B}} \rangle \langle \varphi_i^{\mathcal{B}} \mid$, therefore,

$$F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \langle \varphi^{\mathcal{A}} \mid \rho^{\mathcal{B}} \mid \varphi^{\mathcal{A}} \rangle .$$

This measure of fidelity represents the probability that $\rho^{\mathcal{B}}$ could prove to be the pure state $\mid \varphi^{\mathcal{A}} \rangle$ following a measurement of the observable $\mid \varphi^{\mathcal{A}} \rangle \langle \varphi^{\mathcal{A}} \mid$.

Intuitively, we expect the fidelity to be a symmetric function of its arguments, but symmetry is non-trivial to prove starting from the definition of fidelity. Before proving the symmetry of the quantum fidelity we need to prove Uhlmann theorem [430]. Recall that given a system in an $n$-dimensional Hilbert space, $\mathcal{H}_n$, with the density matrix $\rho$, a purification of $\rho$ is a pure state in an extended Hilbert space, $\mid \varphi \rangle \in \mathcal{H}_n \otimes \mathcal{H}_{n'}$ such that

$$\rho = \text{tr}_{\mathcal{H}_{n'}} \left( \mid \varphi \rangle \langle \varphi \mid \right) .$$

Purification treats $\rho$ as the reduced state of a subsystem; it allows us to perform a measurement and extract classical information without altering the state of a quantum system.

**Theorem.** *(Uhlmann) The fidelity of two mixed quantum states $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ is equal to the maximum value of the modulus of the inner product of any purifications $\mid \varphi^{\mathcal{A}}\rangle$ and $\mid \varphi^{\mathcal{B}}\rangle$, respectively, of the two density matrices.*

$$F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = \max_{(\varphi^{\mathcal{A}}, \varphi^{\mathcal{B}})} \mid \langle \varphi^{\mathcal{A}} \mid \varphi^{\mathcal{B}} \rangle \mid^2 .$$

The proof discussed in this sections is due to Richard Jozsa [221] and avoids the difficulties of the original proof of Uhlmann [430] based on the representation theory of $C^*$-algebras .

The two mixed states $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ are defined on an $n$-dimensional Hilbert space $\mathcal{H}_n$ and their purifications, $\mid \varphi^{\mathcal{A}}\rangle$ and $\mid \varphi^{\mathcal{B}}\rangle$ are defined on the extended space $\mathcal{H}_n \otimes \mathcal{H}'_n$. Without loss of generality we assume that $\mathcal{H}'_n = \mathcal{H}_n$ thus $\varphi^{\mathcal{A}}, \varphi^{\mathcal{B}} \in \mathcal{H}_n \otimes \mathcal{H}_n$. We wish to express the two pure states $\mid \varphi^{\mathcal{A}}\rangle$ and $\mid \varphi^{\mathcal{B}}\rangle$ in terms of the eigenvectors and eigenvalues of the corresponding density matrices, using Schmidt decomposition and

$$\mid \varphi^{\mathcal{A}}\rangle = \sum_i \sqrt{p_i^{\mathcal{A}}} \mid a_i\rangle \otimes \mid c_i\rangle \quad \text{and} \quad \mid \varphi^{\mathcal{B}}\rangle = \sum_i \sqrt{p_i^{\mathcal{B}}} \mid b_i\rangle \otimes \mid d_i\rangle.$$

In these expressions:

1. $A = \{\mid a_1\rangle, \mid a_2\rangle, \ldots \mid a_n\rangle\}$ is the set of orthonormal eigenvectors of $\rho^{\mathcal{A}}$ in $\mathcal{H}_n$; $p_1^{\mathcal{A}}, p_2^{\mathcal{A}}, \ldots, p_n^{\mathcal{A}}$ are the corresponding eigenvalues.
2. $B = \{\mid b_1\rangle, \mid b_2\rangle, \ldots \mid b_n\rangle\}$ is the set of orthonormal eigenvectors of $\rho^{\mathcal{B}}$ in $\mathcal{H}_n$; $p_1^{\mathcal{B}}, p_2^{\mathcal{B}}, \ldots, p_n^{\mathcal{B}}$ are the corresponding eigenvalues.
3. $C = \{\mid c_1\rangle, \mid c_2\rangle, \ldots \mid c_n\rangle\}$ and $D = \{\mid d_1\rangle, \mid d_2\rangle, \ldots \mid d_n\rangle\}$ are two orthonormal bases in $\mathcal{H}'_n$.

There are some unitary transformations $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ that allow us to express the orthonormal bases $B$, $C$, and $D$ as follows:

$$\mid b_i\rangle = \mathbf{B} \mid a_i\rangle, \qquad \mid c_i\rangle = \mathbf{C} \mid a_i\rangle, \qquad \text{and} \qquad \mid d_i\rangle = \mathbf{D} \mid b_i\rangle.$$

We also observe that

$$\sqrt{p_i^{\mathcal{A}}} = \sqrt{\rho^{\mathcal{A}}} \qquad \text{and} \qquad \sqrt{p_i^{\mathcal{B}}} = \sqrt{\rho^{\mathcal{B}}}.$$

Now we express the two pure states as

$$\mid \varphi^{\mathcal{A}}\rangle = \sum_i \sqrt{p_i^{\mathcal{A}}} \mid a_i\rangle \otimes \mid c_i\rangle = \left[\sqrt{\rho^{\mathcal{A}}} \otimes \mathbf{I}\right] \sum_i \mid a_i\rangle \otimes \mathbf{C} \mid a_i\rangle = \sum_i \sqrt{\rho^{\mathcal{A}}} \mathbf{C}^\dagger \mid a_i\rangle \otimes \mid a_i\rangle$$

and

$$\begin{aligned} \mid \varphi^{\mathcal{B}}\rangle &= \textstyle\sum_j \sqrt{p_j^{\mathcal{B}}} \mid b_j\rangle \otimes \mid d_j\rangle = \sum_j \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \mid b_j\rangle \otimes \mid b_j\rangle = \sum_j \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \mathbf{B} \mid a_j\rangle \otimes \mathbf{B} \mid a_j\rangle \\ &= \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \left[\mathbf{B} \otimes \mathbf{I}\right] \textstyle\sum_j \mid a_j\rangle \otimes \mathbf{B} \mid a_j\rangle = \sum_j \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \mathbf{B} \mathbf{B}^\dagger \mid a_j\rangle \otimes \mid a_j\rangle. \end{aligned}$$

Finally,

$$\langle \varphi^{\mathcal{A}} \mid \varphi^{\mathcal{B}} \rangle = \sum_i \langle a_i \mid \mathbf{C}^\dagger \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \mathbf{B} \mathbf{B}^\dagger \mid a_i \rangle = \mathrm{tr}\left[ \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \mathbf{D}^\dagger \mathbf{B} \mathbf{B}^\dagger \mathbf{C}^\dagger \right].$$

It follows that the maximum value of $\mid \langle \varphi^{\mathcal{A}} \mid \varphi^{\mathcal{B}} \rangle \mid$ is $\mathrm{tr} \mid \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \mid$ as $\mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ are unitary transformations. We also can write

$$\mid \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \mid = \left( \left( \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \right) \left( \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \right)^\dagger \right)^{1/2} = \left( \sqrt{\rho^{\mathcal{A}}} \sqrt{\rho^{\mathcal{B}}} \sqrt{\rho^{\mathcal{B}}} \sqrt{\rho^{\mathcal{A}}} \right)^{1/2} = \left( \sqrt{\rho^{\mathcal{A}}} \rho^{\mathcal{B}} \sqrt{\rho^{\mathcal{A}}} \right)^{1/2}.$$

We conclude that $F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = \max_{(\varphi^{\mathcal{A}}, \varphi^{\mathcal{B}})} \mid \langle \varphi^{\mathcal{A}} \mid \varphi^{\mathcal{B}} \rangle \mid^2$. □

**Properties of quantum fidelity.** The first three properties follow immediately from Uhlmann theorem while the others can be proved with relative ease.

(i) $0 \leq F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) \leq 1$;

(ii) $F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = 1$ if and only if $\rho^{\mathcal{A}} = \rho^{\mathcal{B}}$;

(iii) Symmetry: $F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = F\left(\rho^{\mathcal{B}}, \rho^{\mathcal{A}}\right)$;

(iv) Convexity: when $\rho^{\mathcal{A}}, \rho^{\mathcal{B}} \geq 0$ and $p^{\mathcal{A}} + p^{\mathcal{B}} = 1$ then,

$$F\left(\rho, p^{\mathcal{A}} \rho^{\mathcal{A}} + p^{\mathcal{B}} \rho^{\mathcal{B}}\right) \geq p^{\mathcal{A}} F\left(\rho, \rho^{\mathcal{A}}\right) + p^{\mathcal{B}} F\left(\rho, \rho^{\mathcal{B}}\right);$$

(v) Multiplicativity: $F\left(\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}, \rho^{\mathcal{C}} \otimes \rho^{\mathcal{D}}\right) = F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{C}}\right) F\left(\rho^{\mathcal{B}}, \rho^{\mathcal{D}}\right)$;

(vi) $F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right)$ is invariant under unitary transformations on the state space. If $\rho^{\mathcal{A}}_{before}, \rho^{\mathcal{B}}_{before}$ and $\rho^{\mathcal{A}}_{after}$ and $\rho^{\mathcal{B}}_{after}$ are the density operators of systems $\mathcal{A}$ and $\mathcal{B}$, before and after the unitary transformation $\mathbf{U}$ then,

$$F\left(\rho^{\mathcal{A}}_{after}, \rho^{\mathcal{B}}_{after}\right) = F\left(\rho^{\mathcal{A}}_{before}, \rho^{\mathcal{B}}_{before}\right);$$

(vii) Non-decreasing: If $\rho^{\mathcal{A}}_{before}, \rho^{\mathcal{B}}_{before}$ and $\rho^{\mathcal{A}}_{after}$ and $\rho^{\mathcal{B}}_{after}$ are the density operators of systems $\mathcal{A}$ and $\mathcal{B}$, before and after a measurement $\mathbf{M}$ then,

$$F\left(\rho^{\mathcal{A}}_{after}, \rho^{\mathcal{B}}_{after}\right) \geq F\left(\rho^{\mathcal{A}}_{before}, \rho^{\mathcal{B}}_{before}\right).$$

**An alternative definition of quantum fidelity [21].** Let $\mathbf{M}_i$ be a set of POVM measurement operators such that $\sum_i \mathbf{M}_i = I$. The fidelity defines the minimum overlap between the probability distributions $p_A(i)$ and $p_B(i)$, with $p_A(i) \equiv \mathrm{tr}(\rho^{\mathcal{A}} \mathbf{M}_i)$ and $p_B(i) \equiv \mathrm{tr}(\rho^{\mathcal{B}} \mathbf{M}_i)$, for the outcomes of the POVM measurement,

$$F'(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = \min_{\{\mathbf{M}_i\}} \sum_i \sqrt{\mathrm{tr}\left[\rho^{\mathcal{A}} \mathbf{M}_i\right]} \sqrt{\mathrm{tr}\left[\rho^{\mathcal{B}} \mathbf{M}_i\right]}.$$

A POVM that achieves this minimum is called an *optimal POVM*.

The proof that the two expressions of fidelity are equivalent, $F(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}) = F'(\rho^{\mathcal{A}}, \rho^{\mathcal{B}})$, starts with the observation that when $\mathbf{U}$ is a unitary operator, $\mathbf{U}\mathbf{U}^\dagger = I$, then

$$\rho^{\mathcal{A}}\mathbf{M}_i = (\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i(\rho^{\mathcal{A}})^{1/2} = \mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i(\rho^{\mathcal{A}})^{1/2}\mathbf{U}^\dagger = \left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)\left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger$$

and

$$\rho^{\mathcal{B}}\mathbf{M}_i = (\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i(\rho^{\mathcal{B}})^{1/2} = \left((\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right)\left((\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger.$$

The density matrices $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ as well as the measurement operators $\mathbf{M}_i$ are Hermitian. Then,

$$\begin{aligned}
F'(\rho^{\mathcal{A}},\rho^{\mathcal{B}}) &= \min_{\{\mathbf{M}_i\}} \textstyle\sum_i \sqrt{\mathrm{tr}\left[\rho^{\mathcal{A}}\mathbf{M}_i\right]}\sqrt{\mathrm{tr}\left[\rho^{\mathcal{B}}\mathbf{M}_i\right]} \\
&= \min_{\{\mathbf{M}_i\}} \textstyle\sum_i \sqrt{\mathrm{tr}\left[\left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)\left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger\right]} \\
&\qquad \times \sqrt{\mathrm{tr}\left[\left((\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right)\left((\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger\right]} \\
&\geq \textstyle\sum_i \sqrt{\mathrm{tr}\left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)\left(\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger} \\
&\qquad \times \sqrt{\mathrm{tr}\left[(\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right]\left((\rho^{\mathcal{B}})^{1/2}\mathbf{M}_i^{1/2}\right)^\dagger} \\
&\geq \textstyle\sum_i \mid \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\mathbf{M}_i^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid .
\end{aligned}$$

The last expression is based on the Schwartz inequality for the operator inner product, $\mathrm{tr}(AA^\dagger)\mathrm{tr}(BB^\dagger \geq \mid \mathrm{tr}(AB^\dagger) \mid^2$. It is easy to see that

$$\sum_i \mid \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\mathbf{M}_i^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid \geq \mid \sum_i \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\mathbf{M}_i^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid$$

and

$$\mid \sum_i \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}\mathbf{M}_i^{1/2}\mathbf{M}_i^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid = \mid \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid .$$

We address now the question of choosing the linear transformation $\mathbf{U}$ to make this inequality as tight as possible. Given a set of linear operators $\mathbf{U}_k$ and an operator $\mathbf{G}$ then,

$$\max_{U_k} \mid \mathrm{tr}\left[\mathbf{U}_k\mathbf{G}\right] \mid = \mathrm{tr}\sqrt{\mathbf{G}^\dagger\mathbf{G}}.$$

The maximum is obtained for a particular $\mathbf{U}_k$ with the property that $\mathbf{U}_k\mathbf{G} = \sqrt{(\mathbf{G}^\dagger\mathbf{G})}$. It follows that when $\mathbf{G} = (\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{B}})^{1/2}$,

$$\mathbf{U}(\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{B}})^{1/2} = \sqrt{(\rho^{\mathcal{B}})^{1/2}(\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{B}})^{1/2}} = \sqrt{(\rho^{\mathcal{B}})^{1/2}\rho^{\mathcal{A}}(\rho^{\mathcal{B}})^{1/2}} = F\left(\rho^{\mathcal{A}},\rho^{\mathcal{B}}\right).$$

Finally, we conclude that

$$F'\left(\rho^{\mathcal{A}},\rho^{\mathcal{B}}\right) \geq \mid \mathrm{tr}\left[\mathbf{U}(\rho^{\mathcal{A}})^{1/2}(\rho^{\mathcal{B}})^{1/2}\right] \mid = \sqrt{(\rho^{\mathcal{B}})^{1/2}\rho^{\mathcal{A}}(\rho^{\mathcal{B}})^{1/2}} = F\left(\rho^{\mathcal{A}},\rho^{\mathcal{B}}\right).$$

The POVM is optimal when we have equality in the above expression for $F'\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right)$. When $\rho^{\mathcal{B}}$ is invertible then an optimal POVM is a positive operator [21] given by

$$\mathbf{M} = \left(\rho^{\mathcal{B}}\right)^{-1/2} \sqrt{\left(\rho^{\mathcal{B}}\right)^{1/2} \rho^{\mathcal{A}} \left(\rho^{\mathcal{B}}\right)^{1/2}} \left(\rho^{\mathcal{B}}\right)^{-1/2}$$

and $\mathbf{M}_i = \mid i\rangle\langle i \mid$ with $\{\mid i\rangle\}$ an orthonormal eigenbasis of $\mathbf{M}$.

The properties of the fidelity operator will be used in Section 3.13 to show that a general mixed state of a quantum system cannot be broadcast. We are now concerned with the transmission of classical information through a quantum channel and the measurement process necessary to extract the classical information encoded as the state of a quantum system.

## 3.12 Accessible Information in a Quantum Measurement; Holevo Bound

Holevo bound [211] was probably the first major result in quantum information theory; it provides an upper limit of the accessible information in a quantum measurement. Holevo bound gives an upper bound on how well we can infer the value of an unknown random variable $X$ based on the knowledge of another random variable $Y$.

To formulate this result we consider the following setup: a source generates letters $X$ from an alphabet with probabilities $p_i, 1 \leq i \leq n$; Alice encodes the random variable $X$ and transmits it through a quantum channel; Bob performs a measurement of the quantum system and the result of this measurement is a random variable $Y$. The mutual information $I(X;Y)$ measures the amount of information about $X$ that Bob can infer from the measurement. We know from Section 3.3 that

$$I(X;Y) \leq H(X)$$

thus, Bob should attempt to carry out a measurement that maximizes the mutual information, $I(X;Y)$. The question posed is: how close to $H(X)$ can Bob get? In other words: is there an upper bound on $I(X;Y)$?

To answer this question we assume that Alice has a choice of quantum systems in states characterized by the density matrices $\rho_1, \rho_n, \ldots, \rho_n$; to send $X = k$ she chooses one of them, say the one characterized by $\rho_k$ and probability $p_k$. Recall from Chapter 2 that when Bob performs a POVM measurement $\mathbf{M}^Y$ which produces the random variable $Y$ then the probability that $Y = m$ when the system was prepared by Alice in state $X = k$ is

$$\text{Prob}(Y = m \mid X = k) = \text{tr}\left[\mathbf{M}^Y \rho_k\right].$$

We know the distribution of $X$ and $Y$, thus we can calculate $H(X)$ and $H(Y)$; to calculate $H(X;Y)$ we only need

$$\text{Prob}((Y = m, X = k) = p_k \text{tr}\left[\mathbf{M}^Y \rho_k\right].$$

Regardless of the measurement performed by Bob, the mutual information between the measurement outcome $Y$ and the preparation $X$ is limited by the <u>Holevo bound</u>:

$$I(X;Y) \leq S\left(\sum_i p_i\rho_i\right) - \sum_i p_i S(\rho_i).$$

The Holevo bound is achieved if and only if the encoding states $\rho_i$ commute with each other and the receiver, Bob, makes a von Neumann measurement in the basis in which $\rho_i$ are diagonal. As we know, in a von Neumann measurement the system is projected onto one of a complete set of mutually orthogonal states; in Bob's case, this set of states is chosen to be the basis in which the coding states are diagonal.

**Holevo information.** Consider an ensemble of mixed state with the density matrix $\rho = \sum_i p_i\rho_i$ with $\rho_i$ the density matrix of the individual mixed states of the ensemble and $p_i$ the apriori probability of mixed state $i$. Then $S(\rho)$ and $S(\rho_i)$ are the von Neumann entropy of the overall ensemble of states and of the individual state $i$, respectively. The quantity

$$\chi = S\left(\sum_i p_i\rho_i\right) - \sum_i p_i S(\rho_i)$$

is known as the Holevo information, the information available in the quantum ensemble. When the density matrices of the states used as signals to transport classical information over the quantum channel are $\rho_i,\ 1 \leq i \leq n$ with probability $p_i$ and $\mathcal{E}$ is the mapping of the density matrices performed by the channel then the Holevo information gives the classical capacity of a quantum channel and can be expressed as:

$$\chi = S\left(\sum_i p_i\mathcal{E}(\rho_i)\right) - \sum_i p_i S\left(\mathcal{E}(\rho_i)\right)$$

The maximum classical capacity of the channel is the maximum of this expression over all input ensembles

$$C = \chi_{max} = \max_{(p_i,\rho_i)}\left[S\left(\sum_i p_i\mathcal{E}(\rho_i)\right) - \sum_i p_i S\left(\mathcal{E}(\rho_i)\right)\right].$$

The von Neumann entropy is a concave function thus, for any mixed state with density matrix $\rho$ the Holevo information is non-negative

$$\chi(\rho) \geq 0.$$

Consider now a composite quantum system consisting of the pair $(\mathcal{X},\mathcal{Y})$ with the density matrix $\rho^{(\mathcal{X},\mathcal{Y})}$. Let $\chi(\rho^{\mathcal{X}})$ be the Holevo information for system $\mathcal{X}$ and $\chi(\rho^{(\mathcal{X},\mathcal{Y})})$ the Holevo information for the joint system. Schumacher, Westmoreland, and Wootters [370] showed that

$$\chi(\rho^{\mathcal{X}}) \leq \chi(\rho^{(\mathcal{X},\mathcal{Y})}).$$

They observe that: "...since $\chi$ is an upper bound of the accessible information, it is reasonable that this bound does not increase when part of the system is discarded."

To prove that the Holevo bound $\chi$ cannot increase when part of a joint system is discarded we use the relationship between the von Neumann entropy of a composite system in a product state and the Shannon entropy; consider two quantum systems $\mathcal{A}$ and $\mathcal{B}$ in a joint state with a density matrix of the form

$$\rho^{(\mathcal{A},\mathcal{B})} = \sum_i p_i \mid e_i^{\mathcal{A}}\rangle\langle e_i^{\mathcal{A}} \mid \otimes \rho_i^{\mathcal{B}}$$

with $\{\mid e_i^{\mathcal{A}}\rangle\}$ an orthonormal basis for $\mathcal{A}$ and $\rho_i^{\mathcal{B}}$ a set of density operators for system $\mathcal{B}$. The von Neumann entropy of the joint state is

$$S(\rho^{(\mathcal{A},\mathcal{B})}) = H(p) + \sum_i p_i S(\rho_i^{\mathcal{B}})$$

with $H(p)$ the Shannon entropy for the probability distribution of $p_i$.

Consider three systems $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ in a joint state described by a density matrix of the form

$$\rho^{(\mathcal{X},\mathcal{Y},\mathcal{Z})} = \sum_k p_k \rho_k^{(\mathcal{X},\mathcal{Y})} \otimes \mid k^{\mathcal{Z}}\rangle\langle k^{\mathcal{Z}} \mid$$

where $\{\mid k^{\mathcal{Z}}\rangle\}$ is an orthogonal set of states for the system $\mathcal{Z}$.

Recall from Section 3.10 that strong subadditivity requires that

$$S(\rho^{(\mathcal{X},\mathcal{Y},\mathcal{Z})}) + S(\rho^{\mathcal{X}}) \leq S(\rho^{(\mathcal{X},\mathcal{Y})}) + S(\rho^{(\mathcal{X},\mathcal{Z})}).$$

The relation between von Neumann and Shannon entropy allows us to express the elements of this expression as

$$S(\rho^{(\mathcal{X},\mathcal{Y},\mathcal{Z})}) = H(p) + \sum_k p_k S(\rho_k^{(\mathcal{X},\mathcal{Y})}) \qquad \text{and} \qquad S(\rho^{(\mathcal{X},\mathcal{Z})}) = H(p) + \sum_k p_k S(\rho_k^{\mathcal{X}}).$$

We substitute them in the strong subadditivity expression and obtain

$$S(\rho^{\mathcal{X}}) - \sum_k p_k S(\rho_k^{\mathcal{X}}) \leq S(\rho^{(\mathcal{X},\mathcal{Y})}) - \sum_k p_k S(\rho_k^{(\mathcal{X},\mathcal{Y})})$$

or $\chi(\rho^{\mathcal{X}}) \leq \chi(\rho^{(\mathcal{X},\mathcal{Y})})$.

□

**An alternative proof of the Holevo bound.** The proof in [370] is based on a physical model of a POVM measurement process where distinct outcomes are represented by positive operators rather than projections. Three systems are considered in the measurement process: (i) $\mathcal{Q}$ - the quantum system in state $\rho_{before}^{\mathcal{Q}}$; (ii) $\mathcal{A}$ - the apparatus used to perform the measurement; and (iii) $E$ - the environment.

Initially, the quantum system is in state $\rho_{before}^{\mathcal{Q}}$ and the additional systems, the apparatus and the environment are in a joint state $\rho_{before}^{(\mathcal{A},E)}$; the whole ensemble is in the joint state

$$\rho_{before}^{(\mathcal{A},E,\mathcal{Q})} = \rho_{before}^{(\mathcal{A},E)} \otimes \rho_{before}^{\mathcal{Q}}.$$

The measurement process proceeds in two stages: a dynamical evolution representing interactions among the three components captured by the unitary operator $\mathbf{U}$, followed by a second stage, the discarding of the environment. The state of the system after the first stage is

$$\rho_{fs}^{(\mathcal{A},E,\mathcal{Q})} = \mathbf{U}\ \rho_{before}^{(\mathcal{A},E,\mathcal{Q})}\ \mathbf{U}^{\dagger}.$$

The second transformation, the decoupling from the environment implies a partial trace over the system $E$

$$\rho_{after}^{(\mathcal{A},\mathcal{Q})} = \mathrm{tr}_E \left[\rho_{fs}^{(\mathcal{A},E,\mathcal{Q})}\right].$$

If $\rho_{after}^{\mathcal{Q}}$ is the state of the quantum system after the measurement, and $\mid \varphi_i^{\mathcal{A}}\rangle$ is a set of orthogonal states independent of the input state $\rho_{before}^{\mathcal{Q}}$, then the joint state of $\mathcal{A}$ and $\mathcal{Q}$ after the two stage measurement should be:

$$\rho_{after}^{(\mathcal{A},\mathcal{Q})} = \sum_i P_i \mid \varphi_i^{\mathcal{A}}\rangle\langle\varphi_i^{\mathcal{A}} \mid \otimes\rho_{after}^{\mathcal{Q}}.$$

In this expression $P_i$ is the probability of the outcome $i$ of the measurement. The choice of the states $\mid \varphi_i^{\mathcal{A}}\rangle$ is fixed for a given measurement device and the environment it is in and it is not related to the state of the system being measured.The states $\mid \varphi_i^{\mathcal{A}}\rangle$ are macroscopically distinguishable. A given measurement outcome does not necessarily correspond to a pure state of the apparatus system $\mathcal{A}$. If an observer does not completely resolves the state $\mid \varphi_i^{\mathcal{A}}\rangle$, then the measurement is an average over a group of different $i$ and she acquires less information.

Now we pay closer attention to the preparation of the initial state of the quantum system $\mathcal{Q}$. The preparer has a choice of $N$ initial states with density matrices $\rho_1^{\mathcal{Q}}, \rho_2^{\mathcal{Q}}, \ldots, \rho_N^{\mathcal{Q}}$ selected with the probabilities $p_1, p_2, \ldots, p_N$, respectively. Let us assume that she chooses: $\rho_{before}^{\mathcal{Q}} = \rho_k^{\mathcal{Q}}$. Then the corresponding state of the ensemble is

$$\rho_{k,before}^{(\mathcal{A},E,\mathcal{Q})} = \rho_{before}^{(\mathcal{A},\mathcal{Q})} \otimes \rho_k^{\mathcal{Q}}.$$

We wish to compute $\chi^{\mathcal{Q}}$ and we observe that $\chi^{\mathcal{Q}} = \chi^{(\mathcal{A},E,\mathcal{Q})}$ because the joint state $(\mathcal{A}, E)$ is independent of the preparation of the system $\mathcal{Q}$. To compute $\chi^{(\mathcal{A},E,\mathcal{Q})}$ we need $\rho_{before}^{(\mathcal{A},E,\mathcal{Q})}$, the density matrix of the joint state, an average over all possible preparation states

$$\bar{\rho}_{before}^{(\mathcal{A},E,\mathcal{Q})} = \sum_k p_k \rho_{k,before}^{(\mathcal{A},E,\mathcal{Q})}.$$

Then, after the second stage of the measurement process described earlier

$$\rho_{k,after}^{(\mathcal{A},\mathcal{Q})} = \sum_i P(i|k) \mid \varphi_i^{\mathcal{A}}\rangle\langle\varphi_i^{\mathcal{A}} \mid \otimes\rho_{ik,after}^{\mathcal{Q}}$$

with $P(i|k)$ the conditional probability of the measurement outcome $i$ given that the system was initially prepared in the state indexed by $k$ and $\rho_{ik,after}^{\mathcal{Q}}$ the resulting states of system $\mathcal{Q}$ when the preparation $k$ results in the outcome $i$. Again we consider the average state of $\mathcal{A}$ and $\mathcal{Q}$ after the measurement

$$\bar{\rho}_{after}^{(\mathcal{A},\mathcal{Q})} = \sum_k \sum_i P(i,k) \mid \varphi_i^{\mathcal{A}}\rangle\langle\varphi_i^{\mathcal{A}} \mid \otimes\rho_{ik,after}^{\mathcal{Q}}$$

where, the joint probability $P(i,k) = p_k P(i|k)$. Call $P(i)$ the total probability of the measurement outcome with index $i$. Then:

$$P(i) = \sum_k P(i,k).$$

and the final state of $\mathcal{Q}$ averaged over all preparations leading to outcome $i$ is $\bar{\rho}^{\mathcal{Q}}_{i,after} = \sum_k P(k|i)\rho^{\mathcal{Q}}_{ik,after}$. After summing for all $k$, the previous expression of the average joint state of $\mathcal{A}$ and $\mathcal{Q}$ after measurement becomes

$$\bar{\rho}^{(\mathcal{A},\mathcal{Q})}_{after} = \sum_i P(i) \mid \varphi^{\mathcal{A}}_i\rangle\langle\varphi^{\mathcal{A}}_i \mid \otimes \bar{\rho}^{\mathcal{Q}}_{i,after}.$$

The initial and the after the measurement Holevo information of the system consisting of the measuring apparatus $\mathcal{A}$, the environment, $E$, and the quantum system $\mathcal{Q}$ must be the same because the entropy is preserved under unitary transformation:

$$\chi^{(\mathcal{A},E,\mathcal{Q})}_{after} = \chi^{(\mathcal{A},E,\mathcal{Q})}_{before}.$$

Moreover, we have proved earlier that $\chi^{(\mathcal{A},\mathcal{Q})}_{after} \leq \chi^{(\mathcal{A},E,\mathcal{Q})}_{after}$ thus, we conclude that

$$\chi^{(\mathcal{A},\mathcal{Q})}_{after} \leq \chi^{(\mathcal{Q})}_{before}.$$

The definition of Holevo information allows us to express $\chi^{(\mathcal{A},\mathcal{Q})}_{after}$ as

$$\chi^{(\mathcal{A},\mathcal{Q})}_{after} = S(\bar{\rho}^{(\mathcal{A},\mathcal{Q})}_{after}) - \sum_k p_k S(\rho^{(\mathcal{A},\mathcal{Q})}_{k,after}).$$

We denote by $H(Y) = -\sum_i P(i)\log P(i)$ the Shannon entropy of the average probability distribution of $P(i)$ over the outcomes of the measurement; $H(Y|k)$ is the Shannon entropy of the outcomes of the measurement conditioned by the initial preparation with index $k$, and $H(Y|X) = \sum_k p_k H(Y|k)$ is the conditional entropy of the outcomes for all the preparations $k$. Then we express the two von Neumann entropy terms in the expression of $\chi^{(\mathcal{A},\mathcal{Q})}_{after}$ as

$$S(\bar{\rho}^{(\mathcal{A},\mathcal{Q})}_{after}) = H(Y) + \sum_i P(i) S(\rho^{\mathcal{Q}}_{i,after})$$

and

$$S(\bar{\rho}^{(\mathcal{A},\mathcal{Q})}_{k,after}) = H(Y|k) + \sum_i P(i \mid k) S(\rho^{\mathcal{Q}}_{ik,after}).$$

After substituting in the expression

$$\chi^{(\mathcal{Q})}_{before} \geq \chi^{(\mathcal{A},\mathcal{Q})}_{after}$$

it follows that

$$\chi^{\mathcal{Q}}_{before} \geq (H(Y) - H(Y|X)) + \sum_i P(i) \left[ S(\rho^{\mathcal{Q}}_{i,after}) - \sum_k P(k|i) S(\rho^{\mathcal{Q}}_{ik,after}) \right].$$

The Holevo information of the quantum system $Q$ after a measurement conditioned by an outcome indexed by $i$ is

$$\chi^{\mathcal{Q}}_{after,i} = S(\rho^{\mathcal{Q}}_{i,after}) - \sum_k P(k|i) S(\rho^{\mathcal{Q}}_{ik,after})$$

We also observe that the mutual information is $I(X;Y) = H(Y) - H(Y|X)$, thus we have proved that

$$I(X;Y) \leq \chi^{\mathcal{Q}}_{after} - \sum_i P(i) \chi^{\mathcal{Q}}_{after,i}.$$

This result of Schumacher, Westmoreland and Wootters [370] is more general than the Holevo theorem; it shows that the information obtained about the preparation of a system through a measurement process is bounded by the average amount the quantity $\chi$ and decreases during the measurement.

**An intuitive representation of von Neumann entropy before and after a measurement [89].** The entities involved in the reasoning in [370] are:

(i) The preparer - described by the random variable $X$ with the probability distribution $p_i, 1 \leq i \leq N$. The internal state of the preparer (a physical quantum system)is given by the density matrix:

$$\rho^X_{before} = \sum_i p_i \mid x_i \rangle \langle x_i \mid$$

with $\{\mid x_i \rangle\}$ a set of orthonormal preparer states.

(ii) The quantum channel $Q$. The preparer sends through the channel a quantum system in a mixed state $\rho_i$ selected out of a set of $N$ mixed states according to her internal state. The joint state of the preparer and the quantum communication channel is described by the density matrix:

$$\rho^{(X,Q)}_{before} = \sum_i p_i \mid x_i \rangle \langle x_i \mid \otimes \rho_i.$$

This density matrix is block diagonal, the quantum analog of the "grouping" property of Shannon entropy discussed in Section 3.3. As we know, the state of the quantum channel is the partial trace of the joint state:

$$\rho^Q_{before} = \mathrm{tr}_X \left[ \rho^{(X,Q)} \right] = \sum_i p_i \rho_i = \bar{\rho}.$$

(iii) An ancilla $A$. The measurement involves a unitary operation on the ancilla $A$ and the quantum channel $Q$. The information obtained as the result of the measurement is the mutual information between the preparer and the ancilla $H_{after}(X, A)$. Initially the ancilla is in a reference state $\mid 0 \rangle$ and its density matrix is

$$\rho^A_{before} = \mid 0 \rangle \langle 0 \mid .$$

The initial joint state of the ternary system $(X, Q, A)$ can be written as

$$\rho_{before}^{(X,Q,A)} = \rho_{before}^{(X,Q)} \otimes \mid 0\rangle\langle 0 \mid .$$

Now we express the von Neumann entropy of the three components, as well as the joint entropy before and after the measurement:

$$S_{before}(X) = H(p_i), \quad S_{before}(Q) = S(\bar{\rho}), \quad S_{before}(X, Q) = H(p_i) + \sum_i p_i S(\rho_i).$$

The last equality exploits the fact that $\rho_{before}^{(X,Q)}$ is block diagonal. We also observe that the von Neumann entropy of the joint state of $X$, $Q$, and $A$ is

$$S_{before}(X, Q, A) = S_{before}(X, Q)$$

because $A$ is initially in a pure state, $\mid 0\rangle$.

Then, $S_{before}(X; Q)$ (see Figure 78(a)), the quantum mutual entropy, or the mutual entanglement, between the preparer and the quantum channel is equal to the Holevo bound:

$$S_{before}(X; Q) = S_{before}(X) + S_{before}(Q) - S_{before}(X, Q) = S(\bar{\rho}) - \sum_i p_i S(\rho_i) = \chi^{(X)}.$$

The Venn diagram in Figure 78 (a) summarizes this relationships between the von Neumann entropy of $X$ and $Q$ before the measurement and Figure 78 (b) illustrates the joint and the conditional von Neumann entropy of the three subsystems involved after the measurement.

Let us now consider the measurement, a unitary transformation $U^{(Q,A)}$, affecting only the quantum channel and the ancilla. The joint state of the three systems after the measurement is

$$\rho_{after}^{(X,Q,A)} = \left(I^X \otimes U^{(Q,A)}\right) \left(\rho_{before}^{(X,Q)} \otimes \mid 0\rangle\langle 0 \mid\right) \left(I^X \otimes U^{(Q,A)}\right)^\dagger .$$

As noted earlier, we have to find an upper bound for the mutual entanglement between $X$ and $A$, $S_{after}(X; A)$. To do so we first apply the chain rule for mutual entropies

$$S_{after}\left(X; (Q, A)\right) = S_{after}\left(X; A\right) + S_{after}\left(X; Q|A\right).$$

Then we observe that the mutual entanglement between $X$ and the joint system $(Q, A)$ is conserved because the measurement involves a unitary transformation of the joint system $(Q, A)$ and does not affect the state of $X$. Thus,

$$S_{after}\left(X; (Q, A)\right) = S_{before}\left(X; (Q, A)\right) = S_{before}\left(X; Q\right).$$

It follows that

$$S_{after}\left(X; A\right) = S_{before}\left(X; Q\right) - S_{after}\left(X; Q|A\right).$$

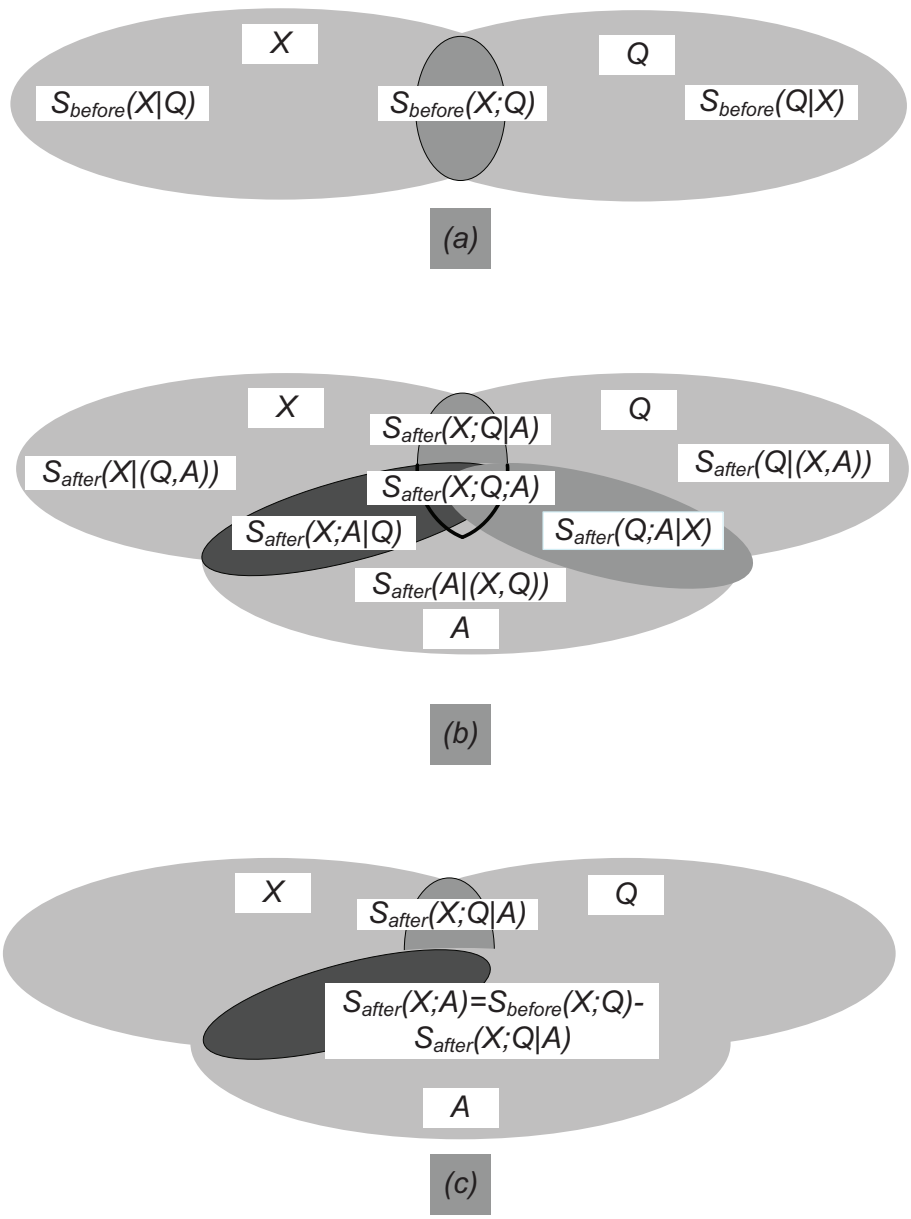



Figure 78: Venn diagrams showing the von Neumann entropy (a) before and (b) after the measurement. Three systems are involved: (i) the preparer $X$; (ii) the quantum channel $Q$; and the ancilla $A$. (c) The mutual entropy between the ancilla and the preparer after the measurement: $S_{after}(X;A) = S_{before}(X;Q) - S_{after}(X;Q|A)$,where $S_{before}(X;Q)$ is in fact $S_{after}(X;(Q,A))$

This equation tells us that the information obtained through the measurement is given by $S_{before}(X;Q)$ (the Holevo bound) reduced by a quantity which is the mutual entropy (after measurement) between the prepare's state $X$ and the state of the quantum channel $Q$, conditional on the observed state of the ancilla $A$ (all states existing after measurement).

Figure 78 (c) illustrates the quantities involved in this expression. Strong subadditivity

of von Neumann entropy implies that mutual entanglement never decreases when extending a system; it follows that $S_{after}\,(X;Q|A) \geq 0$ therefore,

$$S_{after}\,(X;A) \leq S_{before}\,(X;Q)\,.$$

The second term on the right hand side of the expression giving the mutual information between the preparer and the ancilla is the quantum conditional mutual entropy:

$$S_{after}\,(X;Q|A) = S_{after}(X,A) + S_{after}(Q,A) - S_{after}(A) - S_{after}(X,Q,A).$$

This expression is difficult to estimate.

Finally, Cerf and Adami show that a von Neumann measurement allows us to express the von Neumann entropy as a function of classical Shannon entropy [89]:

$$S_{after}\,(X;A) = H(A) - H(A|X) = I(X;A).$$

It follows that the mutual information satisfies the Holevo bound:

$$I(X;A) \leq S_{before}(X;Q) = S(\bar{\rho}) - \sum_i p_i S(\rho_i).$$

This inequality shows that some information about the state before measurement of the system $X$ might still be extractable from the state after measurement of the system $Q$; this is the case when the measurement is "incomplete".

An immediate consequence of the Holevo bound is that no more than $n$ bits of information can be communicated using $n$ qubits. Indeed, recall that given two correlated random variables $X$ and $Y$ the maximum amount of information about $X$ available prior to transmission at the receiving end of a classical communication channel is $H(X)$; after the transmission the average uncertainty is $H(X|Y)$. Therefore, $H(X) - H(X|Y)$ is the maximum amount of information that can be extracted by the recipient of information. $I(X;Y) = H(X) - H(X|Y)$ is the mutual information between $X$ and $Y$. The superdense coding scheme [33] allows the transmission of two bits of information using a single qubit, but in addition to the quantum channel it requires a classical channel to transmit the results of a measurement performed by the sender.

The information about the random variable $X$ could be encoded on $n$ qubits. The state of the $n$ qubit quantum register $Y$ encoding the information about a particular value of the input random variable $X = x$ is described by the density matrix $\rho$ in a $2^n$-dimensional Hilbert space $\mathcal{H}_{2^n}$. Regardless of the choices of $x$ and $\rho$ the POVM measurement of $Y$ produces at most $\chi \leq n$ bits of information about $X$.

**Example.** *Holevo bound for information encoded as the polarization of a photon.* Alice sends Bob information encoded as the polarization of a photon; Bob receives a photon with either vertical polarization, or a polarization at an angle $\theta$, with equal probabilities $p_0 = p_1 = 1/2$. We wish to determine the Holevo bound for the information Bob could extract, to study the dependence of this bound on the angle $\theta$, and what is the optimal measurement.

The possible states of the photon are: $\mid \varphi_1\rangle = \mid 0\rangle$ and $\mid \varphi_2\rangle = \cos\theta \mid 0\rangle + \sin\theta \mid 1\rangle$. The density matrix is

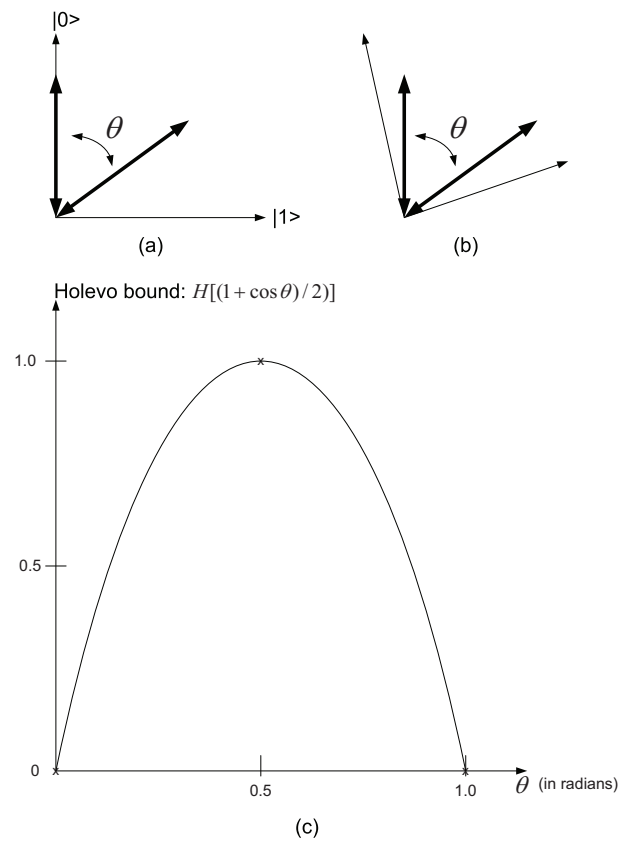



Figure 79: Information encoded as photon polarization. (a) Two non orthogonal states. (b) An optimal measurement. (c) The plot of the Holevo bound, $H\left((1+\cos\theta)/2\right)$, function of the angle between photon polarization. Bob extracts maximum information when $\theta = \pi/2$ which corresponds to orthogonal states.

$$\rho = p_0\rho_0 + p_1\rho_1 = \frac{1}{2}\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \end{pmatrix}\right].$$

Thus,

$$\rho = \frac{1}{2}\begin{pmatrix} 1+\cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & 1-\cos^2\theta \end{pmatrix}.$$

We solve the characteristic equation $\det(\rho - \lambda I) = 0$ to find the eigenvalues of the density matrix

$$\det\left[\frac{1}{2}\begin{pmatrix} 1+\cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & 1-\cos^2\theta \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right] = \det\frac{1}{2}\begin{pmatrix} 1+\cos^2\theta - 2\lambda & \sin\theta\cos\theta \\ \sin\theta\cos\theta & 1-\cos^2\theta - 2\lambda \end{pmatrix}.$$

Then the characteristic equation becomes

$$\left[(1-2\lambda)+\cos^2\theta\right]\left[(1-2\lambda)-\cos^2\theta\right] - \sin^2\theta\cos^2\theta = 0$$

or

$$(1-2\lambda)^2 = \cos^2\theta.$$

The solutions are

$$\lambda_{1,2} = \frac{1 \pm \cos\theta}{2}.$$

We can now express the von Neumann entropy of the photon source as

$$S(\rho) = -\lambda_1 \log \lambda_1 - \lambda_2 \log \lambda_2 = -\frac{1+\cos\theta}{2} \log \frac{1+\cos\theta}{2} - \frac{1-\cos\theta}{2} \log \frac{1-\cos\theta}{2}.$$

Recall that $H(p) = -p \log p - (1-p)\log(1-p)$ is the Shannon entropy of a binary random variable with probability $p$ thus,

$$S(\rho) = H\left(\frac{1+\cos\theta}{2}\right).$$

Figures 79(a), (b), and (c) illustrate the polarization of the two photons, the optimal measurement, and the graph of the Holevo bound function of the angle $\theta$. As expected, Bob extracts maximum information when $\theta = \pi/2$, thus when the two states are orthogonal.

Next we discuss the no broadcasting theorem for mixed quantum states.

## 3.13 No Broadcasting Theorem for General Mixed States

*Broadcasting* is a term used to describe the one-to-many communication where multiple recipients of information receive the same message. There is no limitation regarding broadcasting of classical information. In fact, widely used networking techniques for local area networks (e.g., Ethernet) and for radio and satellite communication use broadcasting. On the other hand, broadcasting a quantum state would allow us to replicate a mixed state, so we should not be surprised that general mixed states cannot be broadcast [21].

In Section 2.1 we discussed the impossibility of cloning pure or mixed non-orthogonal quantum states; cloning would allow us to replicate the states and distinguish between them with an arbitrary high reliability by measuring a large number of replicas of each state. We presented another line of reasoning for the impossibility of cloning mixed quantum states in Section 2.15; cloning an entangled state would lead to a violation of the monogamy of entanglement, the partner of an entangled state would end up being entangled with two states, the original, as well as, the cloned state.

Another argument for the impossibility of cloning entangled states is that neither one of two entangled systems has a definite state of its own, so it is impossible to clone a system whose state is not well defined. Now we reexamine cloning in a more general context provided by quantum information theory.

Schumacher draws a line between *copying* and *transposition*, an important issue for quantum information encoding [369]. If two quantum systems, $\mathcal{A}$ in a general state $\mid \varphi^{\mathcal{A}}\rangle$, and $\mathcal{B}$ in a standard "null" state $\mid \varphi^{\mathcal{B}}\rangle = \mid 0^{\mathcal{B}}\rangle$, are the input and, respectively, the output of a source

Table 7: The states of a quantum system before and after copy and transposition.

| System | Initial State | State after **Copy** | State after **Transposition** |
|---|---|---|---|
| $\mathcal{A}$ | $\mid \varphi^{\mathcal{A}}\rangle$ | $\mid \varphi^{\mathcal{A}}\rangle$ | $\mid 0^{\mathcal{A}}\rangle$ |
| $\mathcal{B}$ | $\mid \varphi^{\mathcal{B}}\rangle = \mid 0^{\mathcal{B}}\rangle$ | $\mid \varphi^{\mathcal{A}}\rangle$ | $\mid \varphi^{\mathcal{A}}\rangle$ |
| $(\mathcal{A},\mathcal{B})$ | $\mid \varphi^{\mathcal{A}}\rangle \otimes \mid 0^{\mathcal{B}}\rangle$ | $\mid \varphi^{\mathcal{A}}\rangle \otimes \mid \varphi^{\mathcal{A}}\rangle$ | $\mid 0^{\mathcal{A}}\rangle \otimes \mid \varphi^{\mathcal{A}}\rangle$ |

or a channel encoder, then *copying* is described by a transformation where the original state of $\mathcal{A}$ is not disturbed and $\mathcal{B}$ is brought into a state corresponding to the original state of $\mathcal{A}$. The *transposition* does not leave a copy behind; the state of $\mathcal{A}$ is transferred to $\mathcal{B}$ and, at the same time, the original state of $\mathcal{A}$, $\mid \varphi^{\mathcal{A}}\rangle$, is erased, Table 7. Quantum information copying is possible if and only if the original states are mutually orthogonal.

We consider a slightly more general problem involving two quantum systems in an $n$-dimensional Hilbert space $\mathcal{H}_n$; the two systems are: $\mathcal{A}$ with two possible initial states and $\mathcal{B}$. The states of the two systems are characterized by the density matrices: $\rho^{\mathcal{A}}_{before} = \rho_i, i \in \{0,1\}$ and $\rho^{\mathcal{B}}_{before}$. The initial state of the composite system $(\mathcal{A},\mathcal{B})$ is

$$\rho^{(\mathcal{A},\mathcal{B})}_{before} = \rho^{\mathcal{A}}_{before} \otimes \rho^{\mathcal{B}}_{before}.$$

The question we pose is if there is a physical process $\mathcal{E}$ such that the state after the transformation, $\rho^{(\mathcal{A},\mathcal{B})}_{after} = \mathcal{E}\left(\rho^{\mathcal{A}}_{before} \otimes \rho^{\mathcal{B}}_{before}\right) \in \mathcal{H}_{n^2}$ has the following property:

$$\mathrm{tr}_{\mathcal{A}}\left[\rho^{(\mathcal{A},\mathcal{B})}_{after}\right] = \mathrm{tr}_{\mathcal{B}}\left[\rho^{(\mathcal{A},\mathcal{B})}_{after}\right] = \rho^{\mathcal{A}}_{before}.$$

If such a transformation, $\mathcal{E}$, exists for the two possible initial states $\rho_0$ and $\rho_1$ of $\mathcal{A}$, then we say that $\mathcal{A}$ can be broadcast.

**Theorem.** *(No-broadcasting) $\mathcal{A}$ can be broadcast if and only if the two initial states of $\mathcal{A}$ commute, $\rho_0 \otimes \rho_1 = \rho_1 \otimes \rho_0$.*

*Proof:* The implication that the state can be broadcast if the two states commute is easy to prove. Indeed, if the two states commute then there is an orthonormal basis $\{\mid e_i\rangle\}$ in which $\rho^{\mathcal{A}}_0$ and $\rho^{\mathcal{A}}_1$ can be simultaneously diagonalized:

$$\rho_0 = \sum_{i=1}^{n} p_{i,0} \mid e_i\rangle\langle e_i \mid \quad \text{and} \quad \rho_1 = \sum_{i=1}^{n} p_{i,1} \mid e_i\rangle\langle e_i \mid .$$

Consider a particular state of $\mathcal{B}$ namely $\rho^{\mathcal{B}} = \mid 1\rangle\langle 1 \mid$ and let $\mathbf{U}$ be a unitary transformation of the joint system $(\mathcal{A},\mathcal{B})$ such that

$$\mathbf{U}\left(\mid e_i\rangle \otimes \mid 1\rangle\right) = \mid e_i\rangle \otimes \mid e_i\rangle, \quad \forall i \in \{1,n\}.$$

In this case,

$$\rho^{(\mathcal{A},\mathcal{B})}_{after} = \mathbf{U}\rho^{(\mathcal{A},\mathcal{B})}_{before}\mathbf{U}^{\dagger}.$$

When $\rho^{\mathcal{A}}_{before} = \rho_0$ then

$$\begin{aligned}\rho_{after}^{(\mathcal{A},\mathcal{B})} &= \mathbf{U}\left(\sum_i p_{i,0} \mid e_i\rangle\langle e_i \mid\right) \otimes \left(\mid 1\rangle\langle 1 \mid\right) \mathbf{U}^\dagger = \sum_i p_{i,0} \left[\mathbf{U}(\mid e_i\rangle \otimes \mid 1\rangle)\right] \left[\mathbf{U}(\mid e_i\rangle \otimes \mid 1\rangle)\right]^\dagger \\ &= \sum_i p_{i,0} \mid e_i\rangle \mid e_i\rangle\langle e_i \mid \langle e_i \mid .\end{aligned}$$

It follows that

$$\mathrm{tr}_{\mathcal{A}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \mathrm{tr}_{\mathcal{B}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \sum_{i=1}^{n} p_{i,0} \mid e_i\rangle\langle e_i \mid = \rho_0.$$

Similarly, when $\rho_{before}^{\mathcal{A}} = \rho_1$ then

$$\begin{aligned}\rho_{after}^{(\mathcal{A},\mathcal{B})} &= \mathbf{U}\left(\sum_i p_{i,1} \mid e_i\rangle\langle e_i \mid\right) \otimes \left(\mid 1\rangle\langle 1 \mid\right) \mathbf{U}^\dagger = \sum_i p_{i,1} \left[\mathbf{U}(\mid e_i\rangle \otimes \mid 1\rangle)\right] \left[\mathbf{U}(\mid e_i\rangle \otimes \mid 1\rangle)\right]^\dagger \\ &= \sum_i p_{i,1} \mid e_i\rangle \mid e_i\rangle\langle e_i \mid \langle e_i \mid .\end{aligned}$$

It follows that

$$\mathrm{tr}_{\mathcal{A}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \mathrm{tr}_{\mathcal{B}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \sum_{i=1}^{n} p_{i,1} \mid e_i\rangle\langle e_i \mid = \rho_1.$$

It is harder to prove that when $\mathcal{A}$ can be broadcast then the two initial states of $\mathcal{A}$ commute. The proof in [21] is based on the following property of fidelity mentioned in Section 3.11: *fidelity does not decrease under partial trace.* We start by assuming that

$$\mathrm{tr}_{\mathcal{A}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \mathrm{tr}_{\mathcal{B}}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\right] = \rho_{before}^{\mathcal{A}}.$$

Let $\{\mathbf{M}_i\}$ be an optimal POVM measurement allowing us to distinguish between the states $\rho_0$ and $\rho_1$; the system$\mathcal{A}$ was initially prepared in one of them. It is easy to see that

$$\mathrm{tr}\left[\rho_{after}^{(\mathcal{A},\mathcal{B})}\left(\mathbf{M}_i \otimes \mathbf{I}\right)\right] = \mathrm{tr}_{\mathcal{A}}\left[\mathrm{tr}_{\mathcal{B}}\left(\rho_{after}^{(\mathcal{A},\mathcal{B})}\right)\mathbf{M}_i\right] = \mathrm{tr}_{\mathcal{A}}\left[\rho_{before}^{\mathcal{A}}\mathbf{M}_i\right].$$

But $(\mathbf{M}_i \otimes \mathbf{I})$ may not be an optimal POVM to distinguish between $\rho_{0,after}^{(\mathcal{A},\mathcal{B})}$ and $\rho_{1,after}^{(\mathcal{A},\mathcal{B})}$ the density matrices of the joint system after the measurement, corresponding to the system $\mathcal{A}$ being prepared in state $\rho_0$ and $\rho_1$, respectively. It follows that, indeed, fidelity cannot decrease under partial trace

$$\begin{aligned}F_{\mathcal{A}}\left(\rho_0, \rho_1\right) &= \sum_i \sqrt{\mathrm{tr}\left[\rho_{0,after}^{(\mathcal{A},\mathcal{B})}\left(\mathbf{M}_i \otimes \mathbf{I}\right)\right]}\sqrt{\mathrm{tr}\left[\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\left(\mathbf{M}_i \otimes \mathbf{I}\right)\right]} \\ &\geq \min_{\{\mathbf{N}_j\}}\left[\sum_j \sqrt{\mathrm{tr}\left[\rho_{0,after}^{(\mathcal{A},\mathcal{B})}\mathbf{N}_j\right]}\sqrt{\mathrm{tr}\left[\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\mathbf{N}_j\right]}\right] \\ &\geq F\left(\rho_{0,after}^{(\mathcal{A},\mathcal{B})}, \rho_{1,after}^{(\mathcal{A},\mathcal{B})}\right).\end{aligned}$$

Similarly,

$$
\begin{aligned}
F_{\mathcal{B}}\left(\rho_0,\rho_1\right) &= \textstyle\sum_i \sqrt{\operatorname{tr}\left[\rho_{0,after}^{(\mathcal{A},\mathcal{B})}\left(\mathbf{I}\otimes\mathbf{M}_i\right)\right]}\sqrt{\operatorname{tr}\left[\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\left(\mathbf{I}\otimes\mathbf{M}_i\right)\right]} \\
&\geq F\left(\rho_{0,after}^{(\mathcal{A},\mathcal{B})},\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\right).
\end{aligned}
$$

We can also prove the opposite inequalities

$$
F_{\mathcal{A}}\left(\rho_0,\rho_1\right) \leq F\left(\rho_{0,after}^{(\mathcal{A},\mathcal{B})},\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\right) \quad \text{and} \quad F_{\mathcal{B}}\left(\rho_0,\rho_1\right) \leq F\left(\rho_{0,after}^{(\mathcal{A},\mathcal{B})},\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\right).
$$

This means that if $\mathcal{A}$ can be broadcast then there are density operators $\rho_{0,after}^{(\mathcal{A},\mathcal{B})}$ and $\rho_{1,after}^{(\mathcal{A},\mathcal{B})}$ describing the joint system $(\mathcal{A},\mathcal{B})$ after the measurement, corresponding to $\rho_0$ and $\rho_1$, respectively, such that

$$
F\left(\rho_{0,after}^{(\mathcal{A},\mathcal{B})},\rho_{1,after}^{(\mathcal{A},\mathcal{B})}\right) = F_{\mathcal{A}}\left(\rho_0,\rho_1\right) = F_{\mathcal{B}}\left(\rho_0,\rho_1\right).
$$

The proof that these equalities imply that $\rho_1$ and $\rho_2$ commute can be found in [21].

Cloning is a special case of broadcasting; when given a process $\mathcal{E}$ we have

$$
\mathcal{E}\left(\rho_{before}^{\mathcal{A}}\otimes\rho_{before}^{\mathcal{B}}\right) = \rho_{before}^{\mathcal{A}}\otimes\rho_{before}^{\mathcal{A}}.
$$

The last equality implies that $\mathcal{A}$ is clonable if and only if

$$
F\left(\rho_0,\rho_1\right) = F\left(\rho_0\otimes\rho_0,\rho_1\otimes\rho_1\right) = \left[F\left(\rho_0,\rho_1\right)\right]^2 \quad\Longrightarrow\quad F\left(\rho_0,\rho_1\right)=0 \;\text{ or }\; F\left(\rho_0,\rho_1\right)=1.
$$

In other words, *cloning is possible if and only if the two initial states of $\mathcal{A}$ are either identical, or orthogonal.*

We are now able to discuss the quantum equivalent of Shannon source coding, Schumacher compression.

## 3.14 Schumacher Compression

Schumacher compression follows closely the arguments for classical source encoding from Section 3.4. The input alphabet of the source $\mathcal{A}$ is determined by the number $D$ of eigenstates of the density operator of one qubit; we assume that $D=2$, though other physical implementations of qubits with $D>2$ may be possible.

We encode sequences of $n$ qubits in $(\mathcal{H}_2)^{\otimes n}$ and the input states of the encoder are

$$
\mid \alpha_k\rangle = \otimes_{i=1}^{n} \mid a_{i,k}\rangle, \quad 1\leq k\leq K.
$$

with probability

$$
P_k = \prod_{i=1}^{n} p_{i,k}.
$$

The density matrix of the input to the encoder,

$$\rho^{\mathcal{A}} = \sum_{k=1}^{K} P_k \rho_k = \sum_{k=1}^{K} P_k \prod_{i=1}^{n} \rho_{i,k},$$

has $d$ eigenvectors $\mid \varphi_i\rangle$, $1 \leq i \leq d$. with probabilities $p_i$, $1 \leq i \leq d$. The source is characterized by von Neumann entropy $S(\rho) = -\text{tr}\left[\rho^{\mathcal{A}} \log \rho^{\mathcal{A}}\right]$. An eigenstate of the system is a sequence of $d$ integers, from the alphabet $(0-D)$, in our case 0s and 1s, and the eigenvalue of this state is the probability that this particular sequence was generated by the source.

We cannot expect that the output of a quantum encoder matches exactly its input when the block length increases, $n \mapsto \infty$, as is the case of classical information, therefore we have to find other ways to correlate the two. Now we use the fidelity between the input and the output states of the encoder as a measure of "goodness" for the transformation of the quantum state performed by the encoder. If $\rho_{out}$ is the output of the encoder then the fidelity is

$$F(\rho^{\mathcal{A}}, \rho_{out}) = \sqrt{(\rho^{\mathcal{A}})^{1/2} \rho_{out} (\rho^{\mathcal{A}})^{1/2}}.$$

By analogy with a classical source we define typical subspaces of the source with von Neumann entropy $S(\rho)$ as subspaces generated by typical sequences of eigenvectors of the density matrix. Call $\Lambda_{typical}$ the subspace generated by typical sequences of eigenvectors of $\rho^{\mathcal{A}}$. Almost all the time the source produces a typical subspace.

We can now formulate Schumacher compression based on the analogy between the probabilities and the eigenvalues

$$\dim\left(\Lambda_{typical}\right) = 2^{nS(\rho^{\mathcal{A}}) + \mathcal{O}(n)}.$$

**Theorem.** *We can compress the quantum information produced by a source $\mathcal{A}$ with von Neumann entropy $S(\rho^{\mathcal{A}})$ up to $\log\left[2^{nS(\rho^{\mathcal{A}})+\mathcal{O}(n)}\right] \approx nS(\rho^{\mathcal{A}})$ qubits; we can send $n$ symbols using $[nS(\rho_{\mathcal{A}}) + \mathcal{O}(n)]$ qubits with a fidelity approaching 1 as $n \to \infty$.*

The Schumacher compression works [389] because two quantum sources: (i) one with the eigenvectors $\mid \varphi_1\rangle, \mid \varphi_2\rangle, \ldots, \mid \varphi_d\rangle$ with probabilities $p_i, 1 \leq i \leq d$ and (ii) one with the states $\mid \alpha_1\rangle, \mid \alpha_2\rangle, \ldots, \mid \alpha_K\rangle$ with probabilities $P_1, P_2, \ldots, P_K$ have the same density matrix $\rho$ and, therefore, the probability of the outcomes will be the same. The density matrix determines the outcomes of any measurement. It follows that the input state of the encoder is almost always close to a typical subspace of the source. More details regarding the encoding process for noiseless quantum channels are presented in Section 3.17.

Finally, we should remember that *the von Neumann entropy does not provide a complete characterization of a source of quantum information;* indeed, two different sources of quantum information, $\mathcal{A}$ and $\mathcal{B}$ may have the same density matrix:

$$\rho^{\mathcal{A}} = \sum_k p_k^{\mathcal{A}} \mid \varphi_k^{\mathcal{A}}\rangle\langle\varphi_k^{\mathcal{A}} \mid \quad \text{and} \quad \rho^{\mathcal{B}} = \sum_k p_k^{\mathcal{B}} \mid \varphi_k^{\mathcal{B}}\rangle\langle\varphi_k^{\mathcal{B}} \mid, \quad \rho^{\mathcal{A}} = \rho^{\mathcal{B}} = \rho.$$

The von Neumann entropy of $\mathcal{A}$ and $\mathcal{B}$ is the same, $S(\rho) = -\text{tr}\left[\rho \log \rho\right]$, but the two sources, $\mathcal{A}$ and $\mathcal{B}$, are different from the quantum information theoretical point of view. The recipient of the information sent by the two sources may be able to discover the identity of the input

state more accurately for one of the sources than for the other. For example, the source $\mathcal{A}$ emits linearly polarized photons at $0^\circ$ and $90^\circ$ with equal probability $p^{\mathcal{A}}_{1,2} = 1/2$; the density matrix of $\mathcal{A}$ is $\rho^{\mathcal{A}} = I/2$ with $I$ the identity matrix, and the entropy of $\mathcal{A}$ is one bit. The states of the two photons are orthogonal, thus perfectly distinguishable and the information can be measured and transmitted as classical information. The source $\mathcal{B}$ emits linearly polarized photons at $0^\circ, 45^\circ, 90^\circ$ and $135^\circ$ with equal probability, $p^{\mathcal{B}}_{1,2,3,4} = 1/4$; the density matrix is $\rho^{\mathcal{B}} = I/2$ but in this case the information cannot be transmitted as classical information and the recipient of information cannot identify with absolute certainty the state of the source $\mathcal{B}$ [35]. The problem of reversible extraction of classical information from quantum information is discussed in Section 3.20.

Next we turn our attention from the source of quantum information to quantum channels.

## 3.15 Quantum Channels

Quantum information theory studies transmission of classical and quantum information over quantum channels from one point to another, or from one time to another; whenever we talk about the "input" and the "output" of the quantum channel we mean the state of the quantum system before and, respectively, after the transformation done by the quantum channel.

To develop a theory of transmission of quantum information we map known concepts from the classical information theory to the quantum domain. Recall from Section 3.7 that the classical noisy channel coding theorem applies to discrete memoryless channels, channels over a finite alphabet and with a finite number of states. The memoryless property of the communication channel implies that the output of the channel is a Markov process, it is affected only by the current input and not by the history of the channel states.

A discrete memoryless quantum channel transforms a quantum system whose state is a vector in a finite-dimensional Hilbert space; the memoryless property has the same meaning as for the classical information, the output of the quantum channel is only determined by the current input. Figure 80 captures the essential features of communication over a quantum channel; the input to a quantum channel is an ensemble of $n$ qubits in a pure state. An encoder maps $n$ input qubits into $m > n$ intermediate systems that may, or may not be qubits, and feeds them to $m$ independent instances of a quantum channel characterized by the superoperator[30] $\mathcal{E}$ describing the transformations of the input states. The output of each instance of the channel is generally a mixed state; the decoder transforms the output of the quantum channels into a mixed state of $n$ qubits.

**Quantum channels and the superoperator formalism.** To capture not only the essential aspects of quantum communication, but also the transformation of the quantum information during storage and processing we consider a more general view of a quantum channel as completely positive, trace preserving maps between spaces of operators.

The density matrix $\rho$ plays an important role in quantum information theory because it captures all the information that can be obtained by an observer about a mixed state and only that; any test to identify the state of a mixed system requires the density matrix. Indeed, if the mixed state comes from an ensemble of states $\{|\varphi_i\rangle : 1 \leq i \leq N\}$ with probabilities $\{p_i : 1 \leq i \leq N\}$, then to determine if the system is in state $|\psi\rangle$ we compute

[30] A superoperator is a linear operator, often positively-defined, acting on a space of linear operators.

$$\sum_{i=1}^{N} p_i \mid \langle \varphi_i \mid \psi \rangle \mid^2 = \mathrm{tr}\left[\rho \mid \psi\rangle\langle\psi \mid\right].$$

If a bipartite system $\mathcal{AB}$ is characterized by the density matrix $\rho$, then to determine if $\mathcal{A}$ is in state $\mid \psi\rangle$ we compute $\mathrm{tr}_{\mathcal{B}}\left[\rho \mid \psi\rangle\langle\psi \mid\right]$. The superoperator, $\mathcal{E}$, is *a trace-preserving completely-positive linear map from input-state density matrix $\rho$ to output-state density matrix $\rho'$, $\rho' = \mathcal{E}(\rho)$*. The superoperator formalism captures the interaction of a system with an environment initially in a pure state, and quantifies the transformation carried out by the channel encoder, the quantum channel, as well as the decoder.

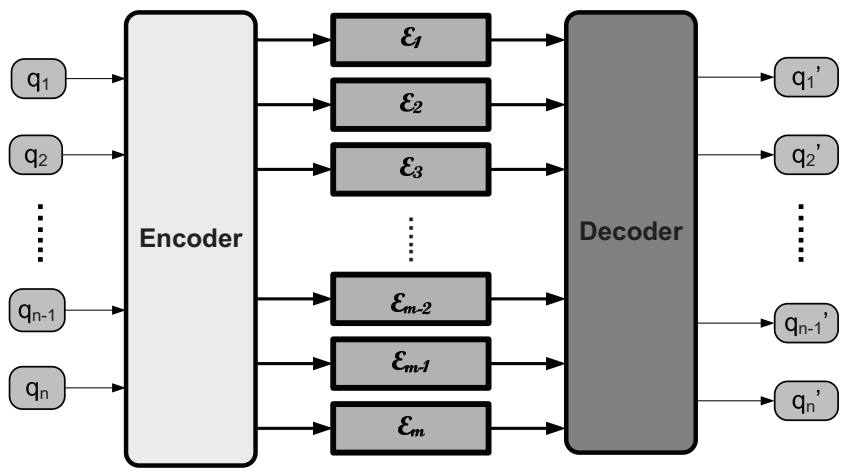



Figure 80: Communication through a quantum channel. The input is an ensemble of $n$ qubits in a pure state $\mid \varphi\rangle$ with the density matrix $\rho = \mid \varphi\rangle\langle\varphi \mid$. The encoder maps the $n$ input qubits into $m > n$ intermediate systems that may not be qubits, and feeds them to $m$ independent instances of a quantum channel characterized by the superoperator $\mathcal{E}$; the output of each instance of the channel is generally a mixed state with the density function of $\rho' = \mathcal{E}(\rho)$. Finally, the decoder transforms the output of the quantum channels into a mixed state of $n$ qubits with the density function $\rho''$. The fidelity of the communication system is $F = \langle\varphi \mid \rho'' \mid \varphi\rangle$.

**The capacity of a quantum channel.** There are at least four definitions of the capacity of a quantum channel based on the nature of the information transmitted (classical or quantum), the auxiliary resources allowed (classical channel, entanglement), and the protocols permitted [388]. The four distinct capacities are [47]: (i) $C$, the ordinary classical capacity for transmitting classical information; (ii) $Q$, the ordinary quantum capacity for transmitting quantum states, typically $Q < C$; (iii) $Q_2$, the classically-assisted quantum capacity for transmitting quantum states with the assistance of a two-way classical channel, $Q \leq Q_2 \leq C$; and (iv) $C_E$, the entanglement-assisted classical capacity, $Q \leq C \leq C_E$.

The capacity $Q$ of the quantum channel is the largest number $Q$ such that for any rate $R < Q$, $\epsilon > 0$, and block sizes $n$ and $m$, there exists an encoding procedure mapping $n$ qubits in a pure state $\mid \varphi\rangle$ with the density matrix $\rho = \mid \varphi\rangle\langle\varphi \mid$ into $m > n$ intermediate systems

that may not be qubits, and feeds them to $m$ independent instances of a quantum channel characterized by the superoperator $\mathcal{E}$; there should also be a decoding procedure mapping the $m$ channel outputs to $n$ qubits such that the original state $\mid \varphi\rangle$ can be recovered with a fidelity $F$ at least $1-\epsilon$.

**Quantum channel fidelity.** A quantum channel can be used to transport either classical or quantum information; the probability that the state at the output of the quantum channel would pass a test for being the same as the state at the input, measures the reliability of the quantum channel by a quantity called *channel fidelity*. The *fidelity of quantum states*, $F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right)$, introduced by Jozsa [221], provides a quantitative measure of the distinguishability of the two states, a measure of a channel fidelity. When $\mid \varphi^{\mathcal{A}}\rangle$ and $\mid \varphi^{\mathcal{B}}\rangle$ are pure states, the fidelity can be expressed as

$$F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = \mid \langle \varphi^{\mathcal{A}} \mid \varphi_{\mathcal{B}}\rangle \mid .$$

In the general case of mixed states the fidelity is

$$F\left(\rho^{\mathcal{A}}, \rho^{\mathcal{B}}\right) = \mathrm{tr}\left[\sqrt{(\rho^{\mathcal{A}})^{1/2}\rho^{\mathcal{B}}(\rho^{\mathcal{A}})^{1/2}}\right].$$

The analog of a classical discrete memoryless noisy channel is a quantum system that interacts unitarily with the environment while in transit from a source to a destination. If the source is an ensemble of pure or mixed states in a Hilbert space of finite dimension $d = 2^n$ with the density matrices $\{\rho_1, \rho_2, \ldots, \rho_{2^n}\}$ with probabilities $p_1, p_2, \ldots, p_{2^n}$, respectively, then, we can transmit $2^n$ classical messages as a sequence of $n$ qubits. If the states $\rho_i$ are orthogonal, we can obtain complete information about the source state by performing a measurement and then we are faced with the classical information transmission scenario. If the source states are non-orthogonal, then no measurement can extract complete information about the state thus, *the best we can achieve is to faithfully reproduce the input state at the receiving end.* Needless to say that in the non-orthogonal state case, the channel must let the quantum system pass through *without* learning anything about the state.

At the receiving end we have a similar problem when we wish to assess the quality of the transmission and eventually correct the errors. Quantum error correction is possible due to the following property of quantum systems: if the bipartite system $(\mathcal{A}, \mathcal{B})$ in state $\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}$ is subject to a transformation $\mathcal{E}$ and if there is an auxiliary system $\mathcal{C}$ in some standard state $\rho^{\mathcal{C}}$ and some unitary operator $\mathbf{U}$ acting on the joint system $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ and if

$$\mathcal{E}\left(\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}\right) = \mathrm{tr}_{\mathcal{C}}\left[\mathbf{U}\left(\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}} \otimes \rho^{\mathcal{C}}\right)\right]$$

then, after applying the transformation $\mathbf{U}$ we can perform a measurement on $\mathcal{C}$ that reveals how close the two states $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ are and then we ignore the auxiliary system $\mathcal{C}$ [21]. The systems $\mathcal{A}$ and $\mathcal{B}$ could be the input and the output of a quantum channel and $\mathcal{E}$ the transformation of the input state density carried out by the channel; then the process described above allows us to determine if the output is in error and to pinpoint the qubit(s) in error. We defer the in-depth discussion of the solution to this problem to Chapter 5, where we analyze quantum error correcting codes.

**Mappings between Hilbert spaces of different dimensions.** The analysis of communication through a quantum channel sometimes requires mappings between Hilbert spaces of different dimensions and the superoperator formalism allows such mappings by adding dummy dimensions to the smaller spaces. Stinespring's Dilation Theorem [412] provides the mathematical justification for mappings between Hilbert spaces of different dimensions. A brief review of basic concepts related to such mappings follows.

If $A$ is a unital $C^*$-algebra and $B(\mathcal{H})$ are the bounded operators on the Hilbert space $\mathcal{H}$, then for every completely positive map $T : A \mapsto B(\mathcal{H})$ there exists another Hilbert space $\mathcal{K}$ and a unital *-homomorphism[31] $L : A \mapsto B(\mathcal{K})$, such that

$$T(a) = VL(a)V^*, \quad \forall a \in A$$

with $V : \mathcal{K} \mapsto \mathcal{H}$ a bounded operator. Then:

$$|| \, T(1) \, || = || \, V \, ||^2 \, .$$

This result from operator theory [412] shows that a completely positive map in a $C^*$-algebra can be represented as a composition of two completely positive maps.

An important consequence of this result for quantum information theory is that *completely positive and trace-preserving quantum operations can be considered as a unitary evolution of a larger (dilated) system of a bounded dimension*, the church of larger Hilbert space according to John Smolin. The interpretation of Stinespring's Dilation Theorem is that any completely positive and trace-preserving map in a Hilbert space (a quantum channel) can be constructed from three operations:

1. Tensoring the input with a second Hilbert system (the ancilla system) in a specified state;

2. A unitary transformation on the larger (*dilated*) space (input plus ancilla system) obtained as a result of step (1); and

3. Reduction to a subsystem.

**Theorem.** *Let $\mathcal{S}_\mathcal{H}$ be the set of states of the Hilbert space $\mathcal{H}$ and $T$ be a completely positive and trace-preserving map, $T : \mathcal{S}_\mathcal{H} \mapsto \mathcal{S}_\mathcal{H}$. Then there exist a Hilbert space $\mathcal{K}$ and a unitary operation $U$ on $\mathcal{S}_\mathcal{H} \otimes \mathcal{S}_\mathcal{K}$, with $\mathcal{S}_\mathcal{K}$ the set of states of the Hilbert space $\mathcal{K}$, such that*

$$T(\rho) = tr_\mathcal{K} U(\rho \otimes \mid 0\rangle\langle 0 \mid) U^\dagger$$

*for all states $\rho \in \mathcal{S}_\mathcal{H}$, where $tr_\mathcal{K}$ denotes the partial trace on the $\mathcal{K}$ system. The dimension of the Hilbert space $\mathcal{K}$ is bounded*

[31] A $^*-$homomorphism $f : A \mapsto B$ is algebra homomorphism compatible with involutions of $A$ and $B$, that is

$$f(a^*) = f(a)^*, \quad \forall a \in A.$$

A homomorphism is unital if $f(1) = 1$

$$\dim \mathcal{K} \leq \dim^2 \mathcal{H}.$$

This is a version of a more general Stinespring's Dilation Theorem restricted to completely positive and trace-preserving maps and finite-dimensional systems; moreover it assumes that the dimensions of the input and output spaces for the mapping $T$ are the same. The triplet $(L, V, \mathcal{K})$ is called a *Stinespring representation* of $T$.

In Stinespring's representation the ancilla system has a natural interpretation as the environment of the physical system under investigation. The output of the channel $T$ arises from a unitary interaction of the input state with the environment, followed by a partial trace over the degrees of freedom of the environment. Almost all information can be extracted from the output channel through a decoding operation if and only if the channel $T$ releases almost no information to the environment. For states, i. e., channels with one-dimensional domain, dilations are called *purifications*.

In summary, a quantum channel is a completely positive trace-preserving linear map; according to Stinespring Dilation Theorem this map can be described by a unitary transformation followed by a partial trace. In the next section we discuss the quantum erasure phenomena.

## 3.16 Quantum Erasure

In our extensive discussion of quantum measurements in Chapter 2 we argue that the density matrix of a system characterizes all possible measurements performed on that system. The quantum erasure experiment discussed in this section raises the question if indeed the density matrix provides a complete description of the state of a subsystem of a composite system.

We first review the concept of interference and Young's double-slit experiment discussed in many textbooks including [284]; then we introduce the coherent and the incoherent superposition states of a quantum system. The term <u>interference</u> refers to the interaction of optic, acoustic, or electromagnetic waves that are correlated, or coherent with each other; interference occurs if the waves come from the same source, or if the waves have about the same frequency. Let us consider a transition between two quantum states $\mid \varphi\rangle, \mid \psi\rangle \in \mathcal{H}_n$ and let $\{\mid e_i\rangle\}$ be an orthonormal basis in $\mathcal{H}_n$, then,

$$\mid \varphi\rangle = \sum_i \alpha_i \mid e_i\rangle \quad \mapsto \quad \mid \psi\rangle = \sum_i \beta_i \mid e_i\rangle$$

As we know from Section 1.7, in Quantum Mechanics the probability of this transition is the square of the modulus of the inner product of the two states

$$\text{Prob}\,(\mid \varphi\rangle \mapsto \mid \psi\rangle) = \mid \langle \varphi \mid \psi\rangle \mid^2 = \mid \sum_i \alpha_i^* \beta_i \mid^2 = \sum_{i,j} \alpha_i^* \alpha_j \beta_j^* \beta_i.$$

The last sum can be separated into two terms thus, the expression becomes

$$\text{Prob}\,(\mid \varphi\rangle \mapsto \mid \psi\rangle) = \sum_i \mid \alpha_i \mid^2 \mid \beta_i \mid^2 + \sum_{i,j,i\neq j} \alpha_i^* \alpha_j \beta_j^* \beta_i.$$

The first term corresponds to a *classical interference* while the second term reflects the *quantum interference* between $i \neq j$ paths of quantum evolution; this is the quantum probability rule.

A qubit could be in a superposition state and we distinguish two types of superposition states, coherent and incoherent. A quantum system in state $\mid \varphi\rangle = \sum_i \alpha_i \mid e_i\rangle \in \mathcal{H}_n$ is in a coherent superposition of basis states $\mid e_i\rangle$ if its density matrix $\rho$ is not diagonal, $\rho \neq (1/n)I_n$. If, in addition, the system is in a pure state then it is said to be *completely coherent*.

If $\rho$ is diagonal, $\rho = (1/n)I_n$, then the system is in an incoherent superposition of states. In particular, a qubit is in an incoherent superposition of states if

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In case of an incoherent superposition, the basis states are prepared independently and there is no definite phase relationship between the basis states, no interference is possible.

A measurement destroys coherence; in his Lecture Notes in Physics [149] Feynman gives a clear example of the interference and how a measurement destroys coherence in a Young's experiment with electrons. When we conduct the experiment without attempting to learn which hole an electron passes through, we observe an interference pattern on the recording media positioned behind the screen with the two holes. We can identify the position of the electrons by illuminating the electrons with a laser beam; the electrons interact with the laser beam and emit a quanta of light allowing us to know precisely the hole each electron is passing through. Now the coherence of the electrons is destroyed and we observe a Gaussian distribution on the recording media.

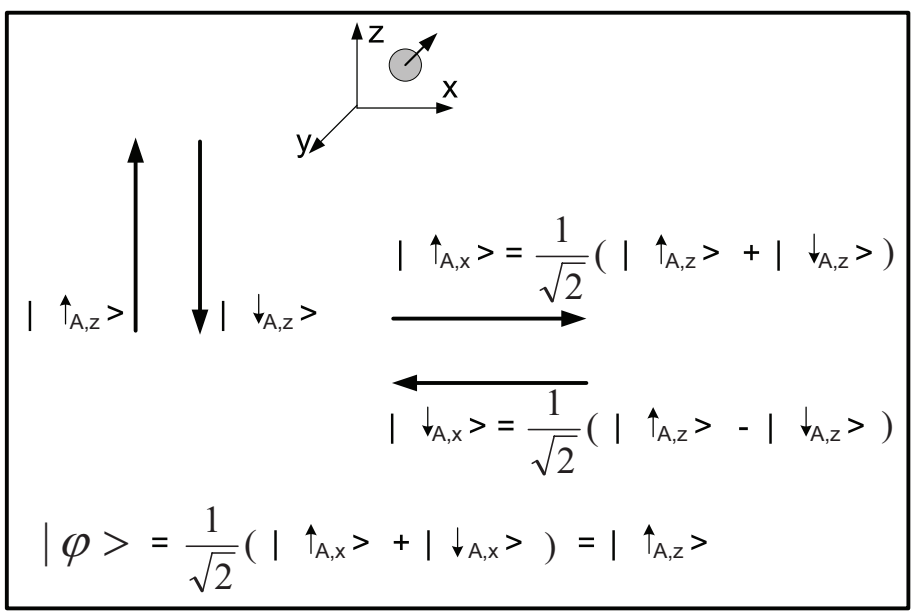



Figure 81: A superposition state $\mid \varphi\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{A,x}\rangle + |\downarrow_{A,x}\rangle) = |\uparrow_{A,z}\rangle$ of a spin one-half particle $A$ where $|\uparrow_{A,x}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{A,z}\rangle + |\downarrow_{A,z}\rangle)$ and $|\downarrow_{A,x}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_{A,z}\rangle - |\downarrow_{A,z}\rangle)$. The projection of the spin of particle $A$ along the $x$-direction is in a coherent superposition state.

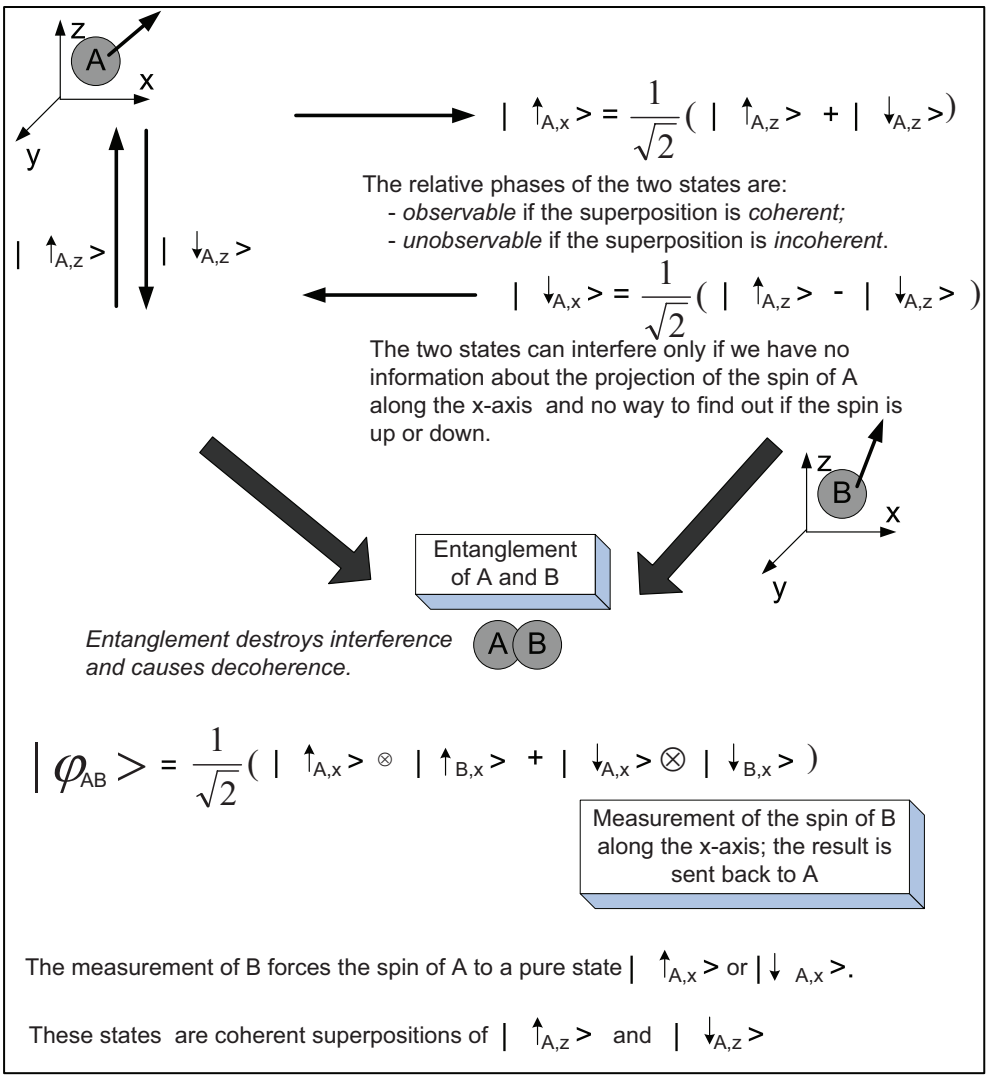



Figure 82: Entanglement destroys interference and causes decoherence. A measurement could restore a coherent superposition.

To discuss the quantum erasure we consider qubits, embodied by 1/2-spin particles, e.g., electrons. The projections of the spin of particle $A$ along the $x, y$, and $z$ directions can be up or down:

$$(|\uparrow_{A,x}\rangle, |\downarrow_{A,x}\rangle)\,, (|\uparrow_{A,y}\rangle, |\downarrow_{A,y}\rangle,)\,(|\uparrow_{A,z}\rangle, |\downarrow_{A,z}\rangle)\,.$$

The orientation of the spin can be measured with a Stern-Gerlach apparatus oriented along one of the three directions. For example, when the Stern-Gerlach apparatus is oriented along the $x$ direction it will measure the projection of the spin along $z$, the direction of the magnetic field. If we denote the basis states of a 1/2-spin system along a particular direction, spin up and spin down as $|\,0\rangle$ and $|\,1\rangle$, respectively, then any superposition state can be expressed as $|\,\varphi\rangle = \alpha_0\,|\,0\rangle + \alpha_1\,|\,1\rangle$. The probability amplitudes $\alpha_0$ and $\alpha_1$ encode the probability of an outcome in the $(|\,0\rangle, |\,1\rangle)$ basis; the relative phases of $\alpha_0$ and $\alpha_1$ are also significant. The two

complex numbers $\alpha_0$ and $\alpha_1$ describe the *3D* orientation of the spin; if $\theta$ is the polar angle and $\phi$ the azimuthal angle then: $\alpha_0 = e^{-i\phi/2}\cos(\theta/2)$ and $\alpha_1 = e^{+i\phi/2}\sin(\theta/2)$. The state $|\,\varphi\rangle$ can be interpreted as a spin pointing in the $(\theta, \phi)$ direction.

If identical copies of the 1/2-spin system are available then the two complex numbers $\alpha_0$ and $\alpha_1$ cannot be determined with a single measurement, or even with any number of measurements, *carried along only one* of the $x$, $y$, or $z$ directions. A vector in a *3D* space can only be determined if we know the projections along all three directions. We know by this time that unitary transformations of a qubit are rotations and there is a transformation that allows us to rotate the spin pointing in the $(\theta, \phi)$ direction to, say, the $z$ direction; in that case only one measurement will be sufficient to determine not only the modulus of the two complex numbers, but also the relative phase between them.

Consider two orthogonal superposition states when the projections of the spin of $A$ in the $x$ direction are up and down, respectively, as in Figure 81:

$$|\uparrow_{A,x}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{A,z}\rangle + |\downarrow_{A,z}\rangle\right) \quad \text{and} \quad |\downarrow_{A,x}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{A,z}\rangle - |\downarrow_{A,z}\rangle\right).$$

A measurement of the projection of the spin along direction $x$ results in one of the two eigenvectors $|\uparrow_{A,x}\rangle$ and $|\downarrow_{A,x}\rangle$ with probability 1/2.

Now we consider the superposition state

$$|\,\varphi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{A,x}\rangle + |\downarrow_{A,x}\rangle\right).$$

When we measure the projection of the spin of the qubit in state $|\,\varphi\rangle$ along the $x$ direction we observe one of the two eigenvectors $|\uparrow_{A,x}\rangle$ and $|\downarrow_{A,x}\rangle$ with probability 1/2. If we measure the spin of the qubit in state $|\,\varphi\rangle$ along the $z$-direction we find $|\uparrow_{A,z}\rangle$ with probability 1. A simple analysis shows that we should not be surprised at all, $|\,\varphi\rangle$ is in fact $|\uparrow_{A,z}\rangle$. Indeed,

$$|\,\varphi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{A,x}\rangle + |\downarrow_{A,x}\rangle\right) = \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}\left(|\uparrow_{A,z}\rangle + |\downarrow_{A,z}\rangle\right) + \frac{1}{\sqrt{2}}\left(|\uparrow_{A,z}\rangle - |\downarrow_{A,z}\rangle\right)\right] = |\uparrow_{A,z}\rangle.$$

In case of a *coherent superposition* we are able to distinguish the relative phases of the two spin states of $A$; the two states can interfere with each other. Such a coherent superposition state is

$$|\uparrow_{A,x}, \downarrow_{A,x}\rangle = \frac{1}{2}\left(|\uparrow_{A,x}\rangle \pm |\downarrow_{A,x}\rangle\right).$$

The two pure states $|\uparrow_{A,z}\rangle$ and $|\downarrow_{A,z}\rangle$ can interfere only if we have no means to find out if the projection of the spin of $A$ along the $x$ direction is up or down.

An *incoherent superposition* state has a density matrix $\rho = 1/2I$. The density matrix of this maximally mixed state can be obtained through different preparations:

$$\rho = \frac{1}{2}\left(|\uparrow_{A,x}\rangle\langle|\uparrow_{A,x}\rangle + |\downarrow_{A,x}\rangle\langle|\downarrow_{A,x}\rangle\right) \quad \text{or} \quad \rho = \frac{1}{2}\left(|\uparrow_{A,z}\rangle\langle|\uparrow_{A,z}\rangle + |\downarrow_{A,z}\rangle\langle|\downarrow_{A,z}\rangle\right).$$

The relative phase of an incoherent superposition is totally unobservable.

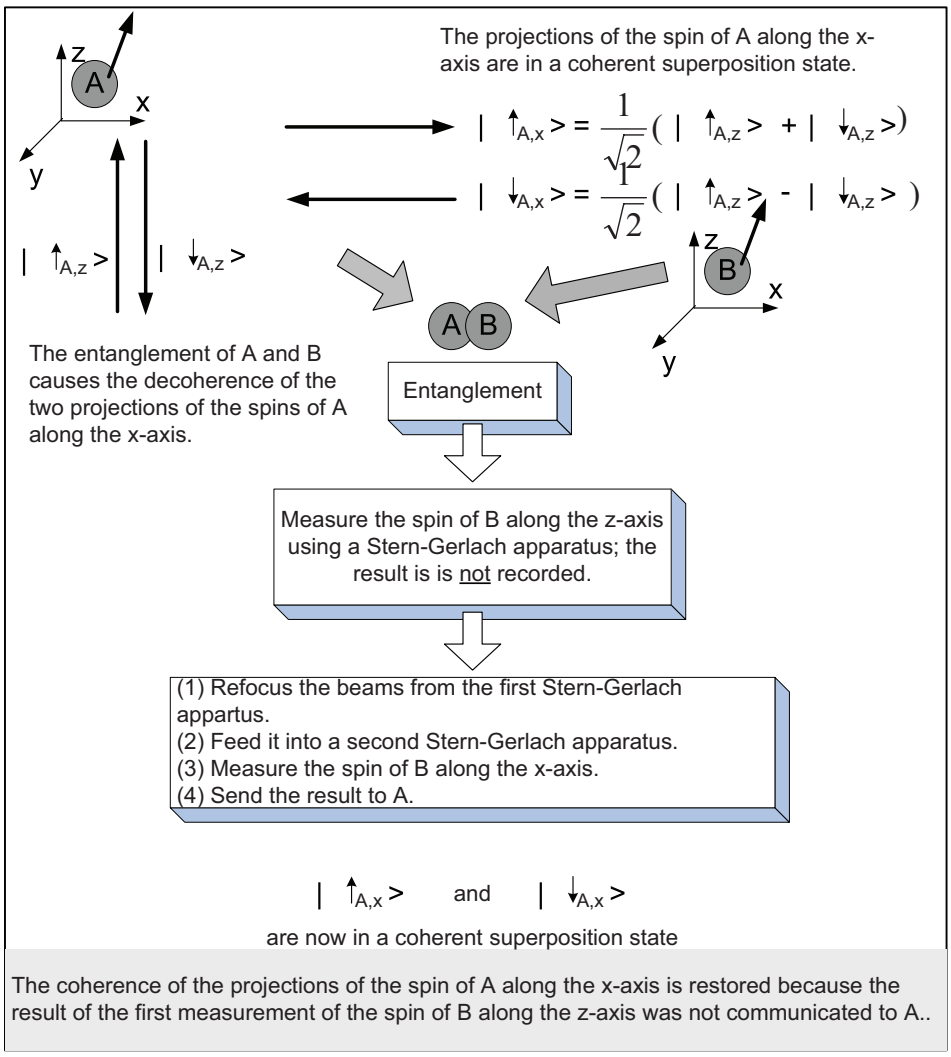



Figure 83: Quantum erasure. While Alice cannot observe the interference between $|\uparrow_{A,x}\rangle$ and $|\downarrow_{A,x}\rangle$ in the ensemble with the density matrix $\rho_A = (1/2)I$, once she gets the results of the measurement along the $z$ direction, this ensemble is forced into a particular preparation, while the density matrix remains the same, $\rho_A = (1/2)I$. Only after receiving the information from Bob, Alice can select a subensemble of her spins that are in the pure state $|\downarrow_{A,z}\rangle$. The information allows Alice to distill *purity* along the $z$-direction from a maximally mixed state along the $x$-direction.

Consider now two qubits $A$ and $B$ and their entangled state, Figure 82,

$$|\,\varphi_{AB}\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_{A,x}\rangle \otimes |\uparrow_{B,x}\rangle + |\downarrow_{A,x}\rangle \otimes |\downarrow_{B,x}\rangle\right).$$

To determine the orientation of the spin of $A$ we perform a measurement on $B$ along the direction $A$ is oriented. For example, Bob could measure the spin of $B$ along $x$ and find out that the state is $|\uparrow_{B,x}\rangle$; this measurement produces a decoherence of $A$ along the $x$-direction, it forces it to choose a certain orientation of the spin, e.g., $|\downarrow_{A,x}\rangle$. The result of the measurement

is sent back to Alice, the preparer of $A$; once this message is received, $A$ is again in a pure state, either $|\uparrow_{A,x}\rangle$, or $|\downarrow_{A,x}\rangle$.

Now we discuss the following experimental setup, Figure 83: Bob feeds particle $B$ into a Stern-Gerlach apparatus oriented to measure the projection of the spin along the $x$ direction and then, without recording the result of the first measurement, refocusses the two emerging beams into a second Stern-Gerlach apparatus oriented for a measurement along the $z$ direction. Bob then communicates to Alice if the projection of the spin of $B$ along the $z$ direction is up or down; after this message reaches Alice the coherence of the states $|\uparrow_{A,x}\rangle$ and $|\downarrow_{A,x}\rangle$ of $A$ is restored. This phenomenon is called quantum erasure; the entanglement causes the loss of coherence between $|\uparrow_{A,z}\rangle$ and $|\downarrow_{A,z}\rangle$ after the first measurement and then, the second measurement along $z$ erases this information and restores the coherence along the $x$-direction.

The explanation is that while Alice cannot observe the interference between $|\uparrow_{A,x}\rangle$ and $|\downarrow_{A,x}\rangle$ in the ensemble with the density matrix $\rho_A = (1/2)I$, once she gets the results of the measurement along the $z$ direction, this ensemble is forced into a particular preparation, while the density matrix remains the same, $\rho_A = (1/2)I$. Only after receiving the information from Bob, Alice can select a subensemble of her spins that are in the pure state $|\downarrow_{A,z}\rangle$. The information allows Alice to distill *purity* along the $z$-direction from a maximally mixed state along the $x$-direction.

Preskill argues [343] that: "the state $\rho_A$ of system $A$ is not the same as $\rho_A$ accompanied by the information Alice has received from Bob." The information provided by Bob about particle $B$ changes the physical description of particle $A$; the state of particle $A$ is described by $\rho_A = (1/2)I$ + "the state of knowledge."

Quantum erasure illustrates that the behavior of particles carrying quantum information can be non-intuitive [343]. The experiments discussed in this section show that *the entanglement destroys interference and causes decoherence and that a measurement could restore a coherent superposition.*

## 3.17 Classical Information Capacity of Noiseless Quantum Channels

We turn our attention to noiseless quantum channels with pure states as input and analyze their capacity to transmit classical information. Recall that the capacity of a classical channel was defined as the highest rate the information can be transmitted through the channel, the maximum of the mutual information between the input $X$ and output $Y$ of the channel, $C = \max I(X;Y)$. The mutual information depends on the fixed conditional probabilities $p_{Y|X}(y|x)$ that quantify the effect of the noise on the letters transmitted by the source; the measurement process for classical information is trivial, a letter received is immediately identified.

There is a significant difference between the capacity of classical and quantum channels models discussed in this section: when the quantum channel is noiseless, the quantum states are transmitted unaltered, but, in the general case, they are non-orthogonal and the recipient of quantum information has a choice on how to perform a measurement to identify the state transmitted by the source; the conditional probabilities arise from the probabilistic nature of quantum measurements.

We wish to determine the maximum rate at which classical information may be transmitted through a noiseless quantum channel when we encode the information into $n$ physical qubits. We consider an ensemble of letter states $\mathcal{A} = \{\mid a_k\rangle\}$ generated with a priori probability $p_{a_k}$. The communication channel is characterized by the letter states it can transmit.

Call $\eta_k$ the *overall frequency* of letter $\mid a_k\rangle$ computed over the entire set of letters used (the cardinality of this set is $n \times N$). If $\nu_{ik}$ is the number of occurrences of letter $\mid a_k\rangle$ in codeword $\mid w_i\rangle$ then:

$$\eta_k = \frac{1}{n} \sum_i p_{w_i} \nu_{ik}.$$

We expect that $\eta_k \approx p_{a_k}$ and we define the *tolerance* of the code $\mathcal{C}^k$ as

$$\tau_{\mathcal{C}^k} = \max_{a_k} \mid \eta_k - p_{a_k} \mid .$$

**Theorem.** *The noiseless quantum channel capacity. Given: (i) the source of quantum information $\mathcal{A}$ with the density matrix $\rho^{\mathcal{A}}$ and von Neumann entropy $S(\rho^{\mathcal{A}})$; (ii) an arbitrarily small number $\delta > 0$. Let $I_\delta$ be the least upper bound on the information per letter transmissible with any code $\mathcal{C}^k$ such that its tolerance is $\tau_{\mathcal{C}^k} \leq \delta$. Then*

$$\lim_{\delta \mapsto 0} I_\delta = S(\rho^{\mathcal{A}}).$$

The proof of this theorem can be found in [205]. If we relax the restriction that letters $\mid a_k\rangle$ are generated with a fixed a priori input probability distribution $p_{a_k}$ and consider this distribution to be variable we can define the capacity of the quantum channel as the maximum von Neumann entropy

$$C = \max_{p_{a_k}} S(\rho^{\mathcal{A}}).$$

This theorem shows that $C$ is the maximum rate at which classical information can be transmitted through a quantum channel and no more. That is formally similar to the statement of the classical channel capacity theorem. Classical channel capacity involves maximizing over the input distribution. Quantum channel capacity implies optimizing over choice both of input distribution and of decoding observable.

We now present a formulation of Schumacher quantum noiseless coding theorem which resembles more closely Shannon source coding theorem [369].

**Theorem.** *Let $\mathcal{A}$ be a quantum source with a letter ensemble described by the density operator $\rho^{\mathcal{A}}$ and let $(\delta, \epsilon) \geq 0$ some arbitrarily small numbers.*

*(i) Suppose that $S(\rho^{\mathcal{A}}) + \delta$ qubits are available per $\mathcal{A}$ letter. Then for sufficiently large $n$, groups of $n$ letters from the source $\mathcal{A}$ can be transposed via the available qubits with fidelity $F > 1 - \epsilon$.*

*(ii) Suppose that $S(\rho^{\mathcal{A}}) - \delta$ qubits are available per $\mathcal{A}$ letter. Then for sufficiently large $n$, groups of $n$ letters from the source $\mathcal{A}$ can be transposed via the available qubits with fidelity $F < \epsilon$.*

*Proof:* The proof is similar with the one for classical information. When $\rho^{\mathcal{A}}$ is the density matrix of the source $\mathcal{A}$, then a group of $n$ physical qubits emitted by the source is characterized by the density matrix $(\rho^{\mathcal{A}})^{\otimes n}$ in $(\mathcal{H}_2)^{\otimes n}$. The Hilbert space $(\mathcal{H}_2)^{\otimes n}$ is partitioned into two orthogonal subspaces, $\Lambda_{typical}$ and $\Lambda_{typical}^{\perp}$ [205] such that:

1. $\Lambda_{typical}$ and $\Lambda_{typical}^{\perp}$ are spanned by the eigenstates of the density matrix $(\rho^{\mathcal{A}})^{\otimes n}$;
2. If $\Pi_{\Lambda}$ and $\Pi_{\Lambda^{\perp}}$ are the projectors onto $\Lambda_{typical}$ and $\Lambda_{typical}^{\perp}$ then,
$$\text{tr}\left[\Pi_{\Lambda}(\rho^{\mathcal{A}})^{\otimes n}\Pi_{\Lambda}\right] > 1-\epsilon \qquad \text{and} \qquad \text{tr}\left[\Pi_{\Lambda^{\perp}}(\rho^{\mathcal{A}})^{\otimes n}\Pi_{\Lambda^{\perp}}\right] < \epsilon.$$
3. The eigenvalues $\lambda_k$ of $(\rho^{\mathcal{A}})^{\otimes n}$ corresponding to the eigenstates of $\Lambda_{typical}$ satisfy the inequality
$$2^{-n\left[S(\rho^{\mathcal{A}})+\delta\right]} \leq \lambda_k \leq 2^{-n\left[S(\rho^{\mathcal{A}})-\delta\right]}.$$
4. The dimension of the typical subspace is bounded:
$$(1-\epsilon)2^{-n\left[S(\rho^{\mathcal{A}})-\delta\right]} \leq \dim\left[\Lambda_{typical}\right] \leq 2^{-n\left[S(\rho^{\mathcal{A}})+\delta\right]}.$$

The typical subspace $\Lambda_{typical}$ consists of "typical" eigenvectors of the density matrix $(\rho^{\mathcal{A}})^{\otimes n}$; these are eigenvectors of $(\rho^{\mathcal{A}})^{\otimes n}$ with frequency close to the probability distribution of the eigenvalues of the density operator $\rho^{\mathcal{A}}$, Figure 84.

The recipient of information must design an optimal POVM measurement to identify the codeword $\mid w_i\rangle$ sent by the source of quantum information. Call $\mid \zeta_k\rangle \in \mathcal{H}_2^{\otimes n}$ the set of vectors chosen to specify the POVM and define the positive operator whose support is the subspace spanned by $\mid \zeta_k\rangle$

$$\mathbf{M} = \sum_k \mid \zeta_k\rangle\langle\zeta_k \mid .$$

Then $\mathbf{M}^{1/2}$ exists and it is invertible in the subspace spanned by $\mid \zeta_k\rangle$ and allows us to define the vectors

$$\mid \mu_k\rangle = \mathbf{M}^{-1/2} \mid \zeta_k\rangle.$$

It is easy to prove that the corresponding positive operators $\mid \mu_k\rangle\langle\mu_k \mid$ sum up to identity on this subspace

$$\sum_k \mid \mu_k\rangle\langle\mu_k \mid = \sum_k \mathbf{M}^{-1/2} \mid \zeta_k\rangle\langle\zeta_k \mid \mathbf{M}^{-1/2} = \mathbf{M}^{-1/2}\left(\sum_k \mid \zeta_k\rangle\langle\zeta_k \mid\right)\mathbf{M}^{-1/2} = 1.$$

These operators are the outcome operators of a POVM which is called a "pretty good measurement"; this POVM is used by the receiver to distinguish among the vectors $\mid \zeta_k\rangle$. A variation of it will also be used in Section 3.22. When the codewords $\mid w_i\rangle \in \mathcal{C}$ are sent with an equal probability $p_i = 1/N$ with $N$ equal to the cardinality of the code $\mathcal{C}$, $N = \mid \mathcal{C} \mid$, then

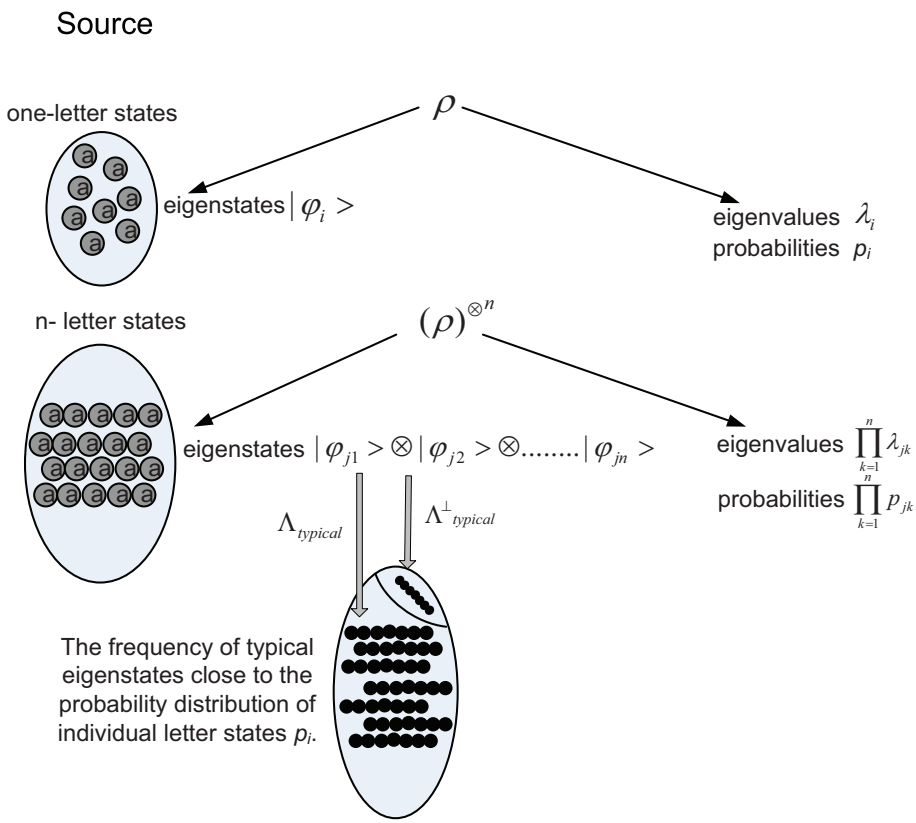



Figure 84: The construction of "typical subspaces" for Schumacher quantum noiseless encoding. The single letter source has the density matrix $\rho$ with eigenvectors $\mid \varphi_i\rangle$ and eigenvalues $\lambda_i$ with probability $p_i$. The quantum code encodes a single logical qubit into $n$ physical ones. The eigenvectors of $\rho^{\otimes n}$, the density matrix of $n$-letter codewords are tensor products of the eigenvectors of $\rho$, namely $\mid \varphi_{j1}\rangle \otimes \mid \varphi_{j2}\rangle \otimes \ldots \otimes \mid \varphi_{jn}\rangle$. The "typical subspace" $\Lambda_{typical}$ is spanned by the "typical eigenvectors" of the density matrix $\rho^{\otimes n}$. The "typical eigenvectors" of $\rho^{\otimes n}$ have a frequency close to the probability distribution of the eigenvalues of the density operator $\rho$; indeed, each eigenvector is associated with one eigenvalue and "almost" all eigenvectors are typical.

the recipient of information performs the POVM measurement and obtains the outcome $\mid \mu_i\rangle$ with probability

$$p\left[(\mid \mu_i\rangle)|(\mid w_i\rangle)\right] = \mathrm{tr}\left[\mid \mu_i\rangle\langle\mu_i \mid w_i\rangle\langle w_i \mid\right] = \mid \langle\mu_i \mid w_i\rangle \mid^2 .$$

Then the expected probability of error is

$$p_{err} = 1 - \sum_i \frac{1}{N} \mid \langle\mu_i \mid w_i\rangle \mid^2 = \frac{1}{N}\sum_i \left(1 - \langle\mu_i \mid w_i\rangle\right)\left(1 + \langle\mu_i \mid w_i\rangle\right).$$

The modulus of the inner products between the codeword sent, $\mid w_i\rangle$, and the outcome of the measurement, $\mid \mu_i\rangle$, is a measure of the error of the POVM measurement for $\mid w_i\rangle$

$$\langle \mu_i \mid w_i \rangle = \begin{cases} 0 & \text{if the vectors are orthogonal} \\ 1 & \text{if the vectors are identical} \\ > 0 \text{ and } < 1 & \text{otherwise} \end{cases}$$

The inner product is at most equal to unity, thus $(1 + \langle \mu_i \mid w_i \rangle) \leq 2$. It follows that

$$p_{err} \leq \frac{2}{N} \sum_i (1 - \langle \mu_i \mid w_i \rangle) .$$

The recipient of information employs an observable to distinguish between the signal states/codewords that, with very small exceptions, are typical sequences; therefore, instead of the signal states $\mid w_i \rangle$ we consider $\mid \sigma_i \rangle$, the projection of these states on the typical subspace $\Lambda_{typical}$:

$$\mid \sigma_i \rangle = \Pi_\Lambda \mid w_i \rangle.$$

The following diagram illustrates the relationship between the codeword, its projection, and the outcome of the POVM measurement:

| | | |
|---|---|---|
| $\mid w_i \rangle$ - codeword | $\Longrightarrow$ | $\mid \mu_i \rangle$ - result of POVM |
| $\Downarrow$ | | |
| $\mid \sigma_i \rangle = \Pi_\Lambda \mid w_i \rangle$ - projection of $\mid w_i \rangle$ on $\Lambda_{typical}$ | | |

Let $S$ be the complex $N \times N$ matrix $S = [S_{ij}]$ with positive eigenvalues; the elements of the matrix $S$ are inner products of the $N$ projections of $\mid w_i \rangle$ on $\Lambda_{typical}$, where $S_{ij} = \langle \sigma_i \mid \sigma_j \rangle$. The vectors $\mid \mu_k \rangle$ which are contained in the subspace $\Lambda_{typical}$ are constructed such that

$$\langle \mu_i \mid w_j \rangle = \langle \mu_i \mid \sigma_j \rangle = (\sqrt{S})_{ij}.$$

The vectors $\mid \mu_k \rangle$ together with the projection onto the subspace perpendicular to all the vectors $\mid \sigma_i \rangle$ define the POVM of the receiver.

The expected probability of error can be expressed as

$$p_{err} \leq \frac{2}{N} \sum_i \left[ 1 - (\sqrt{S})_{ii} \right] .$$

The following inequality is known: when $x \geq 0$ then the square root function has a parabola as a lower bound $\sqrt{x} \geq 1/2(3x - x^2)$. It follows that $\sqrt{S} \geq 1/2(3S - S^2)$ for the matrix $S$ with non-negative eigenvalues. Call the norm of the projected codewords $\kappa_i = \langle \sigma_i \mid \sigma_i \rangle = S_{ii}$. Then,

$$(\sqrt{S})_{ii} \geq \frac{3}{2}\kappa_i - \frac{1}{2}\kappa_i^2 - \frac{1}{2} \sum_{j \neq i} S_{ij} S_{ji}.$$

The expected probability of error for code $\mathcal{C}^k$ can now be written as

$$p_{err} \leq \frac{2}{N} \sum_i \left( 1 - \frac{3}{2}\kappa_i + \frac{1}{2}\kappa_i^2 + \frac{1}{2} \sum_{j \neq i} S_{ij} S_{ji} \right) = \frac{2}{N} \sum_i \left( (1 - \kappa_i)(1 - \frac{\kappa_i}{2}) + \frac{1}{2} \sum_{j \neq i} S_{ij} S_{ji} \right)$$

or,

$$p_{err} \leq \frac{2}{N} \sum_i \left( (1-\kappa_i) + \frac{1}{2} \sum_{j \neq i} S_{ij} S_{ji} \right).$$

To prove Schumacher's noiseless coding theorem we have to show that when $N$ is sufficiently large so that $\log N \approx nS(\rho^{\mathcal{A}})$, then there exists a code and the expected probability of error for this code becomes increasingly small, $p_{err} \mapsto 0$.

We consider a random code with $n$ qubit codewords; each codeword $\mid w_i\rangle$, $1 \leq i \leq N$, is a sequence of $n$ letter states generated using the a priori probabilities for the letters. The probability of a codeword $\mid w_i\rangle = \mid a_1\rangle \mid a_2\rangle \ldots \mid a_n\rangle$ is $P_{w_i} = p(a_1)p(a_2)\ldots p(a_n)$. Each codeword is generated independently from the other codewords. Let $\mathcal{C} = \{\mathcal{C}^1, \mathcal{C}^2, \ldots, \mathcal{C}^k, \ldots,\}$ be the set of the random codes constructed by the procedure we just described. Now we compute averages over all such codes for various quantities, for example,

$$\overline{\mid w_i\rangle\langle w_i \mid} = (\rho^{\mathcal{A}})^{\otimes n}.$$

The average over all random codes of the expected probability of error is then

$$\overline{p_{err}} \leq \frac{2}{N} \sum_i \left( (1-\overline{\kappa_i}) + \frac{1}{2} \sum_{j \neq i} \overline{S_{ij} S_{ji}} \right) = 2(N-\overline{\kappa_i}) + (N-1)\overline{S_{ij}S_{ji}} < 2(N-\overline{\kappa_i}) + N\overline{S_{ij}S_{ji}}.$$

We have used the fact that the $N$ terms in the summation with index $i$ and the $(N-1)$ terms in the summation with index $j$ have the same value. Now we can compute the averages involved in the previous expression. First, recall that $\kappa_i = \langle \sigma_i \mid \sigma_i\rangle$ and $\mid \sigma_i\rangle = \Pi_\Lambda \mid w_i\rangle$, therefore,

$$\overline{\kappa_i} = \overline{\mathrm{tr}\left[\Pi_\Lambda \mid w_i\rangle\langle w_i \mid \Pi_\Lambda\right]} = \mathrm{tr}\left[\Pi_\Lambda \overline{\mid w_i\rangle\langle w_i \mid} \Pi_\Lambda\right] = \mathrm{tr}\left[\Pi_\Lambda (\rho^{\mathcal{A}})^{\otimes n} \Pi_\Lambda\right].$$

The projection of a codeword on the typical subspace $\Pi_\Lambda$ satisfies the inequality

$$\mathrm{tr}\left[\Pi_\Lambda (\rho^{\mathcal{A}})^{\otimes n} \Pi_\Lambda\right] \geq 1 - \epsilon.$$

Thus,

$$\overline{\kappa_i} \geq 1 - \epsilon.$$

Recall that $S_{ij} = \langle w_i \mid \sigma_j\rangle$. Thus,

$$S_{ij}S_{ji} = \langle w_i \mid \sigma_j\rangle\langle w_j \mid \sigma_i\rangle = \langle w_i \mid \Pi_\Lambda \mid w_j\rangle\langle w_j \mid \Pi_\Lambda \mid w_i\rangle.$$

Then:

$$\begin{aligned} \overline{S_{ij}S_{ji}} &= \overline{\langle w_i \mid \Pi_\Lambda \mid w_j\rangle\langle w_j \mid \Pi_\Lambda \mid w_i\rangle} = \overline{\mathrm{tr}\left[\Pi_\Lambda \mid w_i\rangle\langle w_i \mid\mid w_j\rangle\langle w_j \mid \Pi_\Lambda\right]} \\ &= \mathrm{tr}\left[\Pi_\Lambda \left((\rho^{\mathcal{A}})^{\otimes n}\right)^2 \Pi_\Lambda\right]. \end{aligned}$$

It follows that the average over all random codes of the expected probability of error is

$$\overline{p_{err}} < 2(N - \overline{\kappa_i}) + N\overline{S_{ij}S_{ji}} < 2\epsilon + N\text{tr}\left[\Pi_\Lambda \left((\rho^{\mathcal{A}})^{\otimes n}\right)^2 \Pi_\Lambda\right].$$

If we sum the eigenvalues of $(\rho^{\mathcal{A}})^{\otimes n}$ only over the typical subspace $\Lambda_{typical}$ we obtain

$$\text{tr}\left[\Pi_\Lambda \left((\rho^{\mathcal{A}})^{\otimes n}\right)^2 \Pi_\Lambda\right] < (\dim\left[\Lambda_{typical}\right]) \left(2^{-n(S(\rho^{\mathcal{A}})-\delta)}\right)^2 < 2^{-n(S(\rho^{\mathcal{A}})-3\delta)}.$$

We have taken into account the fact that $(\dim\left[\Lambda_{typical}\right]) \leq 2^{n\left[S(\rho^{\mathcal{A}})+\delta\right]}$. Finally,

$$\overline{p_{err}} < 2\epsilon + N2^{-n(S(\rho^{\mathcal{A}})-3\delta)}.$$

From the family of codes $\mathcal{C}$ we can then select one, $\mathcal{C}^k$, such that

$$p_{err} < 2\epsilon + N2^{-n(S(\rho^{\mathcal{A}})-3\delta)}.$$

This concludes the proof of quantum noiseless coding theorem.
The last relationship shows that if $n$ is very large, much larger than needed to form the typical subspace $\Lambda$, the sender can make $N$ as large as $N = 2^{n(S(\rho^{\mathcal{A}})-4\delta)}$ and still have an average probability of error $p_{err} < 3\epsilon$. In this case the sender would use $S(\rho^{\mathcal{A}}) - 4\delta$ bits to encode a letter. Therefore, the existence of codes that allow the transmission at an asymptotic rate of $S(\rho^{\mathcal{A}}$ bits per letter with arbitrarily low error can be proven without actually construction such a code.

In the proof just presented it was assumed [205] that the sender codewords $\mid w_i\rangle$ are pure states and it was shown that for an appropriate choice of the sender's code and of the receiver decoding observable, for large enough $n$, codewords of $n$ letters may be used to transmit $nS(\rho_{\mathcal{A}})$ bits of information (or $S(\rho_{\mathcal{A}})$ bits per letter) with arbitrarily low probability of error. The theorem was also proved for codewords that are mixed states [374]; the general procedure is that of [205] using averages over random codes and a similar prescription for the decoding POVM observables, but enforcing stronger typicality conditions on various quantities associated with the channel. Since the mixed state codewords may be interpreted as the outputs of a noisy channel, the result is applied to the case of classical information capacity of noisy channels discussed later on in this chapter.

The analysis of noisy quantum channels requires the characterization of the relationship between the states at the input and the output of the channel. Next, we introduce the concepts of entropy exchange, entanglement fidelity, and coherent information necessary for the study of noisy quantum channels.

## 3.18 Entropy Exchange, Entanglement Fidelity, and Coherent Information.

Classical error correction uses a syndrome to identify the bit in error of a codeword and reverses it. In this section we discuss the theoretical underpinning of quantum error correction, namely the process to determine the error syndrome, which identifies the position of the qubits in error and the type of error; this process, analyzed in Chapter 5, consists of a measurement of a register of ancilla qubits entangled with the qubits of the codeword.

Quantum error correction requires a multi-qubit measurement which does not disturb the quantum information carried by the signaling qubits and retrieves information about the error. The outcome of the measurement tells us not only which physical qubit was affected, but also, the way it was affected; the error could be either a bit flip, or a phase flip, or both, corresponding to the Pauli matrices X, Z, and Y, respectively. The measurement of the syndrome has a projective effect; even if the error due to the noise was arbitrary, it can be expressed as a superposition of basis operations given by the Pauli matrices and the identity matrix.

We show how an environment which interacts unitarily with the quantum system $\mathcal{Q}$ can be constructed and discuss the purification of the mixed state of the system $\mathcal{Q}$; purification of system $\mathcal{Q}$ in a mixed state means to identify a reference system $\mathcal{R}$ such that the composite system $\mathcal{P} = (\mathcal{Q}, \mathcal{R})$ is in a pure state. The concepts introduced in this section allow us to quantify the interactions between the system $\mathcal{Q}$, the environment $E$, and the reference system $\mathcal{R}$: the *entropy exchange* is a measure of the information exchanged between the system and the environment during the interaction with the environment; the *entanglement fidelity* is a measure of the overlap between the initial purification and the joint state of the system $\mathcal{Q}$ and the reference system $\mathcal{R}$ after the transmission of $\mathcal{Q}$. The *coherent information* measures the degree to which the quantum coherence is preserved by the superoperator $\mathcal{E}$ associated with a quantum channel, it is the correspondent of the mutual information.

An important concept related to quantum information processing is that of *trace-preserving operation* introduced in Section 2.14. Informally, a trace-preserving quantum transformation takes place when a quantum system interacts with the environment $E$ and no measurement is performed either on the quantum system, or the environment. If the operation is denoted by $\mathcal{E}$ and the quantum system $\mathcal{Q}$ is in a state characterized by the density matrix $\rho$, a trace-preserving operation requires that $\mathrm{tr}\,[\mathcal{E}(\rho)] = 1$. The linear transformation $\mathcal{E}(\rho)$ enjoys *complete positivity* [208] if there exist operators $\mathbf{A}_i$ such that

$$\mathcal{E}(\rho) = \sum_i \mathbf{A}_i \rho \mathbf{A}_i^\dagger \quad \text{with} \quad \sum_i \mathbf{A}_i^\dagger \mathbf{A}_i \leq I.$$

When $\sum_i \mathbf{A}_i^\dagger \mathbf{A}_i = I$ then $\mathrm{tr}\,[\mathcal{E}(\rho)] = 1$ for all density operators $\rho$ and the transformation is *trace-preserving*. It can be shown that the following three conditions are equivalent [257]:

(i) $\mathcal{E}$ is a trace-preserving, completely positive linear map on the density operators of $\mathcal{Q}$.

(ii) $\mathcal{E}$ has a unitary representation;

(iii) $\mathcal{E}$ has a normalized operator-sum representation.

When classical information about the quantum system is made available through a measurement, then the operation is *non trace-preserving*.

Transmission of information through a quantum communication channel implies a state change described by the superoperator $\mathcal{E}$. If $\rho$ is the density matrix of the *input state* then the density matrix of the *output state* is $\mathcal{E}(\rho)$. Figures 85(a) and (b) illustrate the transformation of the input state of the quantum channel when a closed system is subject to a unitary transformation and a general transformation, respectively. In case of the unitary transformation $\mathbf{U}$, the density operator of the channel output state is

$$\mathcal{E}(\rho) = \mathbf{U}\rho\mathbf{U}^\dagger.$$

In the case of a general transformation, the density operator of the normalized channel output state is

$$\frac{\mathcal{E}(\rho)}{\mathrm{tr}\left[\mathcal{E}(\rho)\right]}.$$

The trace in the denominator is included to preserve the trace condition for the input state $\mathrm{tr}(\rho) = 1$.

The question we pose is whether we can model a trace-preserving operation as a unitary operation. The answer provided in [22] requires the introduction of an *environment*, the quantum system interacts with unitarily as in Figure 85(c) and it is based on the theorem discussed next.

**Theorem.** *If we are given a quantum system $\mathcal{Q}$ with the density matrix $\rho$ in a $d$-dimensional Hilbert space $\mathcal{H}_d$ and a trace preserving operation $\mathcal{E}$ it is possible to construct an environment $E$ with the following properties:*

- *The environment is at most of dimensionality $d^2$;*
- *The environment is initially in a pure state with the density matrix $\epsilon = \mid \psi\rangle\langle\psi \mid$;*
- *The quantum system $\mathcal{Q}$ and the environment $E$ are initially uncorrelated.*

*Then there exists a unitary transformation $\mathbf{U}$ of the quantum system $\mathcal{Q}$ and the environment $E$ such that $\mathcal{E}(\rho) = tr_E\left[\mathbf{U}(\rho \otimes \epsilon)\mathbf{U}^\dagger\right]$.*

In other words, *a trace-preserving quantum operation can always be modelled as a unitary evolution by constructing an environment which interacts unitarily with the quantum system;* conversely, any unitary interaction between the quantum system and an initially uncorrelated environment gives risen to a trace-preserving quantum operation.

When the input to the communication channel, the quantum system $Q$, is in a mixed state, it is useful to introduce in addition to the environment $E$ a reference system $R$, such that $\mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle$, the joint state $(\mathcal{Q} \otimes \mathcal{R})$ is a pure state and that

$$\rho^{\mathcal{Q}}_{before} = \mathrm{tr}_R\left[\mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle\langle\varphi^{(\mathcal{Q},\mathcal{R})} \mid\right].$$

In Section 2.4 we discussed purification of mixed states, a concept very useful to study information transmission through noisy quantum channels. Recall that purification of a system $\mathcal{Q}$ in a mixed state means to identify another system $\mathcal{R}$ such that the composite system $\mathcal{P} = (\mathcal{Q}, \mathcal{R})$ is in a pure state. We now recognize that $\mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle$ is a purification of the state $\rho^{\mathcal{Q}}_{before}$.

The analysis of noisy quantum channels requires the characterization of the relationship between the states at the input and the output of the channel.

**Entropy exchange.** To model a noisy quantum channel Schumacher [371] and Lloyd [276] introduced the entropy exchange. If $\mathcal{E}$ is the quantum transformation of the input state $\rho$, the entropy exchange, $S_e$, is defined as the von Neumann entropy of the environment $E$ after the transformation $\mathcal{E}$, if the environment starts out in a pure state

$$S_e(\rho, \mathcal{E}) = S(\rho^{(\mathcal{Q}\mathcal{R})}_{after}) = S(\rho^{E}_{after}).$$

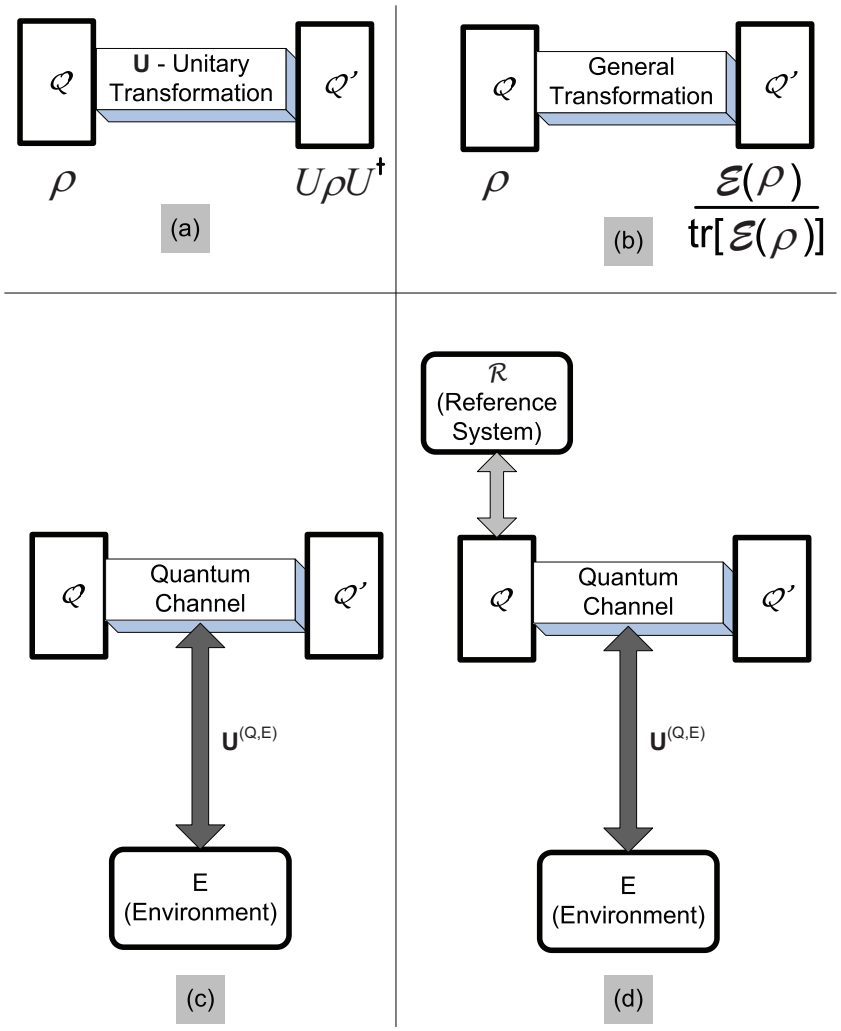



Figure 85: Quantum channel models. (a) The channel performs a unitary transformation $\mathbf{U}$ of quantum state $\mathcal{Q}$ described by the density matrix $\rho$ into state $\mathcal{Q}'$ described by $\mathcal{E}(\rho) = \mathbf{U}\rho\mathbf{U}^\dagger$; (b) The channel performs a general transformation with $\frac{\mathcal{E}(\rho)}{\mathrm{tr}[\mathcal{E}(\rho)]}$ the normalized output state. (c) The environment $E$ allows us to model a trace-preserving operation as a unitary operation based on the representation theorem for trace-preserving quantum operations. (d) A reference system $\mathcal{R}$ allows purification of a mixed state of the quantum system $\mathcal{Q}$.

Initially the environment $E$ is in a pure state as a result of a purification operation; $\rho^E_{after}$ is the density matrix of $E$ after the operation. When

$$\mathcal{E}(\rho) = \sum_i \mathbf{A}_i \rho \mathbf{A}_i^\dagger$$

then the entropy exchange can be expressed as

$$S_e(\rho, \mathcal{E}) = S(w) = -\mathrm{tr}\left[w \log w\right]$$

with $w = [w_{ij}]$ the density operator of the environment after transformation, a matrix with elements in an orthonormal basis given by

$$w_{ij} = \frac{\text{tr}\left[\mathbf{A}_i \rho \mathbf{A}_j^\dagger\right]}{\text{tr}\left[\mathcal{E}(\rho)\right]},$$

where $\{A_i\}$ are operation elements of $\mathcal{E}$. The entropy exchange is an intrinsic property of the system $Q$; it is a measure of the information exchanged between $Q$ and the environment during the transformation $\mathcal{E}$. *The entropy exchange limits the amount of information that could be acquired by an eavesdropper in a quantum cryptographic protocol.*

**Entanglement fidelity.** Schumacher [371] introduced the entanglement fidelity to measure how well a quantum system $\mathcal{Q}$ with the density matrix $\rho$ and its entanglement with other states are preserved by a noisy quantum channel. Recall that when $\mid \varphi\rangle$ is a pure state of a quantum system the fidelity $F$ of an arbitrary, possibly mixed state with density matrix $\rho$ is defined as $F = \langle\varphi \mid \rho \mid \varphi\rangle$. The fidelity $F$ is a measure of how close $\rho$ is to $\mid \varphi\rangle\langle\varphi \mid$.

The entanglement fidelity is a measure of the overlap between the initial purification and the joint state of the system $\mathcal{Q}$ and the reference system $\mathcal{R}$ after the transmission of $\mathcal{Q}$:

$$F_e(\rho, \mathcal{E}) = \langle\varphi^{(\mathcal{Q},\mathcal{R})} \mid \left[(\mathcal{I}_\mathcal{R} \otimes \mathcal{E})\left(\mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle\langle\varphi^{(\mathcal{Q},\mathcal{R})} \mid\right)\right] \mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle$$

with $\mid \varphi^{(\mathcal{Q},\mathcal{R})}\rangle$ the joint state of the quantum system $\mathcal{Q}$ and the reference system $\mathcal{R}$ before the transmission, $\mathcal{I}_\mathcal{R}$ the identity transformation of the reference system $\mathcal{R}$, and $\mathcal{E}$ a trace-preserving operation. When

$$\mathcal{E}(\rho) = \sum_i \mathbf{A}_i \rho \mathbf{A}_i^\dagger$$

then entanglement fidelity can be expressed as

$$F_e(\rho, \mathcal{E}) = \frac{\sum_i \mid \text{tr}\left[\mathbf{A}_i \rho\right] \mid^2}{\text{tr}\left[\mathcal{E}(\rho)\right]}.$$

The entanglement fidelity depends only on $\rho$ and $\mathcal{E}$ and not on the details of the purification.

Entanglement fidelity is a more stringent requirement than average pure state fidelity; as we shall see shortly the entanglement fidelity is a lower bound on the average fidelity of pure states. Thus, *if we are able to ensure a high degree of the entanglement fidelity for an ensemble of pure states, then individual pure states will be reproduced with high fidelity at the output of the quantum channel.* A critical point is that in quantum error correction we wish to apply an error correction procedure to a subset of a quantum register even though the subset may later become entangled with other parts of the quantum computer. The properties of the entanglement fidelity [22, 371]:

(i) The entanglement fidelity is a real number between 0 and 1, $0 \leq F_e(\rho, \mathcal{E}) \leq 1$. $F_e(\rho, \mathcal{E}) = 1$ if and only if for all pure states $\mid \varphi\rangle$ in the support of $\rho$ the following condition is satisfied: $\mathcal{E}(\mid \varphi\rangle\langle\varphi \mid) = \mid \varphi\rangle\langle\varphi \mid$.

(ii) The entanglement fidelity is a lower bound on the fidelity $F_e(\rho, \mathcal{E}) \leq F(\rho, \mathcal{E}(\rho))$.

(iii) Given an ensemble of quantum systems in states $\{\mid \varphi_i\rangle, p_i\}$ with the density matrix $\rho = \sum_i p_i \mid \varphi_i\rangle\langle\varphi i \mid$ the entanglement fidelity is a lower bound on the average fidelity of the pure states, $F_p$,

$$F_e(\rho, \mathcal{E}) \leq F_p = \sum_i p_i \langle\varphi_i \mid \mathcal{E}\left(\mid \varphi_i\rangle\langle\varphi i \mid\right) \mid \varphi_i\rangle.$$

(iv) If the pure state fidelity is $F_p = 1 - \eta$, for small $\eta \geq 0$, then the entanglement fidelity is [246]

$$F_e(\rho, \mathcal{E}) \geq 1 - \frac{3}{2}\eta.$$

(v) *Continuity.* The *continuity lemma* provides an upper bound of the entanglement fidelity when the channel input is perturbed. If $\mathcal{E}$ is a trace-preserving quantum operation, $\rho$ is the density matrix, and $\Delta$ is a zero trace Hermitian operator then,

$$\mid F_e\left(\rho + \Delta, \mathcal{E}\right)) - F_e(\rho, \mathcal{E}) \mid \quad \leq \quad 2\mathrm{tr}\left(\mid \Delta \mid\right) + \mathrm{tr}\left(\mid \Delta \mid\right)^2.$$

This lemma gives bounds on the change in entanglement fidelity when the input state is perturbed. A detailed proof of the lemma can be found in [22].

**Coherent information.** The coherent information is the quantum correspondent of the mutual information in classical Information Theory [372] and is defined as

$$I(\rho, \mathcal{E}) = S\left(\frac{\mathcal{E}(\rho)}{\mathrm{tr}\left[\mathcal{E}(\rho)\right]}\right) - S_e(\rho, \mathcal{E})$$

where the quantum operation $\mathcal{E}$ takes the input state $\rho$ to an output state $(\mathcal{E}(\rho)/\mathrm{tr}\left[\mathcal{E}(\rho)\right])$.

The coherent information $I(\rho, \mathcal{E})$ can be viewed as a measure of the degree to which the quantum coherence is preserved by the operation $\mathcal{E}$; in other words, a measure of the degree of quantum entanglement retained by the systems $\mathcal{Q}$ and $\mathcal{R}$ after this operation.

The role of the coherent information will be discussed in the next few sections. Now we just state several of the properties of coherent information [22]:

(i) Given the quantum operator $\mathcal{E}$ and the non-negative $\lambda$

$$I(\rho, \lambda\mathcal{E}) = I(\rho, \mathcal{E}).$$

(ii) *Convexity of coherent information*:

$$I\left(\rho, \mathcal{E}\right) \leq \sum_i p_i I\left(\rho, \mathcal{E}_i\right).$$

The coherent information $I\left(\rho, \mathcal{E}\right)$ is convex when the quantum operation $\mathcal{E} = \sum_i p_i \mathcal{E}_i$ is a trace-preserving convex sum of trace-preserving operations $\mathcal{E}_i$ with $p_i \geq 0$ and $\sum_i p_i = 1$. The proof of this property of coherent information follows immediately from the concavity of conditional von Neumann entropy discussed in Section 3.10.

(iii) *Generalized convexity theorem for coherent information*:

$$I\left(\rho, \sum_i \mathcal{E}_i\right) \leq \frac{\sum_i \operatorname{tr}\left[\mathcal{E}_i\left(\rho\right)\right] I\left(\rho, \mathcal{E}_i\right)}{\operatorname{tr}\left[\sum_i \mathcal{E}_i\left(\rho\right)\right]}$$

where $\mathcal{E}_i$ are quantum operations, can be proved based on convexity property of coherent information.

(iv) *Additivity of independent channels:*

$$I\left(\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n, \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_n\right) = \sum_i I\left(\rho_i, \mathcal{E}_i\right)$$

where $\rho_1, \rho_2, \ldots, \rho_n$ are density operators of $n$ quantum systems (qubits) and $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ are quantum operators which correspond to the environment acting independently on these systems. The evolution operator $\mathcal{E}$ acting on the joint state of the $n$ systems (qubits) $\rho = \rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$ factorizes as $\mathcal{E} = \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_n$.

The discussion in this section is very important for quantum error correction. As we shall see in Chapter 5 we separate the space of encoded states and the space of errors into orthogonal subspaces of the Hilbert space. To measure the quantum register in error we perform a unitary transformation involving also ancilla qubits. Such a trace-preserving completely positive map, or superoperator, allows us to determine a syndrome that uniquely identifies the error. In this process we do not disturb the encoded state and subsequently we can correct the error using the information stored in the ancilla qubits.

From this brief presentation it follows that we are also interested in how well subspaces of a Hilbert space can be transmitted through a quantum channel. Given a subspace with projector $\mathbf{P}$ then the *subspace fidelity* is defined as

$$F_s(\mathbf{P}, \mathcal{E}) = \min_{|\varphi\rangle} \langle \varphi \mid \mathcal{E}\left(\mid \varphi\rangle\langle\varphi \mid\right) \mid \varphi\rangle$$

where the minimization is over all states $\mid \varphi\rangle$ in the subspace whose projector is $\mathbf{P}$. The entanglement fidelity is close to 1 when the subspace fidelity is close to 1; this follows from the definition of subspace fidelity and the fourth property of the entanglement fidelity.

## 3.19 Quantum Fano and Data Processing Inequalities

In this section we introduce the quantum equivalent of two important properties of classical information discussed in Section 3.6. Recall that Fano inequality gives us a lower bound of the error when we estimate one random variable using another one correlated with it. The data processing inequality considers a source with entropy $H(X)$ and two cascaded channels. We assume that the source provides the input $X$ to a first communication channel and $Y$, the output of the first channel, is the input of a second channel. $X$, $Y$, and $Z$, the output of the second channel, form a Markov chain. Then the mutual information between the source and the output of the second channel, $I(X, Z)$ cannot be larger than the mutual information between the source and the output of the first channel, $I(X, Y)$, and, in turn $I(X, Y)$ cannot be larger than $H(X)$, the entropy of the source. An inequality with a similar interpretation was proven by Schumacher and Nielsen in [372] for quantum transformations.

We consider a quantum system characterized by the density matrix $\rho$ as input to a quantum channel characterized by the transformation $\mathcal{E}$; the output of the channel is $\mathcal{E}(\rho)$. Recall from Section 3.18 that $S_e(\rho, \mathcal{E})$, the entropy exchange, is defined as the von Neumann entropy of the environment after the transformation of the state due to the channel; the entanglement fidelity, $F_e(\rho, \mathcal{E})$ measures how well a quantum system with the density matrix $\rho$ and its entanglement with other states are preserved by a noisy quantum channel.

**The quantum Fano inequality.** This inequality relates the entropy exchange and the entanglement fidelity [373] as

$$S_e(\rho, \mathcal{E}) \leq h(F_e(\rho, \mathcal{E})) + [1 - F_e(\rho, \mathcal{E})] \log(d^2 - 1).$$

In this expression $d$ is the dimension of the Hilbert space of the quantum system $\mathcal{Q}$ and $h(p)$ is the Shannon entropy of a binary random variable, $h(p) = -p \log p - (1-p) \log(1-p)$.
*Proof:* The proof is based on the grouping property of Shannon entropy:

$$H(p_1, p_2, \ldots, p_{d^2}) = h(p_1) + (1 - p_1) H\left(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}, \ldots, \frac{p_{d^2}}{1-p_1}\right).$$

The proof [373] considers a set of $d^2$ orthonormal basis states $\{| \varphi_i \rangle\}$ of the joint system $(\mathcal{Q}, \mathcal{R})$ such that the first state in the set is $| \varphi_1 \rangle = | \varphi^{(\mathcal{Q},\mathcal{R})} \rangle$, the entangled state, and $p_i = \langle \varphi_i | \rho_{after}^{(\mathcal{Q},\mathcal{R})} | \varphi_i \rangle$. It can be shown that

$$S_e(\rho, \mathcal{E}) \leq H(p_1, p_2, \ldots, p_{d^2})$$

where $H(p_i)$ is the Shannon information of the set $\{p_i\}$. With the observation that by definition $p_1 = F_e(\rho, \mathcal{E})$ and that $H\left(\frac{p_2}{1-p_1}, \frac{p_3}{1-p_1}, \ldots, \frac{p_{d^2}}{1-p_1}\right) \leq \log(d^2 - 1)$, the inequality is evident.

The quantum Fano inequality shows that for a given quantum operation if the exchange entropy is large, the entanglement fidelity must necessarily be small, indicating that the entanglement between $\mathcal{Q}$ and $\mathcal{R}$ has not been well preserved.

**The quantum data processing inequality.** Consider a setup like the one in Figure 86 where given trace-preserving quantum operations $\mathcal{E}_1$ and $\mathcal{E}_2$ define the quantum process

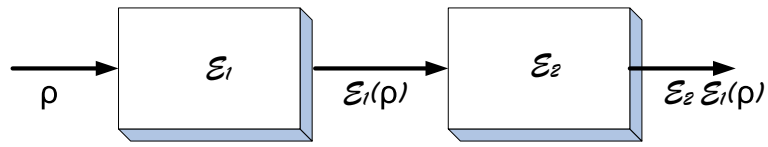
$$\rho \rightarrow \mathcal{E}_1(\rho) \rightarrow (\mathcal{E}_2 \circ \mathcal{E}_1)(\rho),$$

then $S(\rho)$, the von Neumann entropy of the source, cannot be exceeded by $I(\rho, \mathcal{E}_1)$, the coherent information between the source and the state after transformation $\mathcal{E}_1$ and, in turn, $I(\rho, \mathcal{E}_1)$ cannot be exceeded by $I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$, the coherent information between the state after transformation $\mathcal{E}_1$ and the state after transformation $\mathcal{E}_2 \circ \mathcal{E}_1$.

$$S(\rho) \geq I(\rho, \mathcal{E}_1) \geq I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1).$$

Schumacher and Nielsen also proved [372] that it is possible to reverse the transformation $\mathcal{E}$ if and only if no information is lost during the transformation, the entropy of the source is equal to the coherent information between the source and the output of the channel characterized by $\mathcal{E}$:

$$S(\rho) = I(\rho, \mathcal{E}).$$




$$S(\rho) \geq I(\rho, \ \mathcal{E}_1 \ ) \geq I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$$

Figure 86: The quantum data processing inequality. A quantum system with the density matrix $\rho$ is transmitted through two cascaded quantum channels characterized by the super-operators $\mathcal{E}_1$ and $\mathcal{E}_2$; the output of the first channel, $\mathcal{E}_1(\rho)$, becomes the input to a second channel. The density matrix of the system after the two trace-preserving transformations is $(\mathcal{E}_2 \circ \mathcal{E}_1)(\rho)$. The coherent information between the original input and the output of the second channel, $I(\rho, (\mathcal{E}_2 \circ \mathcal{E}_1)(\rho))$, cannot be larger than the coherent information between the original input and the output of the first channel, $I(\rho, \mathcal{E}_1(\rho))$ and, in turn, $I(\rho, \mathcal{E}_1(\rho))$ cannot be larger than $S(\rho)$, the von Neumann entropy of the source. The inequality $I(\rho, \mathcal{E}_1) \geq I\left(\rho, \mathcal{E}_2 \circ \mathcal{E}_1\right)$ holds even when $\mathcal{E}_1$ is trace-preserving but $\mathcal{E}_2$ is not.

*Proof:* To prove the first inequality $S(\rho) \geq I(\rho, \mathcal{E}_1)$ we consider in addition to the quantum system $\mathcal{Q}$, a reference system $\mathcal{R}$, and the environment $E$, see Figure 85(d). We start from the definition of coherent information and distinguish $\rho_{before} = \rho$, the state of the quantum system before, and $\rho_{after\ \mathcal{E}_1}$, the state after the transformation $\mathcal{E}_1$. The coherent information is

$$I(\rho, \mathcal{E}) = S\left(\mathcal{E}_1(\rho)\right) - S_e(\rho, \mathcal{E}_1)$$

with $S\left(\mathcal{E}_1(\rho)\right) = S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}\right)$ and $S_e\left(\rho, \mathcal{E}_1\right) = S\left(\rho^{(E)}_{after\ \mathcal{E}_1}\right)$. Now we observe that the state of the joint system $(\mathcal{Q}, \mathcal{R}, E)$ was a pure state before the transformation $\mathcal{E}_1$ and remains pure after. Indeed, the joint state $\mid \varphi^{(\mathcal{Q},\mathcal{R})}_{before}\rangle$ before the transformation is a purification of $\rho$, the possibly mixed state of quantum system $\mathcal{Q}$; the environment is initially in a pure state $\mid \varphi^{(E)}_{before}\rangle$. The representation theorem for trace-preserving transformations tells us that a trace-preserving operation can be regarded as a unitary transformation by adding an environment initially in a pure state that the system interacts unitarily with. It follows that $S(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}) = S(\rho^{(\mathcal{R},E)}_{after\ \mathcal{E}_1})$ and that similar relations exist among all permutations of $\mathcal{Q}$, $\mathcal{R}$, and $E$. Now by virtue of subadditivity of von Neumann entropy,

$$S\left(\rho^{(\mathcal{R},E)}_{after\ \mathcal{E}_1}\right) \leq S\left(\rho^{(\mathcal{R})}_{after\ \mathcal{E}_1}\right) + S\left(\rho^{(E)}_{after\ \mathcal{E}_1}\right).$$

Then,

$$I(\rho, \mathcal{E}) = S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}\right) - S\left(\rho^{(E)}_{after\ \mathcal{E}_1}\right) \leq S\left(\rho^{(\mathcal{R})}_{after\ \mathcal{E}_1}\right) + S\left(\rho^{(E)}_{after\ \mathcal{E}_1}\right) - S\left(\rho^{(E)}_{after\ \mathcal{E}_1}\right) = S\left(\rho^{(\mathcal{R})}_{after\ \mathcal{E}_1}\right).$$

Using similar arguments as before we see that $S\left(\rho^{(\mathcal{R})}_{after\ \mathcal{E}_1}\right) = S(\rho)$ thus we conclude that

$$I(\rho, \mathcal{E}) \leq S\left(\rho\right).$$

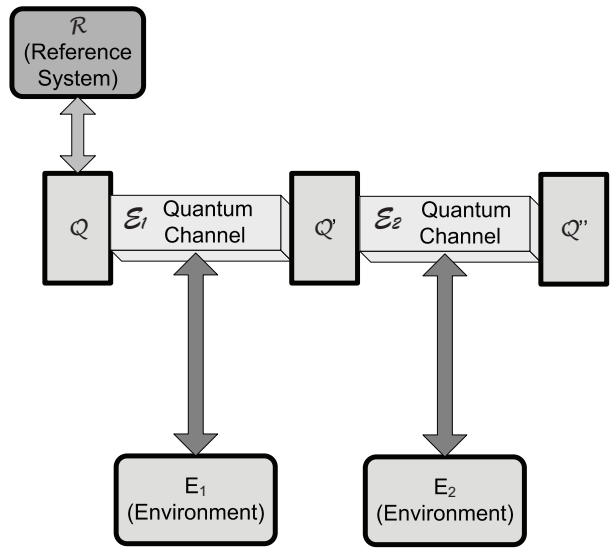



Figure 87: The configuration used to prove the second part of the quantum data processing inequality. The quantum systems $\mathcal{Q}$, $\mathcal{Q}'$, and $\mathcal{Q}''$ together with a reference system $\mathcal{R}$ and two environments $E_1$ and $E_2$ are shown; the state of quantum systems $\mathcal{Q}$, $\mathcal{Q}'$, and $\mathcal{Q}''$ is described by the density matrices $\rho, \rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}$ and $\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_2}$, respectively.

To prove the second inequality we consider the configuration in Figure 87 consisting of quantum systems $\mathcal{Q}$, $\mathcal{Q}'$, and $\mathcal{Q}''$ together with a reference system $\mathcal{R}$ and two environments $E_1$ and $E_2$. The state of quantum systems $\mathcal{Q}$, $\mathcal{Q}'$, and $\mathcal{Q}''$ is described by the density matrices $\rho, \rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}$ and $\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_2}$, respectively. The strong subadditivity of von Neumann entropy applied to this system requires that

$$S\left(\rho^{(\mathcal{R},E_1,E_2)}_{after\ \mathcal{E}_2}\right) + S\left(\rho^{(E_1)}_{after\ \mathcal{E}_2}\right) \leq S\left(\rho^{(\mathcal{R},E_1)}_{after\ \mathcal{E}_2}\right) + S\left(\rho^{(E_1,E_2)}_{after\ \mathcal{E}_2}\right).$$

We now express the individual terms of this expression. First, we observe that the joint state, $(\mathcal{Q}, \mathcal{R}, E_1, E_2)$, of the quantum system $\mathcal{Q}$, the reference system $\mathcal{R}$, and the two environments, $E_1$ and $E_2$, is pure at every stage in this chain of transformations, thus, it follows that

$$S\left(\rho^{(\mathcal{R},E_1,E_2)}_{after\ \mathcal{E}_2}\right) = S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_2}\right).$$

From the definition of entropy exchange it follows that

$$S\left(\rho^{(E_1)}_{after\ \mathcal{E}_2}\right) = S\left(\rho^{(E_1)}_{after\ \mathcal{E}_1}\right) = S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_1\right)$$

and

$$S\left(\rho^{(E_1,E_2)}_{after\ \mathcal{E}_2}\right) = S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_2 \circ \mathcal{E}_1\right).$$

Finally, we observe that neither the reference system $\mathcal{R}$, nor the environment $E_1$ are involved in the transformation $\mathcal{E}_2$ Thus,

$$\rho^{(\mathcal{R},E_1)}_{after\ \mathcal{E}_2} = \rho^{(\mathcal{R},E_1)}_{after\ \mathcal{E}_1}.$$

Then from the purity of the joint system $(\mathcal{Q}, \mathcal{R}, E_1)$ after the first stage it follows that

$$S\left(\rho^{(\mathcal{R},E_1)}_{after\ \mathcal{E}_2}\right) = S\left(\rho^{(\mathcal{R},E_1)}_{after\ \mathcal{E}_1}\right) = S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}\right).$$

After substituting in the strong subadditivity relation we conclude that

$$S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_2}\right) + S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_1\right) \leq S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}\right) + S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_2 \circ \mathcal{E}_1\right)$$

or,

$$S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_2}\right) - S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_2 \circ \mathcal{E}_1\right) \leq S\left(\rho^{(\mathcal{Q})}_{after\ \mathcal{E}_1}\right) - S_e\left(\rho^{(\mathcal{Q})}, \mathcal{E}_1\right).$$

From the definition of the coherent information it follows that

$$I\left(\rho, \mathcal{E}_2 \circ \mathcal{E}_1\right) \leq I(\rho, \mathcal{E}_1).$$

This concludes the proof of the quantum data processing inequality.

## 3.20 Reversible Extraction of Classical Information from Quantum Information

We have seen in Section 2.13 that quantum states can be teleported, transmitted without distortion by separating quantum from classical information and using two separate communication channels, a classical and a quantum one. The classical information is uncorrelated with the arbitrary quantum input state and this prevents an eavesdropper to observe non-orthogonal input states without disturbing them; the scheme allows input states to be transmitted with perfect fidelity.

Classical information has distinct advantages, it can be transmitted with an arbitrarily small probability of error and the recipient can recover the state of the source with almost certainty, while the recipient of quantum information has to perform an optimal measurement in order to maximize the mutual information subject to Holevo bound

$$I(X;Y) \leq S(\rho) - \sum_a p_a \rho_a$$

where $p_a$ are the probabilities and $\rho_a$ are the density matrices of the *signal states* $\mid a\rangle$ [325].

The example at the end of Section 3.14 illustrates the advantage of using a classical channel to transmit quantum information and, at the same time, shows that a neat separation of classical and quantum information is possible when the source states can be partitioned into mutually orthogonal subsets. Indeed, consider a three-dimensional Hilbert space with the orthonormal basis $\{\mid e_1\rangle, \mid e_2\rangle, \mid e_3\rangle\}$; the source $\mathcal{A}$ transmits the state $\mid e_1\rangle$ with probability $p_1 = 1/2$ and a random state in the subspace spanned by $\mid e_2\rangle$ and $\mid e_3\rangle$ with probability $p_2 = 1/2$ and has a von Neumann entropy of 1.5 bits [35]. The information produced by $\mathcal{A}$ can be neatly separated by the following encoding scheme: (i) determine first by an incomplete

measurement if the qubit is in state $\mid e_1\rangle$ and transmit one bit of classical information 1 if the state is $\mid e_1\rangle$; else, (ii) transmit both a bit 0 and the qubit carrying the state in the subspace spanned by $\mid e_2\rangle$ and $\mid e_3\rangle$. This encoding scheme allows the source state to be recovered with perfect fidelity.

The Schumacher compression and the quantum noiseless coding theorem tell us how to reduce the number of qubits required to encode the information produced by a quantum source viewed as an ensemble of pure, but not all orthogonal states, $\mid \varphi_a\rangle$ with probabilities $p_a$. The question we address is if it is advantageous to extract some classical information from the quantum information and transmit both of them as bits and qubits, respectively, through a pair of classical and quantum channels; could we improve on Schumacher compression by acting on the quantum part of the information in this scheme?

The separation of classical and quantum information and the compression of quantum information may, or may not, be always feasible and advantageous. For example, consider two sources, $\mathcal{B}$ emitting linearly polarized photons at $0°, 45°, 90°$ and $135°$ with equal probability and $\mathcal{C}$ emitting linearly polarized photons at $0°, 1°, 90°, 91°$ with equal probability. The former has a high entropy and cannot be compressed. A classical measurement of the later allows us to distinguish with high, though not perfect reliability between the two groups of states $0°, 1°$ and $90°, 91°$ and each one of these two groups has an entropy far less than one bit.

The following theorem answers the question whether we can improve on Schumacher compression by acting on the quantum part of the information after a reversible extraction of classical information from the quantum information [35].

**Improvement on Schumacher compression.** A quantum source consisting of the ensemble of states $\mid \varphi_a\rangle$ with probabilities $p_a$, $\mathcal{A} = \{\mid \varphi_a\rangle, p_a\}$ can be encoded into, and then recovered with perfect fidelity from, an intermediate representation consisting of orthonormal distinguishable states $\mid X_j\rangle$ holding classical information and intermediate quantum information non-orthogonal states $\mid \xi_{ja}\rangle$, such that the mean entropy of the intermediate quantum states is less than that of the source ensemble $\mathcal{A}$, if and only if the states of the source fall into two or more mutually orthogonal subsets of non-zero-probability.

We assume that the states of the source, $\mid \varphi_a\rangle \in \mathcal{H}_{source}$ are normalized, but not orthogonal, and that in the process of encoding we use an ancilla system initially in state $\mid \zeta\rangle \in \mathcal{H}_{ancilla}$. The composite system is then in a superposition of states $\mid X_j\rangle$ that will allow us to extract classical information and normalized, but non-orthogonal states, $\mid \xi_{ja}\rangle$, representing residual information not captured by the $\mid X_j\rangle$ states

$$\mid \Phi\rangle = \sum_j \alpha_{ja} \mid \xi_{ja}\rangle \otimes \mid X_j\rangle \quad \text{with} \quad \sum_j \mid \alpha_{ja} \mid^2 = 1 \quad \text{and} \quad \mid \Phi\rangle \in \mathcal{H}_{source} \otimes \mathcal{H}_{ancilla}.$$

The encoding proceeds as follows: we first measure the "classical" sub-system, in other words we determine if $\mid X_j\rangle$ is an eigenstate of the system and obtain a classical result $j$ with probability $P(j|a) = \mid \alpha_{ja} \mid^2$ when the source state was $\mid \varphi_a\rangle$. It is trivial to prove that the condition is sufficient, namely that when the states of the source fall into two or more mutually orthogonal subsets of non-zero-probability then we can recover with perfect fidelity the original states. Indeed, if the source states are separable into mutually orthogonal subsets, there is a non-demolition measurement which allows us for each original state $\mid \varphi_a\rangle$ to determine the subset of states it belongs to, without disturbing $\mid \varphi_a\rangle$. The proof that this condition

is necessary requires to show that the intermediate ensemble $\mathcal{A}' = \{|\ \xi_{ja}\rangle, p_a\}$ has the same entropy as the original one, $\mathcal{A} = \{|\ \varphi_a\rangle, p_a\}$. The proof is left as an exercise.

Thus, this theorem tells us that no compression of the quantum information is possible if the source ensemble of states is not separable into orthogonal subsets and if the decoding operation is perfectly faithful.

The question left unanswered by this theorem is what happens when instead of absolute fidelity we impose only asymptotic fidelity in decoding.

## 3.21 Noisy Quantum Channels

We can think about a quantum channel as the physical medium that enables the transfer of a quantum system from a sender to a receiver. The channel is noiseless when the quantum system is undisturbed during the transfer and it is noisy when the quantum system interacts with some other system, e.g., the environment, during the transfer.

Before we discuss noisy quantum channels we should take a closer look at the "noise," the undesirable transformations suffered by a quantum system during the transfer from the sender to the receiver. Consider two Hilbert spaces $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$; a quantum operation $\mathbf{N}$ is a map from the set of density operators in $\mathcal{H}_{in}$ to the set of density operators in $\mathcal{H}_{out}$ with three properties:

(i) If a system is characterized by the density matrix $\rho$, the probability that the process represented by $\mathbf{N}$ occurs is

$$0 \leq \mathrm{tr}\left[\mathbf{N}\left(\rho\right)\right] \leq 1, \qquad \forall \rho \in \mathcal{H}_{in}.$$

The normalized quantum state after the mapping $\mathbf{N}$ is

$$\frac{\mathbf{N}\left(\rho\right)}{\mathrm{tr}\left[\mathbf{N}\left(\rho\right)\right].}$$

While $\mathrm{tr}\left[\rho\right] = 1$, the quantum operation $\mathbf{N}$ may not preserve the trace property of the density matrix; in the general case $\mathrm{tr}\left[\mathbf{N}\left(\rho\right)\right] \neq 1$. To identify the outcome of the mapping $\mathbf{N}$ we perform a measurement and the probability of an outcome is $0 \leq \mathrm{tr}\left[\mathbf{N}\left(\rho\right)\right] \leq 1$. The result of the mapping is a valid density matrix up to a normalization factor.

(ii) $\mathbf{N}$ is a convex linear map

$$\mathbf{N}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathbf{N}\left(\rho_i\right).$$

The convexity of the linear map is a property relevant to ensembles of states $\{\rho_i, p_i\}$ with density matrices $\rho_i$ and probabilities $p_i$. There are two ways to look at the mapping $\mathbf{N}$:

- it is applied to $\rho$, the density matrix of the entire system,

$$\mathbf{N} : \rho \mapsto \frac{\mathbf{N}\left(\rho\right)}{\mathrm{tr}\left[\mathbf{N}\left(\rho\right)\right]} \quad \text{with} \quad \rho = \sum_i p_i \rho_i;$$

- it is applied to the individual states described by the density matrices $\rho_i$

$$\mathbf{N} : \rho_i \mapsto \frac{\mathbf{N}(\rho_i)}{\mathrm{tr}\left[\mathbf{N}(\rho_i)\right]}.$$

Call $p(\mathbf{N})$ the probability that the system in state $\mathbf{N}(\rho)$ was initially in state $\rho$ and $p(i|\mathbf{N})$ the probability that the system in state $\mathbf{N}(\rho_i)$ was initially in state $\rho_i$. According to condition (i) above these two probabilities are, respectively, $p(\mathbf{N}) = \mathrm{tr}\left[\mathbf{N}(\rho)\right]$ and $p(\mathbf{N}|i) = \mathrm{tr}\left[\mathbf{N}(\rho_i)\right]$. The condition that two results of applying $\mathbf{N}$ should be identical means that

$$\mathbf{N}(\rho) = \mathrm{tr}\left[\mathbf{N}(\rho)\right] \sum_i p(i|\mathbf{N}) \frac{\mathbf{N}(\rho_i)}{\mathrm{tr}\left[\mathbf{N}(\rho_i)\right]}.$$

But according to Bayes' rules,

$$p(i|\mathbf{N})\, p(\mathbf{N}) = p(\mathbf{N}|i)\, p_i.$$

This implies that

$$p(i|\mathbf{N}) = \frac{p(\mathbf{N}|i)\, p_i}{p(\mathbf{N})} = p_i \frac{\mathrm{tr}\left[\mathbf{N}(\rho_i)\right]}{\mathrm{tr}\left[\mathbf{N}(\rho)\right]}.$$

It follows immediately that

$$\mathbf{N}(\rho) = \mathbf{N}\left(\sum_i p_i \rho_i\right) = \sum_i p_i \mathbf{N}(\rho_i).$$

(iii) $\mathbf{N}$ is a completely positive map; if $\mathbf{A}$ is a positive operator then $\mathbf{N}(\mathbf{A})$ is positive. Moreover, if we have an additional space $\mathcal{H}_{aux}$ and $(\mathbf{I} \otimes \mathbf{N})(\mathbf{A})$ acts on the combined space $\mathcal{H}_{aux} \otimes \mathcal{H}_{in}$ then $(\mathbf{I} \otimes \mathbf{N})(\mathbf{A})$ is positive; in this expression $\mathbf{I}$ is the identity transformation on $\mathcal{H}_{aux}$.

The following proposition relates the operator-sum representation discussed in Section 2.14 with the properties defining quantum operations; the proof of this proposition can be found in [308].

**Proposition.** *A map* $\mathbf{N}$ *satisfies conditions (i) -(iii) if and only if*

$$\mathbf{N} = \sum_i \mathbf{E}_i \rho \mathbf{E}_i^\dagger$$

*with* $\{\mathbf{E}_i\}$ *a set of operators which map the input Hilbert space to the output Hilbert space and* $\sum_i \mathbf{E}_i \mathbf{E}_i^\dagger \leq \mathbf{I}$.

**Example.** *The partial trace is a quantum operation.* Consider two quantum systems $\mathcal{Q} \in \mathcal{H}_\mathcal{Q}$ and $\mathcal{R} \in \mathcal{H}_\mathcal{R}$ and let $\{|\, q_i\rangle\}$ be an orthonormal basis in $\mathcal{H}_\mathcal{Q}$ and $\rho$ a density operator of $\mathcal{Q}$; let $\{|\, j\rangle\}$ be an orthonormal basis for $\mathcal{H}_\mathcal{R}$ and $\rho^\mathcal{R} =|\, j\rangle\langle j'\, |$ a density matrix of $\mathcal{R}$. If $(\mathcal{Q}, \mathcal{R}) \in \mathcal{H}_{\mathcal{Q}\mathcal{R}}$ is a composite system of $\mathcal{Q}$ and $\mathcal{R}$ we define the linear operator $\mathbf{E}_i : \mathcal{H}_{\mathcal{Q}\mathcal{R}} \mapsto \mathcal{H}_\mathcal{Q}$ as

$$\mathbf{E}_i \left( \sum_j \lambda_i \mid q_j \rangle \otimes \mid j \rangle \right) \equiv \lambda_i \mid q_i \rangle$$

with $\lambda_i$ complex numbers. Now we define $\mathbf{N}$ as a quantum operation with operation elements $\mathbf{E}_i$

$$\mathbf{N}(\rho) = \sum_i \mathbf{E}_i \rho \mathbf{E}_i^\dagger.$$

Then,

$$\mathbf{N}\left(\rho \otimes \rho^{\mathcal{R}}\right) = \mathbf{N}\left(\rho \otimes \mid j \rangle \langle j' \mid\right) = \delta_{j,j'} \cdot \rho.$$

But,

$$\rho = \mathrm{tr}_{\mathcal{R}}\left[\rho \otimes \mid j \rangle \langle j' \mid \ \right].$$

We conclude that

$$\mathbf{N}\left(\rho \otimes \mid j \rangle \langle j' \mid\right) = \mathrm{tr}_{\mathcal{R}}\left[\rho \otimes \mid j \rangle \langle j' \mid \ \right].$$

Both $\mathbf{N}$ and the trace are linear operations thus we conclude that $\mathbf{N} = \mathrm{tr}_{\mathcal{R}}$.

There are several types of quantum operations carried out by noisy quantum channels: bit-flip, phase-flip, bit-phase flip, state depolarization, amplitude and then phase damping; in all instances some states are left invariant by the quantum operation but the Bloch sphere is deformed.

A <u>bit-flip channel</u> transforms the qubit state $\mid \varphi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$ into $\mid \psi \rangle = \alpha_1 \mid 0 \rangle + \alpha_0 \mid 1 \rangle$ with probability $p$. The basis states are transformed as $\mid 0 \rangle \mapsto \mid 1 \rangle$ and $\mid 1 \rangle \mapsto \mid 0 \rangle$; the operation elements are

$$E_0 = \sqrt{p}\ \sigma_I = \sqrt{p} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad E_1 = \sqrt{1-p}\ \sigma_x = \sqrt{1-p} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

After going through a bit-flip channel, the pure states represented by points on the $x$ axis of a Bloch sphere are not affected, while the ones in the $y - z$ plane are contracted by a factor $1 - 2p$. A <u>phase-flip channel</u> transforms the qubit state $\mid \varphi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$ into $\mid \psi \rangle = \alpha_0 \mid 0 \rangle - \alpha_1 \mid 1 \rangle$ with probability $p$; the operation elements are

$$E_0 = \sqrt{p}\ \sigma_I = \sqrt{p} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad E_1 = \sqrt{1-p}\ \sigma_z = \sqrt{1-p} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

After going through a phase-flip channel, the pure states represented by points on the $z$ axis of a Bloch sphere are not affected, while the ones in the $(x - y)$-plane are contracted by a factor $1 - 2p$. A <u>bit-phase flip channel</u> transforms the qubit state $\mid \varphi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$ into $\mid \psi \rangle = \alpha_1 \mid 0 \rangle - \alpha_0 \mid 1 \rangle$ with probability $p$; the operation elements are

$$E_0 = \sqrt{p}\ \sigma_I = \sqrt{p} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad E_1 = \sqrt{1-p}\ \sigma_y = \sqrt{1-p} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

After going through a bit-phase flip channel, the pure states represented by points on the $y$ axis of a Bloch sphere are not affected, while the ones in the $x-z$ plane are contracted by a factor $1-2p$.

Consider a quantum system $\mathcal{A} \in \mathcal{H}_n$ with density matrix $\rho^{\mathcal{A}}$. A depolarizing channel characterized by the quantum operation $\mathbf{N}$ replaces $\mathcal{A}$ with the completely mixed state $\frac{1}{n}I$ with probability $p$ and leaves the system's state unchanged with probability $(1-p)$

$$\mathbf{N}\left(\rho^{\mathcal{A}}\right) = \frac{p}{n}I + (1-p)\rho^{\mathcal{A}}.$$

To study the effect of the depolarizing channel on a single qubit with density matrix $\rho$ we set $n=2$ and note that according to Problem 6 at the end of Chapter 1

$$\frac{1}{2}\sigma_I = \frac{1}{4}\left(\rho + \sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z\right)$$

with $\sigma_I, \sigma_x, \sigma_y$ and $\sigma_z$ the Pauli matrices. Then,

$$\mathbf{N}\left(\rho\right) = (1-\frac{3}{4}p)\rho + \frac{1}{4}p\left(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z\right).$$

It follows that the operation elements are:

$$\sqrt{1-\frac{3}{4}p}\,\sigma_I, \quad \frac{\sqrt{p}}{2}\sigma_x, \quad \frac{\sqrt{p}}{2}\sigma_y, \quad \frac{\sqrt{p}}{2}\sigma_z.$$

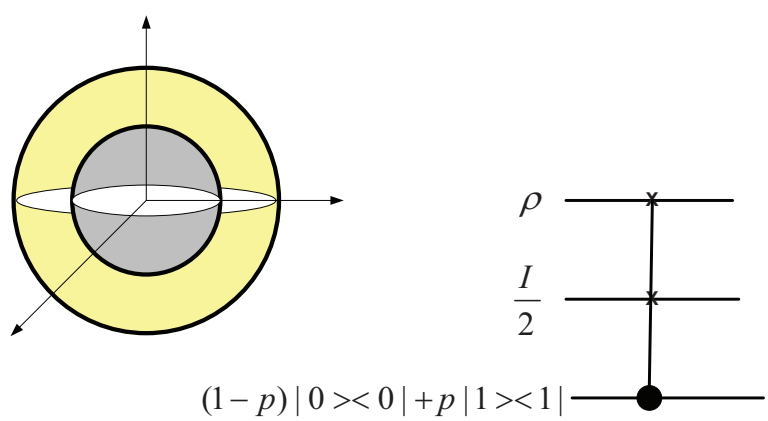



Figure 88: The effect of a depolarizing channel on one qubit. (Left) The Bloch sphere representation of a depolarizing channel with p=0.5. The sphere contracts uniformly depending on $p$. (Right) The quantum circuit simulating a depolarizing channel. The second and the third qubit represent the environment and simulate the channel. The third qubit is in a mixture of state $\mid 0\rangle$ with probability $1-p$ and of state $\mid 1\rangle$ with probability $p$ and acts as a control qubit, it swaps the original qubit and the second qubit in a completely mixed state $I/2$.

When a quantum system looses energy during its interaction with the environment we talk about amplitude damping. For example, consider the interaction of a photon in a superposition state $\mid \varphi\rangle = \alpha_0 \mid 0\rangle + \mid 1\rangle$ with a beam splitter that performs the unitary operation

$B = \exp\left[\theta\left(\alpha_0^\dagger\alpha_1 - \alpha_0\alpha_1^\dagger\right)\right]$ where $\alpha_0, \alpha_1$ and $\alpha_0^\dagger, \alpha_1^\dagger$ are the photon annihilation and creation operators, respectively, and $\theta$ is the angle of incidence of the photon. The output of the beam splitter is

$$\mid \psi\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \cos(\theta) \mid 01\rangle + \sin(\theta) \mid 10\rangle.$$

The quantum operation describing the transformation of the density matrix $\rho$ of the photon is

$$\mathbf{N}(\rho) = E_0\rho E_0 + E_1\rho E_1$$

with $E_i = \langle i \mid B \mid 0\rangle$, and $\gamma = \sin^2(\theta)$

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \quad \text{and} \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

The $E_1$ operation mimics the loss of energy to the environment, it changes $\mid 1\rangle$ into $\mid 0\rangle$; the $E_0$ operation reduces the amplitude of the $\mid 1\rangle$ state and leaves the $\mid 0\rangle$ state unchanged. In fact, energy is not lost to the environment, but the environment perceives that it is more likely that the system is in state $\mid 0\rangle$ than in state $\mid 1\rangle$. In this case the state $\mid 0\rangle$ is left invariant.

When the quantum operation describes the energy dissipation to the environment at finite temperature the process is called generalized amplitude damping; then the quantum operation describing the transformation of the density matrix due is

$$\mathbf{N}(\rho) = E_0\rho E_0 + E_1\rho E_1 + E_2\rho E_2 + E_3\rho E_3$$

with the operation elements for a single qubit

$$E_0 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad E_1 = \sqrt{p}\begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}$$

and

$$E_2 = \sqrt{1-p}\begin{pmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{pmatrix}, \quad E_3 = \sqrt{1-p}\begin{pmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{pmatrix}$$

The density matrix of the resulting stationary state $\rho_{sty} = \mathbf{N}(\rho)$

$$\rho_{sty} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}$$

satisfies the relation $\mathbf{N}(\rho_{sty}) = \rho_{sty}$. The generalized amplitude damping describes, for example, the $T_1$ relaxation process due to the coupling of spins to a lattice system which in thermal equilibrium at a temperature much higher than the spin temperature.

The phenomenon of phase damping implies a loss of information without the loss of energy, a pure quantum mechanical process unique to quantum information. When a system evolves in time the energy eigenstates of the system do not change, but they accumulate a phase proportional to the eigenvalue of the corresponding eigenstate. This happens for example

when photons are scattered randomly through a waveguide; it also happens when the states of an atom are perturbed due to interaction with distant charges. When a system evolves for an unknown period of time, information about the relative phase between energy eigenstates is lost. Assume that we have a qubit in state $\mid \varphi\rangle = a_0 \mid 0\rangle + a_1 \mid 1\rangle$ and we apply to it a rotation operation $R_z(\theta)$ with a random angle of rotation, a so-called *phase kick* operation. The density matrix of the output state is obtained by averaging over the angle $\theta$

$$\rho = \begin{bmatrix} \mid a \mid^2 & ab^* e^{-\lambda} \\ a^* b e^{-\lambda} & \mid b \mid^2 \end{bmatrix}$$

with $\lambda$ proportional to the time of rotation. A characteristic of phase damping operation is the fact that the off-diagonal terms of the density matrix decay exponentially with time. The phase damping amount increases with time for all states except the states along the $z$-axis which remain invariant.

The operation elements of the phase damping operation for two interacting oscillators, one representing the quantum system, the other the environment and with an interaction Hamiltonian $H = \chi a^\dagger a(b + b^\dagger)$ are

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} \qquad E_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{bmatrix}$$

with $\lambda = 1 - \cos^2(\chi \delta t)$ for an interaction time interval $\delta t$. These operation elements can be transformed through a unitary recombination into a new set of operation elements very similar to those of the phase flip channel

$$E_0' = \sqrt{k} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad E_1' = \sqrt{1-k} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

with $k = \frac{1}{2}(1 + \sqrt{1-\lambda})$. This shows that the phase damping quantum operation is exactly the same as the phase flip channel, thus phase errors that once were thought of as continuous can be treated discretely; that makes the quantum error-correction of phase errors possible. The phase damping is often referred to as the *spin-spin relaxation* process.

## 3.22 Holevo-Schumacher-Westmoreland Noisy Quantum Channel Encoding Theorem

When we discussed classical noisy channels we addressed two related questions: (i) Is it possible to encode the input so that the probability of error is arbitrarily low? and (ii) If such an encoding scheme exists, what is the maximum rate at which classical information can be transmitted over the channel using this encoding scheme? Similar questions, but framed differently, are discussed in this section for noisy quantum channels.

As before, we consider discrete memoryless quantum channels; the state of the quantum system is a vector in a finite-dimensional Hilbert space and the memoryless property requires that the mean of the output of the quantum channel is only determined by the current input. This time the quantum channel is noisy, it applies a transformation $\mathbf{N}$ to its input; if $\rho^{\mathcal{A}}$ is the density matrix of the input, then the density matrix of the output is $\rho^{out} = \mathbf{N}(\rho^{\mathcal{A}})$ and

the states at the input and the output of the channel are in the general case mixed states. The basic reasoning follows closely the one in Section 3.17; we consider product states and typical subspaces and we are concerned with the probability of error associated with specific encoding and decoding schemes.

To give a mathematical description of the quantum channel means to succinctly describe the properties of the transformation $\mathbf{N}$. The channel maps density matrices to density matrices thus, the mapping must be *linear, trace-preserving*, and *positive*. Linearity is required by basic principles of quantum mechanics; the channel must preserve the trace of the density matrices and map density matrices that are positive semi-definite matrices to positive semi-definite matrices. The mapping $\mathbf{N}$ must be *completely positive*, it must be positive even when tensored with the identity matrix [388].

We assume that the input to the noisy quantum channel is an ensemble of mixed states (we shall refer to them as letter states, or letters) $\mid q_i\rangle$ with density matrix $\rho_i$ and von Neumann entropy $S(\rho_i)$; the letters are selected with a priori probability $p_i$. The average density matrix and the von Neumann entropy of the ensemble are, respectively,

$$\omega = \sum_i p_i \rho_i \quad \text{and} \quad S(\omega) = -\text{tr}\left[\omega \log \omega\right].$$

We now consider strings of $n$ such letters:

$$\mid a\rangle = \mid q_{a,1}\rangle \otimes \mid q_{a,2}\rangle \otimes \ldots \otimes \mid q_{a,n}\rangle.$$

The string $\mid a\rangle$ occurs with a priori probability:

$$P_a = p_{a,1} p_{a,2} \ldots p_{a,n}$$

and has the density matrix:

$$\rho_a = \rho_{a,1} \otimes \rho_{a,2} \otimes \ldots \otimes \rho_{a,n}.$$

The density matrix $\rho_a$ has a set of $1 \leq k \leq n$ orthogonal eigenvectors $\mid e_{a,k}\rangle$ with the corresponding eigenvalues $\gamma_{k|a}$. Since $\text{tr}(\rho) = 1$ these eigenvalues $\gamma_{k|a}$ form a probability distribution. The index $k$ identifying the $n$-dimensional eigenvector $\mid e_{a,k}\rangle$ plays a significant role in quantum error correction and is called *the syndrome of the codeword* $\mid a\rangle$ *identified by the density matrix* $\rho_a$. The probability of the eigenvector $\mid e_{a,k}\rangle$ is

$$P_{a,k} = P_a \ \gamma_{k|a}.$$

This expression suggests that we can look at $P_{a,k}$ as a joint probability and at $\gamma_{k|a}$ as the conditional probability that the eigenvector $\mid e_{a,k}\rangle$ is observed when the string $\mid a\rangle$ was sent. Using these notations we can express $\rho$, the density matrix of the average state of the channel input as

$$\rho = \sum_a P_a \rho_a = \omega^{\otimes n} = \sum_a \sum_k P_{a,k} \mid e_{a,k}\rangle\langle e_{a,k} \mid$$

with $\omega^{\otimes n}$ the tensor product $\omega^{\otimes n} = \omega \otimes \omega \otimes \ldots \otimes \omega$.

The main result discussed in this section is known as the Holevo-Schumacher-Westmoreland (HSW) noisy quantum channel encoding theorem proven independently by Holevo [212] and by Schumacher and Westmoreland [374].

**Theorem.** *Consider a noisy quantum channel characterized by the trace-preserving transformation* $\mathbf{N}$*; the input of the channel is an ensemble of mixed letter states with the a priori probability* $p_i$*, the density matrix* $\rho_i$*, and the von Neumann entropy* $S(\rho_i)$*. The average density matrix of the ensemble is* $\omega$*, with* $\omega = \sum_i p_i \rho_i$*, and the corresponding von Neumann entropy is* $S(\omega)$*. Given arbitrarily small positive constants* $\epsilon, \delta \geq 0$*, for sufficiently large* $n$ *there exist:*

*(i) a code* $\mathcal{C}$ *whose codewords are strings of* $n$ *letters and*

*(ii) a decoding observable such that: the information per letter is at least* $\zeta - \delta$ *with* $\zeta = S(\omega) - \sum_j p_j S(\rho_j)$ *and the probability of error is* $p_{err} \leq \epsilon$*.*

*The capacity of the noisy quantum channel to transmit classical information is*

$$C_1 = \max_{\{p_i, \rho_i\}} \left[ S(\omega) - \sum_i p_i S(\rho_i) \right],$$

*where the maximum is over all ensembles of possible input mixed states* $\{p_i, \rho_i\}$*.*

We use the notation $C_1$ rather than $C$ to emphasize the fact that this is the *one shot* classical capacity, the maximum amount of classical information sent through a single use of the channel by an optimal choice of the source states and an optimal measurement. Recall from Section 3.18 that the syndrome measurement tells us as much as possible about the error affecting a codeword, but nothing at all about the information carried by the codeword; otherwise the measurement would destroy any quantum superposition of the logical qubits with other qubits in the system.

Instead of considering the ensemble of mixed states $\mid q_i\rangle$ with density matrices $\rho_i$ as the input of the quantum noisy channel we can think of them as the output of the channel with $\mid q_i^{in}\rangle$ as input [374]

$$\mid q_i\rangle = \mathbf{N}\left(\mid q_i^{in}\rangle\right), \quad \rho_i = \mathbf{N}(\rho_i^{in}), \quad \text{and} \quad \omega = \mathbf{N}\left(\omega^{in}\right).$$

Then, the effect of the state transformation due to the noisy channel appears explicitly [308] as

$$\zeta(\mathbf{N}) = S\left(\mathbf{N}\left(\omega^{in}\right)\right) - \sum_i p_i S\left(\mathbf{N}\left(\rho_i^{in}\right)\right)$$

and

$$C_1 = \max_{\{p_i, \rho_i\}} \left[ S\left(\mathbf{N}\left(\omega^{in}\right)\right) - \sum_i p_i S\left(\mathbf{N}\left(\rho_i^{in}\right)\right) \right].$$

In $\zeta(\mathbf{N})$ we recognize the expression of the Holevo bound for a quantum system subjected to a trace preserving transformation $\mathbf{N}$. The proof of this theorem, like the proof of the coding theorem for noiseless channels in Section 3.17, involves several aspects: (i) the typicality of codewords consisting of $n$ signals from the input alphabet; (ii) the question of encoding and decoding strategies; and (iii) the probability of error.

Given $\epsilon, \delta > 0$ we have to find an integer $n$ large enough such that the typicality condition on strings of letters of length $n$ is satisfied. Call $\mid e_i\rangle$ and $\lambda_i$ the eigenvectors and, respectively, the eigenvalues of $\rho$. First, we have to show that there is a typical subspace $\Lambda$ spanned by the eigenvectors of $\rho$ such that,

$$\mathrm{tr}\left[\rho\Pi\right] > 1 - \epsilon$$

where $\Pi$ is the projection onto $\Lambda$ of the original states. In addition, the typical subspace $\Lambda$ must satisfy the property

$$\mathrm{tr}\left[\rho^2\Pi\right] < 2^{-n(S(\omega)-2\delta)}.$$

To bound the error we require that the eigenvalues of $\rho$ corresponding to eigenvectors in the typical subspace be restricted to a very narrow range of values:

$$2^{-n(S(\omega)+\delta)} \leq \lambda_i \leq 2^{-n(S(\omega)-\delta)}, \quad \forall \mid e_i\rangle \in \Lambda.$$

To provide a precise definition of typicality in this context we introduce several new concepts. Recall that we have three probability distributions: (i) for individual letters - letter $\mid q_i\rangle$ occurs with probability $p_i$; (ii) for strings of $n$ letters - string $\mid a\rangle$ occurs with probability $P_a$ ; and (iii) the joint probability for string-syndrome pairs - the pair $(\mid a\rangle, k)$ occurs with probability $P_{a,k}$. Then the Shannon entropy of the individual letters of the input alphabet, of the strings of letters, and of the string-syndrome pairs are defined as

$$H(X) = \sum_i p_i \log p_i, \quad H(A) = -\sum_a P_a \log P_a, \quad \text{and} \quad H(A,K) = -\sum_a \sum_k P_{a,k} \log P_{a,k}.$$

These quantities are related to one another:

$$H(A) = nH(X)$$

and

$$H(A,K) = H(A) + \sum_a P_a S(\rho_a) = n\left(H(X) + \sum_i p_i S(\rho_i)\right).$$

Now we can provide a more precise definition of typicality. First, we say that the string $\mid a\rangle$ is a *typical string* if its probability is in a narrow range

$$2^{-n(H(X)-\delta)} < P_a < 2^{-n(H(X)+\delta)}, \quad \forall \mid a\rangle \in \Lambda.$$

In addition,

$$\sum_{\mid a\rangle \in \Lambda} P_a > 1 - \epsilon, \quad \forall \mid a\rangle \in \Lambda.$$

Second, we say that $(\mid a\rangle, k)$ is a *typical string-syndrome pair* if the joint probability $P_{a,k}$ is in a narrow range as well

$$2^{\left[-n\left(H(X)+\sum_i p_i S(\rho_i)+\delta\right)\right]} < P_{a,k} < 2^{\left[-n\left(H(X)+\sum_i p_i S(\rho_i)-\delta\right)\right]}.$$

We also require that

$$\sum_{|a\rangle,k} P_{a,k} > 1-\epsilon, \qquad \forall \text{ string-syndrome pairs } (\mid a\rangle, k).$$

Third, if $\mid a\rangle$ is a typical string then the set of $k$, such that $(\mid a\rangle, k)$ are typical string-syndrome pairs, is called the set of *relatively typical syndromes* for $\mid a\rangle$. The corresponding probability, $p_{k|a} = P_{a,k}/P_a$, should also be in a narrow range

$$2^{\left[-n\left(\sum_i p_i S(\rho_i)+2\delta\right)\right]} < p_{k|a} < 2^{\left[-n\left(\sum_i p_i S(\rho_i)-2\delta\right)\right]}.$$

The last inequality can be expressed as

$$2^{[-n(S(\omega)-\zeta+2\delta)]} < p_{k|a} < 2^{[-n(S(\omega)-\zeta-2\delta)]}.$$

A code $\mathcal{Q}$ is a subset of the set the strings of length $n$ of letters; codewords are labelled using the Greek alphabet, e.g., $\mid \alpha\rangle, \mid \beta\rangle, \ldots$, while we use the Roman alphabet, e.g., $\mid a\rangle, \mid b\rangle, \ldots$, to denote strings of $n$ letters that are not codewords. The code $\mathcal{Q}$ consists of $N$ codewords used with the same frequency.

The decoding procedure will be based on the "pretty good measurement" discussed in Section 3.17 but now we shall attempt to identify the codewords, as well as the syndromes. The decoding observable is a POVM described by a set of positive operators summing to unity; for each codeword-syndrome pair, $(\mid \alpha\rangle, k)$, we have a vector $\mid \mu_{\alpha,k}\rangle$ such that the projector $\mid \mu_{\alpha,k}\rangle\langle\mu_{\alpha,k} \mid$ is an element of the decoding POVM.

The probability of errors can be derived using a logic similar to the one for noiseless quantum channels:

$$p_{err} = 1 - \frac{1}{N}\sum_{a,k} p_{k|a} \mid \langle\mu_{\alpha,k} \mid e_{\alpha,k}\rangle \mid^2$$

with $\mid e_{\alpha,k}\rangle$ eigenvectors of the density matrix $\rho$ corresponding to the codewords. To provide an upper bound for $p_{err}$ Schumacher and Westmoreland use arguments mirroring closely the ones in Section 3.17. The derivation is fairly involved and the interested reader may consult [374] for the details; the final result for an upper bound for the probability of error averaged over all random codes is

$$p_{err} \leq 9\epsilon + N2^{-n(\zeta-4\delta)}.$$

For sufficiently large $n$ one can choose $N$ nearly as big as $2^{n\zeta}$ codewords and still have a small probability of error. The theorem shows that good codes for noisy quantum channels exist without addressing the question how to construct such codes.

## 3.23 Capacity of Noisy Quantum Channels

In Section 3.15 we define the capacity $Q$ of a quantum channel $\mathcal{N}$ characterized by the superoperator $\mathbf{N}$ as the largest number $Q$ such that, for any rate $R < Q$ and for any small and positive $\epsilon > 0$, the original state $\mid \varphi\rangle$ can be recovered by the receiver with a fidelity $F \geq 1 - \epsilon$. While this definition resembles closely the one for classical channels, determining the capacity of quantum channels proves to be more challenging than expected due to the inherent properties of quantum information that preclude the direct mapping of classical methods to studying the quantum channel capacity. Moreover, the presence of ancillary resources, such as a classical channel and entanglement, affect in subtle ways the capacity of a quantum channel.

Quantum information cannot be cloned and measurements generally alter the quantum state thus, encoding and decoding are conceptually more complex and will affect the channel capacity. Non-orthogonal pure and mixed states cannot be reliably distinguished thus, it is not possible to attain perfect fidelity, a reason to expect lower quantum channel capacity when the input states are non-orthogonal. The means to transport classical information using entangled states should be investigated "de nuovo", as entangled states do not have a classical corespondent.

As mentioned in Section 3.15, there are several definitions of the capacity of a quantum channel based on the nature of the information transmitted (classical or quantum), the ancillary resources allowed (classical channel, entanglement), and the protocols permitted [387]. The four distinct capacities as defined in [47] are: (i) $C$, the ordinary capacity for transmitting classical information, defined as the largest asymptotic rate of transmitting reliably classical information using an encoder and a decoder characterized by the superoperators $\mathbf{E}$ and $\mathbf{D}$, respectively; (ii) $Q$, the quantum capacity for transmitting quantum states, defined as the largest asymptotic rate of transmitting reliably qubits under similar conditions as in the definition of $C$; (iii) $Q_2$, the classically-assisted quantum capacity for transmitting qubits with the assistance of a two-way classical channel; and (iv) $C_E$, the entanglement-assisted classical capacity.

In Section 3.22 we have seen that according to the HSW theorem the capacity of the noisy quantum channel for transmitting classical information is

$$C_1 = \max_{\{p_i, \rho_i\}} \left[ S(\omega) - \sum_i p_i S(\rho_i) \right]$$

where the input of the channel is an ensemble of mixed letter states with probability $p_i$, the density matrix $\rho_i$, and the von Neumann entropy $S(\rho_i)$; the average density matrix of the ensemble is $\omega$, where $\omega = \sum_i p_i \rho_i$, and the corresponding von Neumann entropy is $S(\omega)$. $C_1$ is maximum amount of classical information sent through a single use of the channel by an optimal choice of the source states and an optimal measurement, the *one shot* classical capacity [374].

A common framework for various capacities of a quantum channel $\mathcal{N}$, with an encoding subprotocol, $\mathbf{E}$ and a decoding subprotocol $\mathbf{D}$ is introduced [47]. The encoder $E$ receives an $n$-qubit state, $\mid \varphi\rangle \in \mathcal{H}_n$ and produces $m \geq n$ possibly entangled states as inputs to the channel $N$; the decoder $D$ receives the $m$ possibly entangled qubits and produces $n \leq m$

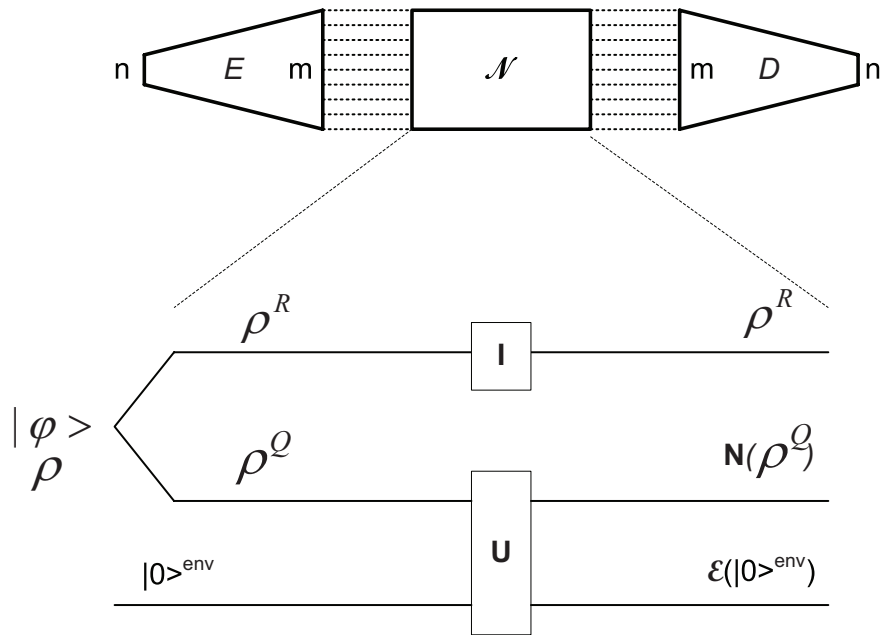



Figure 89: A quantum channel $\mathcal{N}$, with an encoding subprotocol, $\mathbf{E}$ and a decoding subprotocol $\mathbf{D}$; the encoder $E$ receives an $n$-qubit state, $\mid \varphi\rangle \in \mathcal{H}_n$ and produces $m \geq n$ possibly entangled states as inputs to the channel $\mathcal{N}$; the decoder $D$ receives the $m$ possibly entangled qubits and produces $n \leq m$ qubits as output. The noisy channel $\mathcal{N}$ could be viewed as a unitary interaction $\mathbf{U}$ of the quantum system $Q$ with an environment initially in state $\mid 0\rangle$ which remains in a mixed state $\mathcal{E}(\rho^{env})$ after the interaction. The input to $\mathcal{N}$ is a mixed state $\mid \varphi\rangle$ with the density matrix $\rho$ which can be imagined as part of an entangled pure state of a joint system between $Q$ and a purifying reference system $R$. The system $R$ is not physically acted upon, thus the operation on it is $I$. The output joint state is $\mathbf{N} \otimes \mathbf{I}\left(\rho^Q \otimes \rho^R\right)$ and has the same entropy as $\mathcal{E}(\rho^{env})$.

qubits as output, Figure 89. Then the generalized capacity $C_X(\mathbf{N})$ of a quantum channel is defined asymptotically as

$$C_X(\mathcal{N}) = \lim_{\epsilon \mapsto 0} \lim \sup_{m \mapsto \infty} \left\{ \frac{n}{m} : \exists_{\mathbf{E}} \exists_{\mathbf{D}} \ \ \forall_{|\varphi\rangle \in \mathcal{H}_n} \quad F(\mid \varphi\rangle, \mathbf{E}, \mathbf{D}, \mathbf{N}) > 1 - \epsilon \right\}.$$

with $F(\mid \varphi\rangle, \mathbf{E}, \mathbf{D}, \mathbf{N},)$ the fidelity of the output state relative to the input state $\mid \varphi\rangle$. The input and output von Neumann entropy of the noisy quantum channel model in Figure 89 are, respectively,

$$S\left(\rho^Q\right) = -\mathrm{tr}\left[\rho^Q \log \rho^Q\right] \qquad \text{and} \qquad S\left(\mathbf{N}(\rho^Q)\right).$$

The definition of the generalized capacity of a quantum channel is strikingly similar with the one for a classical binary channel $N$ with the encoding and decoding subprotocols $\mathbf{E}$ and $\mathbf{D}$, respectively,

$$C(\mathcal{N}) = \lim_{\epsilon \mapsto 0} \lim \sup_{m \mapsto \infty} \left\{ \frac{n}{m} : \exists_{\mathbf{E}} \exists_{\mathbf{D}} \ \ \forall x \in \{0,1\} \quad F(x, \mathbf{E}, \mathbf{D}, \mathbf{N}) > 1 - \epsilon \right\}.$$

For the classical capacity of a classical binary channel the encoding, $\mathbf{E}$, and decoding, $\mathbf{D}$, protocols and the channel $\mathbf{N}$ are restricted to be classical stochastic maps. For the quantum capacity the encoding and decoding protocols are completely positive trace preserving maps from $\mathcal{H}_2^{\otimes n}$ to the input space of $N^{\otimes m}$ and from the output space of $N^{\otimes m}$ back to the space $\mathcal{H}_2^{\otimes n}$ of the output.

The methodology to study the capacity of a given channel $M$ is sometimes based on the use of the asymptotic efficiency of a channel $N$, with known capacity, to simulate channel $M$; this method requires a comprehensive understanding of the properties of the two channels. The asymptotic capacity of channel $\mathcal{N}$ with the superoperator $\mathbf{N}$ to simulate channel $\mathcal{M}$ with the superoperator $\mathbf{M}$ is defined as

$$C_X(\mathcal{N},\mathcal{M}) = \lim_{\epsilon\mapsto 0}\lim\sup_{m\mapsto\infty}\left\{\frac{n}{m} : \exists_{\mathbf{E}}\exists_{\mathbf{D}} \ \ \forall|\ \varphi\rangle \in \mathcal{H}_{\mathbf{M}}^{\otimes n} \quad F\left(\mathbf{M}^{\otimes n}(|\ \varphi\rangle), \mathbf{E}, \mathbf{D}, \mathbf{N}\right) > 1-\epsilon\right\},$$

where the encoder $\mathbf{E}$ and decoder $\mathbf{D}$ enable the acceptance of an input $|\ \varphi\rangle$ in a space $\mathcal{H}_{\mathbf{M}}^{\otimes n}$ which is the tensor product of $n$ copies of the input Hilbert space of the channel $\mathbf{M}$ to be simulated and make $m$ forward uses of the simulating channel $\mathbf{N}$ to produce an output state the fidelity of which is compared with that of the state which would have been generated by sending the input $|\ \varphi\rangle$ through $\mathbf{M}^{\otimes n}$.

In Section 3.24 we shall discuss the classical channel simulation of a quantum channel. Such simulation involves some forward classical communication; the *forward classical communication costs (FCCC)* measures the capacity of the classical channel needed to simulate the quantum channel.

In some cases exact expressions for the capacities of different types of quantum channels are known, while in many other cases only upper and lower bounds have been derived. Typically, $Q < C$ and $Q \leq C \leq C_E$. In [42] a distinction is made between $Q_1$, the capacity of a quantum channel assisted by a forward classical channel and $Q_2$, the capacity of a quantum channel assisted by a two-way classical channel. It has been shown [38] that a forward classical channel does not increase the capacity of the quantum channel, thus $Q_1(N) = Q(N)$. By definition $Q \leq Q_2$ and it is conjectured that $Q_2 \leq C$.

**Example.** *The quantum depolarizing channel.* QDC is the quantum analog of the binary symmetric channel; QDC replaces with probability $p$ the incoming qubit state with a random qubit state without informing the receiver of the identity of the qubit in error. The bounds for the capacity of a depolarizing channel [38, 39, 42, 138] are summarized in Figure 90.

**Example.** *The* 2/3*-depolarizing quantum channel.* The channel transmits a qubit state intact with probability 1/3 and replaces it with a random qubit state with probability $p = 2/3$. This quantum channel can be simulated by a classical channel using the following procedure: a third party selects a random orientation $R$ and communicates it to the sender and the receiver; the sender measures the qubit along the direction $R$ and sends through the classical channel the result of the measurement, either a 0 or a 1 thus, the FCCC of this simulation is 1 bit. The classical capacity of the 2/3-depolarizing quantum channel can be measured when the binary 0 is encoded as the quantum state $|\ 0\rangle$ and 1 is encoded as $|\ 1\rangle$ in some basis and the measurement is carried out by the receiver in the same basis. The 2/3-depolarizing

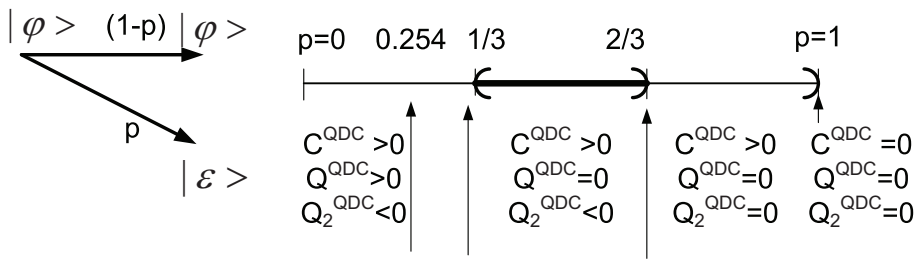



Figure 90: QDC, quantum depolarizing channel - the state $\mid \varphi\rangle$ of the incoming qubit it replaced with probability $p$ with a random state $\mid \epsilon\rangle$. $C^{QDC}$, $Q^{QDC}$, and $Q_2^{QDC}$ are the capacities of the quantum depolarizing channel for transmitting: classical information, quantum states, and quantum states with the assistance of a two-way classical channel, respectively. Quantum capacity depends on the probability $p$ of depolarization error: (i) when $p < 0.25408$, then $(C^{QDC}, Q^{QDC}, Q_2^{QDC}) > 0$; (ii) when $1/3 < p < 2/3$ then $(C^{QDC}, Q_2^{QDC}) > 0$ but $Q^{QDC} = 0$; (iii) when $2/3 \leq p < 1$ then $C^{QDC} > 0$ but $Q^{QDC} = Q_2^{QDC} = 0$; (iv) when $p = 1$ then $C^{QDC} = Q^{QDC} = C_2^{QDC} = 0$;

quantum channel can be simulated by a classical binary symmetric channel (BSC) with the error probability $p_{BSC} = 1/3$; the capacity of this BSC is 0.0817 bits thus, the asymptotic classical capacity of the 2/3-depolarizing quantum channel is close to 0.0817 bits.

**Example.** *The quantum erasure channel.* QEC is the quantum analog of the binary erasure channel discussed in Section 3.7; the incoming qubit is replaced with probability $p$ by a qubit in a special state, the *erasure state*, a state orthogonal to $\mid 0\rangle$ and $\mid 1\rangle$ states, and the receiver is informed about the identity of the qubit in error. The capacities of the QEC are known: the capacity to transmit classical information is $C^{QEC} = 1-p$; the one way quantum capacity is $Q^{QEC} = \max(0, 1-2p)$, and the two-way quantum capacity is $Q_2^{QEC} = 1-p$.

It is easy to establish that $C^{QEC} = 1-p$. The fraction of non-erased qubits for quantum erasure channels is $(1-p)$ and, according to Holevo's bound, $C^{QEC}$ cannot be larger than $(1-p)$. Recall that the capacity of a classical erasure channel is $(1-p)$; a classical channel emulating a quantum erasure channel with the input state in the basis $(\mid 0\rangle, \mid 1\rangle)$ and the output state in the basis $(\mid 0\rangle, \mid 1\rangle, \mid 2\rangle)$ is equal to $C = 1-p$.

To show that $Q^{QEC} = \max(0, 1-2p)$ we show first that $Q \geq 1-2p$. If we use a one-way hash coding scheme then the encoder adds two redundancy bits for each erased qubit and uses large block sizes; then the decoder recovers the phase and the amplitude of the erased qubits with probability one [40]. Now we show that when $p \geq 1/2$ then $Q^{QEC} = 0$; Alice uses the following scheme to send qubits through a $(2p-1)$ QEC channel to two receivers: she tosses a random coin and sends to Bob the information qubit and to Charles a qubit in a pure erasure state, or viceversa, depending on the outcome of the toss. This scheme is equivalent to the one when Alice is connected to Bob and to Charles via two $p \geq 1/2$ quantum erasure channels. The capacity of such channels must be zero to prevent cloning. A linear interpolation between a noiseless channel and a 50% QEC shows that $(1-2p)$ is indeed an upper bound for $Q$.

To prove that the two-way quantum capacity of a quantum erasure channel $Q_2^{QEC} = 1-p$ we consider a teleportation scheme like the one described in Section 2.13. Teleportation

requires EPR pairs e.g., pairs in state $\beta_{00} = 1/\sqrt{2}(\mid 00\rangle + \mid 11\rangle)$ and a classical channel. To simulate $n$ instances of a quantum erasure channel with erasure probability $p$ we use $n(1-p)$ EPR pairs to teleport information from the sender to the receiver. If the capacity of the channel were larger than $(1-p)$, the sender and the receiver would be able to deterministically increase their entanglement by local actions and communication, a proposition contrary to the laws of Quantum Mechanics [40]. This leads us to the next topic, the entanglement-assisted capacity of quantum channels.

## 3.24 Entanglement-Assisted Capacity of Quantum Channels

The discovery of superdense coding [33] followed by quantum teleportation [34] generated a renewed interest in quantum information theory and specifically in the effect of the entanglement on the capacity of quantum channels. Superdense coding doubles the capacity of a noiseless quantum channel, $C_E = 2C$ and, to our surprise, the enhancement of the capacity due to entanglement increases for some channels, as the noise level increases.

Superdense coding, Figure 91(bottom), uses an EPR pair in the state:

$$\mid \beta_{11}\rangle = \frac{1}{\sqrt{2}}\left(\mid 01\rangle - \mid 10\rangle\right).$$

The sender has one of the qubits of the pair and the receiver the other qubit of the pair. To communicate one of the four possible two-bit messages, $00, 01, 10$, or $11$, the sender applies to her own half of the EPR pair one of four transformations given by the four Pauli matrices, $\sigma_I, \sigma_x, \sigma_y$ and $\sigma_z$ and then sends the qubit. The receiver is now in the possession of an EPR pair in one of the four Bell states:

$$\mid \beta_{11}\rangle = \frac{1}{\sqrt{2}}\left(\mid 01\rangle - \mid 10\rangle\right), \quad \mid \beta_{10}\rangle = \frac{1}{\sqrt{2}}\left(\mid 00\rangle - \mid 11\rangle\right),$$

$$\mid \beta_{00}\rangle = \frac{i}{\sqrt{2}}\left(\mid 00\rangle + \mid 11\rangle\right), \quad \mid \beta_{01}\rangle = \frac{1}{\sqrt{2}}\left(\mid 01\rangle + \mid 10\rangle\right).$$

The receiver decides which one of the four messages has been sent after she identifies the Bell state:

$$\mid \beta_{11}\rangle \mapsto 00, \qquad \mid \beta_{10}\rangle \mapsto 01, \qquad \mid \beta_{00}\rangle \mapsto 10, \qquad \mid \beta_{01}\rangle \mapsto 11.$$

As we can see, the capacity of a noiseless quantum channel is doubled by the presence of entanglement.

Quantum teleportation, Figure 91(top), is in some sense the reciprocal of superdense coding. Quantum teleportation allows the state $\mid \varphi\rangle$ to be transmitted with the aid of an EPR pair and classical communication. The sender sends two bits of classical information and the receiver applies a transformation to its half of the EPR pair, guided by the two bits of classical information, as we have seen in Section 2.13.

Superdense coding and teleportation show that prior entanglement between the sender and the receiver does increase the capacity of a quantum channel. On the other hand, entanglement does not increase the capacity of a classical channel. The reason is the so called

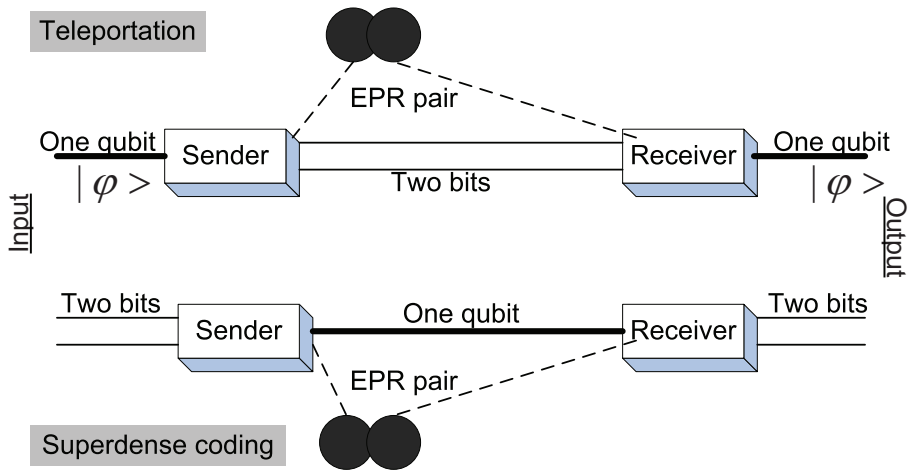



Figure 91: Superdense coding and quantum teleportation use entanglement as an ancillary resource for quantum communication. Superdense coding allows the sender to transmit one qubit instead of two bits of classical information with the aid of an EPR pair. Quantum teleportation allows the state $|\varphi\rangle$ to be transmitted with the aid of an EPR pair and classical communication; the sender sends two bits of classical information and the receiver applies a transformation to her half of the EPR pair, guided by the two bits of classical information. The quantum channel used for superdense coding and the classical channel used for teleportation are noiseless.

*constraint on causality* which does not allow us to increase the expectation of any local variable of the other system; its violation would allow us to send messages to the past.

In this section we consider a strategy called *measure/re-prepare* (MR) to simulate a quantum channel with the aid of a forward classical communication channel: (a) the sender and the receiver use the same orthonormal basis and the sender prepares the quantum channel input in a pure state in this basis and sends the information regarding the state through a classical channel; (b) the receiver then prepares a quantum state in the same orthonormal basis using the classical information she got. The amount of classical information sent through the classical channel in this case is called <u>forward classical communication cost</u>, $FCCC_{MR}$.

We begin our discussion of the entanglement-assisted capacity of quantum channels with an example when the exact expression of $C_E$ is known. The *d-dimensional depolarizing channel with depolarization probability* $p$, $D_p^{(d)}$, transmits a $d$-state of a quantum system unaltered with probability $(1-p)$ and randomizes the input state with probability $p$. We show that asymptotically, when $p \mapsto 1$, the entanglement-assisted capacity of the channel $D_p^{(d)}$ is $C_E = (d+1)C_1(D_p^{(d)})$ [45].

First, we calculate $C_E(D_p^{(d)})$, the entanglement-assisted capacity of $D_p^{(d)}$. In this case, as in the case of the erasure channel, we use the following methodology to obtain an exact value for $C_E$: we consider two quantum communication protocols that allow us to establish lower and upper bounds for the entanglement-assisted capacity of a specific type of quantum channel $\mathcal{N}$ characterized by the superoperator $\mathbf{N}$. If the lower and the upper bounds coincide, then we obtain exact expressions for the capacity of quantum channel $\mathcal{N}$.

The choices for the communication protocols used to establish the lower and the upper

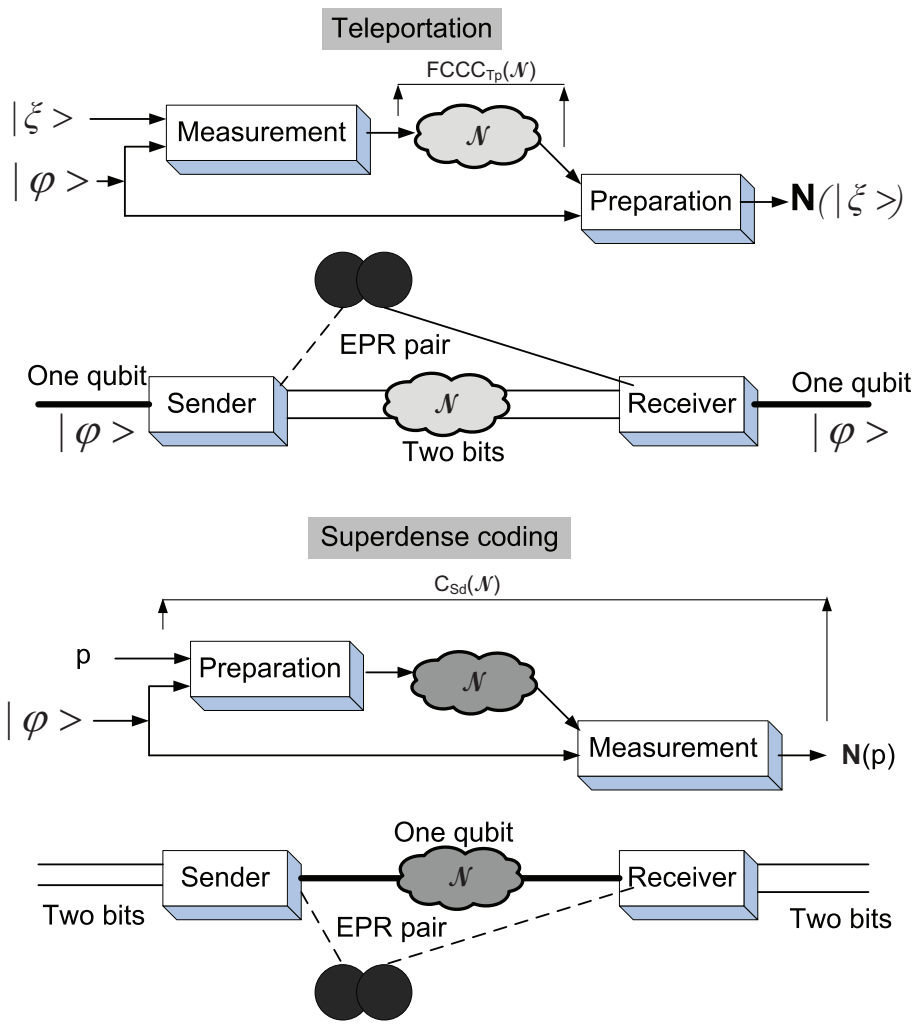



Figure 92: Superdense coding and quantum teleportation used to simulate the $d$-dimensional depolarizing channel with depolarization probability $p$. Superdense coding allows us to establish a lower bound for $C_E$; the noiseless quantum channel used to transport the sender's half of the EPR pair is substituted by a noisy quantum channel, $\mathcal{N}$. $C_{Sd}(\mathcal{N})$ is the lower bound on the entanglement-assisted capacity using the superdense coding protocol with channel $\mathcal{N}$. Quantum teleportation allows us to establish an upper bound for $C_E$; the noiseless classical channel between the sender and the receiver is substituted by a noisy one over a $d^2$-letter alphabet. $FCCC_{Tp}$, the classical communication cost of simulating a quantum channel with teleportation, cannot be smaller than the classical capacity of the quantum channel.

bounds for the entanglement-assisted capacity of the quantum channel $N$ are superdense coding, $Sd$, and teleportation, $Tp$, depicted in Figure 91. To establish a lower bound for $C_E$ we consider a modified setup for the superdense coding in Figure 92 (bottom); we substitute the noiseless quantum channel used to transport the sender's half of the EPR pair by a noisy quantum channel, $\mathcal{N}$. We denote by $C_{Sd}(\mathcal{N})$ the lower bound on the entanglement-assisted capacity estimated for the superdense coding protocol for channel $\mathcal{N}$.

For the upper bound we consider a teleportation setup, Figure 92 (top), where the noiseless

classical channel between the sender and the receiver is substituted by a noisy one over a $d^2$-letter alphabet. Due to causality considerations, $FCCC_{Tp}$ the classical communication cost of simulating a quantum channel with teleportation cannot be smaller than the classical capacity of the quantum channel. This argument justifies the use of the teleportation protocol for the upper bound of the entanglement-assisted capacity of $D_p^{(d)}$.

It can be shown [45] that there are one-to-one mappings between a $d$-dimensional quantum channel with depolarization probability $p$ and $d^2$-ary symmetric channel with randomizing probability $p$; the two mappings are done via superdense coding and teleportation. Moreover, the upper and lower bounds for the entanglement-assisted capacity of the quantum channel $D_p^{(d)}$ obtained through simulation of $D_p^{(d)}$ with superdense coding and teleportation coincide

$$C_E\left(D_p^{(d)}\right) = C_{sd} = FCCC_{Tp} = 2\log d - H_{d^2}\left(1 - p\frac{d^2-1}{d^2}\right).$$

To compute $C_1(D_p^{(d)})$ we use the measure-reprepare (MR) strategy and consider a quantum channel with an input consisting of an arbitrary orthonormal basis $\{\mid 1\rangle, \mid 2\rangle, \ldots, \mid d\rangle\} \in \mathcal{H}_d$. We assign an equal probability $1/d$ to each basis vector at the input and we perform a von Neumann measurement of the output of the channel in the same basis. This quantum channel behaves like a $d$-ary symmetric classical channel with probability of error $p$ thus,

$$C_1\left(D_p^{(d)}\right) = \log d - H_d\left(1 - p\frac{d-1}{d}\right).$$

In this expression $H_d(q)$ is the Shannon entropy of a $d$-ary distribution consisting of an element of probability $q$ and $(d-1)$ elements of probability $(1-p)/(d-1)$ each

$$H_d(q) = -q\log q - (1-q)\log\left(\frac{1-q}{d-1}\right).$$

Clearly, when $p \mapsto 1$ then $C_1(D_p^{(d)}) \mapsto \log d$. When $p \geq d/(d+1)$ the $d$-dimensional depolarizing channel can be simulated classically; the amount of classical forward communication required:

$$FCCC_{MR}\left(D_p^{(d)}\right) = \log d - H_d\left(d - p(d - \frac{1}{d})\right).$$

We continue our discussion of the entanglement-assisted capacity of a quantum channel with the presentation of the general case. Consider a quantum channel $N$ mapping the input Hilbert space $\mathcal{H}_{in}$ to $\mathcal{H}_{out}$ and assume that the sender, Alice, and the receiver, Bob, have an unlimited supply of EPR pairs. If $C_E$ denotes the entanglement-assisted classical capacity of $N$ then $C_E$ is the maximal quantum mutual information

$$C_E = \max_{\rho\in\mathcal{H}_{in}}\left[S(\rho) + S\left(\mathbf{N}\left(\rho\right)\right) - S\left(\left(\mathbf{N}\otimes\mathbf{I}\right) \mid \psi_{pur}\rangle\right)\right].$$

In this expression, that resembles the expression of capacity of a classical channel, $S(\rho) \in \mathcal{H}_{in}$ and $S\left(\mathbf{N}\left(\rho\right)\right)$ are the von Neumann entropy of the input and, respectively, the output of the quantum channel; $S\left(\left(\mathbf{N}\otimes\mathbf{I}\right) \mid \psi_{pur}\rangle\right)$ is the von Neumann entropy of a purification $\mid \psi_{pur}\rangle$ of

the input density matrix $\rho$, a partial trace over a reference system $R$, $\mathrm{tr}_R\left[\mid \psi_{pur}\rangle\langle\psi_{pur} \mid\right] = \rho$. In this model half of the input, the one corresponding to the entangled state Bob holds at the beginning of the exchange, is sent through the reference system represented by the identity channel $I$ and the other half, Alice's half, is sent through the noisy channel $N$, Figure 89.

The analysis of the entanglement-assisted capacity of a quantum channel in [47] proves first a lower bound for $C_E$:

$$C_E \geq \max_{\rho \in \mathcal{H}_{in}} \left[S(\rho) + S\left(\mathbf{N}\left(\rho\right)\right) - S\left(\left(\mathbf{N} \otimes \mathbf{I}\right) \mid \psi_{pur}\rangle\right)\right].$$

The proof uses frequency-typical subspaces, a technique common in classical information theory. Then it proves an upper bound:

$$C_E \leq \max_{\rho \in \mathcal{H}_{in}} \left[S(\rho) + S\left(\mathbf{N}\left(\rho\right)\right) - S\left(\left(\mathbf{N} \otimes \mathbf{I}\right) \mid \psi_{pur}\rangle\right)\right].$$

The proofs in [47] are quite intricate and will not be presented here. In conclusion, the entanglement-assisted capacity $C_E$ of a quantum channel is given by an expression similar to that for a purely classical channel as a maximum, over the channel input states $\rho$, of the entropy of the channel input plus the entropy of the channel output minus their joint entropy. Here, the joint entropy is defined as the entropy of an entangled purification of $\rho$ after half of it has passed through the channel.

## 3.25 Additivity and Quantum Channel Capacity

The channel capacity quantifies the amount of classical or quantum information that can be protected through error correction a topic mentioned in Section 3.26 and discussed in depth in Chapters 4 and 5. While a classical channel has a definite capacity to carry classical information, it may not be possible to find a simple expression for the capacity of a quantum channel [391].

In 2008, Smith and Yard showed that *communication is possible over zero-capacity quantum channels* [399]. This result raised new questions regarding quantum channel capacity as a complete description of a quantum channel's ability to transmit quantum information. In 2009, Matthew Hastings provided an answer to one of the most important open questions in quantum information theory showing that the minimum entropy output of a communication channel is not additive [204]. We now discuss briefly these two results.

There are several classes of *zero-capacity channels* [399] useless for quantum information transmission when used alone; one class consists of quantum channels when the joint state of the output and the environment is symmetric under interchange and information transmission would violate the no-cloning theorem. The other class consists of the so called *entanglement-binding* channels which produce only weakly entangled states. The amazing result is that when we combine one channel from each of these classes with a channel from the other the resulting channel capacity is positive under some conditions. This means that when two parties attempt to communicate on channels $\mathcal{N}_1$ and $\mathcal{N}_2$ using separate encoders, $E_1$ and $E_2$ as well as separate decoders, $D_1$ and $D_2$, the communication fails; when the two channels are used in parallel the communication is possible, as shown in Figure 93.

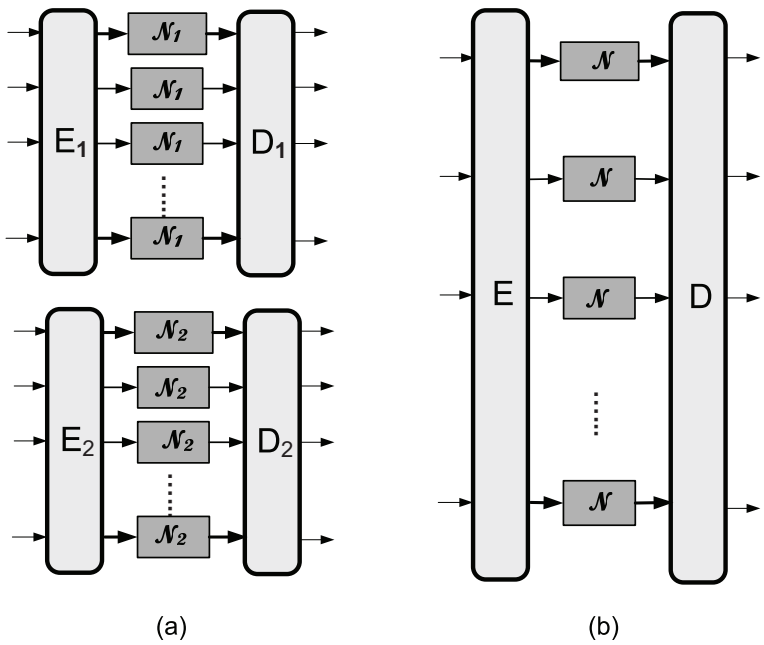



Figure 93: Communication over zero-capacity quantum channels. (a) Two parties attempt to communicate over zero-capacity channels $\mathcal{N}_1$ and $\mathcal{N}_2$ using separate encoders, $E_1$ and $E_2$, as well as separate decoders, $D_1$ and $D_2$. (b) The sender's encoder, $E$ has access to the inputs of all channels and the receiver's decoder $D$ acts jointly on the output of all channels; communication is possible [399].

This result requires two new concepts introduced in [398], $\mathcal{P}(\mathcal{N})$, the private capacity of a quantum channel $\mathcal{N}$ and $\mathcal{Q}_{\mathcal{A}}(\mathcal{N})$, the assisted capacity of $\mathcal{N}$. The private capacity is the rate at which the channel $\mathcal{N}$ can be used to send classical data secure against eavesdropping for quantum key distribution protocols [32]; the assisted capacity reflects the free use of arbitrary symmetric channels to assist quantum communication over $\mathcal{N}$. The assisted capacity is at least as large as half of channel's private capacity [399].

The additivity conjecture we discuss next would make possible to compute the classical capacity of a quantum channel. Recall from Section 3.12 that the classical capacity of a quantum channel $\mathcal{N}$ characterized by the mapping $\mathcal{E}$ of the densities $\rho_i$ of an ensemble of mixed states with $p_i$ the apriori probability of mixed state $i$ is given by the Holevo expression

$$\chi(\mathcal{N}) = S\left(\sum_i p_i \mathcal{E}(\rho_i)\right) - \sum_i p_i S\left(\mathcal{E}(\rho_i)\right).$$

with $S(\rho)$ and $S(\rho_i)$ the von Neumann entropy of the overall ensemble of states and of the individual state $i$, respectively.

Given two quantum channels, $\mathcal{N}_1$ and $\mathcal{N}_2$, the *additivity of the classical information capacity* requires an answer to the question if

$$\max\left[\chi(\mathcal{N}_1) \otimes \chi(\mathcal{N}_2)\right] = \max\left[\chi(\mathcal{N}_1)\right] + \max\left[\chi(\mathcal{N}_2)\right].$$

Shor has shown [390] that there are three additivity problems equivalent with the additivity of the Holevo channel capacity. The first, the *additivity of the minimum output entropy of a quantum channel*, reflects the desire to minimize the second term in the expression of Holevo capacity and is formulated as

$$\min_{\rho} [S(\mathcal{N}_1 \otimes \mathcal{N}_2)(\rho)] = \min_{\rho} S(\mathcal{N}_1(\rho)) + \min_{\rho} S(\mathcal{N}_1(\rho))$$

The other two problems are the *additivity of the entanglement of formation,* and *strong superadditivity of the entanglement of formation* [390].

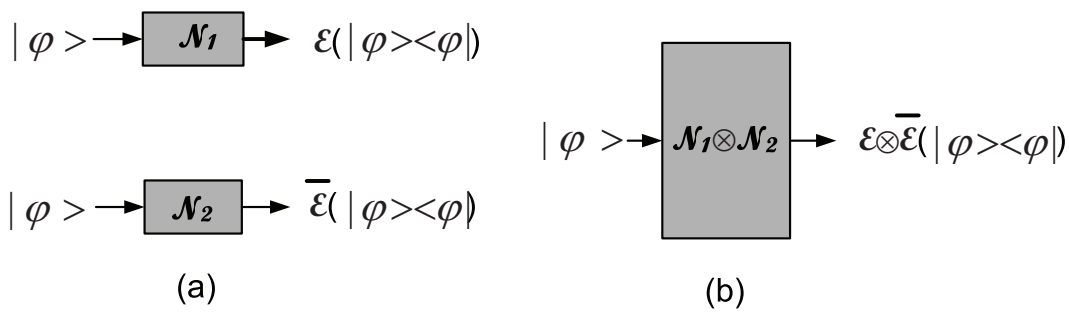



Figure 94: Superadditivity of the minimum output entropy. (a) Minimization of the output entropy $S(\mathcal{E}(\rho))$ and $S(\bar{\mathcal{E}}(\rho))$ of two channels $\mathcal{N}_1$ and its complex conjugate, $\mathcal{N}_2$ with a pure state $\mid \varphi\rangle$ as input and possibly a mixed state with density $\rho$ at the output. (b) The entangled input state $\mid \varphi\rangle$ is transformed according to the mapping $\mathcal{E} \otimes \bar{\mathcal{E}}$ and the entropy of the output state is minimized [204].

Hastings constructs a counter-example for the additivity of the minimum output entropy and shows that the additivity conjecture is false [204]. The basic idea is to model a situation when the unitary evolution of a system is determined by an unknown state of the environment. This is done with a pair of channels, $\mathcal{N}_1$, and its complex conjugate, $\mathcal{N}_2$, Figure 94; each channel randomly chooses a unitary transformation $\mathbf{U}_i$, $1 \leq i \leq D$ from a small set and applies that transformation to the input density $\rho$:

$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i^\dagger \rho U_i \quad \text{and} \quad \bar{\mathcal{E}}(\rho) = \sum_{i=1}^{D} p_i \bar{U}_i^\dagger \rho \bar{U}_i.$$

In this expression $U_i$ are $N \times N$ unitary matrices chosen at random from the Haar measure, the probabilities $p_i$ are also randomly chosen and roughly equal, and $1 << D << N$. The main result in [204] is: *when $D$ and $N$ are sufficiently large then there is a non-zero probability that the random choice of $U_i$ and $p_i$ will lead to a channel with a mapping $\mathcal{E} \otimes \bar{\mathcal{E}}$ such that*

$$S^{min}(\mathcal{E} \otimes \bar{\mathcal{E}})(\rho) < S^{min}(\mathcal{E}(\rho)) + S^{min}(\bar{\mathcal{E}}(\rho)) = 2S^{min}(\mathcal{E}(\rho)).$$

The superadditivity of the minimum output entropy and the equivalence of the four additivity problems indicate that the problem of quantum channel capacity remains open.

We conclude the chapter with an overview of applications of information theory.

## 3.26 Applications of Information Theory

Classical information theory is critical for the development of modern computing and communication systems. Error control mechanisms enable reliable transport of information in the Internet and allow us to build reliable computing and communication systems. Other important applications of classical information theory are in the area of secure communications and information storage. Since the early 1990s quantum information theory has produced significant results in areas such as quantum error correcting codes, quantum data compression, superdense coding, quantum teleportation, and entanglement concentration; these results give us hope that we will be able to compute and communicate more reliably and securely in the future. Table 8 summarizes the applications of classical and quantum information [46].

Table 8: Application of classical and quantum information theory.

| Application | Classical information | Quantum information |
|---|---|---|
| Error-correction | Classical error-correcting codes. | Quantum error-correcting codes. |
| Fault-tolerance | von Neumann's analysis: error probability per gate smaller than a threshold. | The decoherence time per gate below a threshold estimated to be $10^{-6}$ to $10^{-2}$ seconds. |
| Communication | Bits over classical channels. | Qubits over quantum channels. |
| Entanglement-assisted communication | Not possible. | Superdense coding, Quantum teleportation. |
| Cryptographic key distribution | Classical protocols are insecure. | Quantum protocols are secure |
| Cryptographic protocols | Insecure against attacks by quantum computers. | Secure against attacks by quantum computers. |
| Two-party bit commitment | Known protocols insecure against attacks by quantum computer. | Known protocols insecure against attacks by quantum computer. |

Error correction is one of the most important application of information theory and, in turn, a critical element for fault-tolerance. A preview of classical and quantum error correction, topics discussed in depth in the next two chapters, gives us some idea of the similarities as well as the dissimilarities of classical and quantum information processing and of the challenges posed by practical applications of quantum information theory.

Quantum error correction is possible; indeed, consider systems $\mathcal{A}$ and $\mathcal{B}$, the input and the output of a quantum channel, and $\mathcal{E}$ the transformation of the input state density, $\rho^{\mathcal{A}}$, carried out by the channel. If the bipartite system $(\mathcal{A}, \mathcal{B})$ in the product state $\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}$ is subject to a transformation $\mathcal{E}$, if there is an auxiliary system $\mathcal{C}$ in some standard state $\rho^{\mathcal{C}}$ and some unitary operator $\mathbf{U}$ acting on the joint system $(\mathcal{A}, \mathcal{B}, \mathcal{C})$, and if

$$\mathcal{E}\left(\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}\right) = \mathrm{tr}_{\mathcal{C}}\left[\mathbf{U}\left(\rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}} \otimes \rho^{\mathcal{C}}\right)\right]$$

then, after applying the transformation $\mathbf{U}$, we can perform a measurement on $\mathcal{C}$ that reveals how close the two states $\rho^{\mathcal{A}}$ and $\rho^{\mathcal{B}}$ are, and then ignore the auxiliary system $\mathcal{C}$.

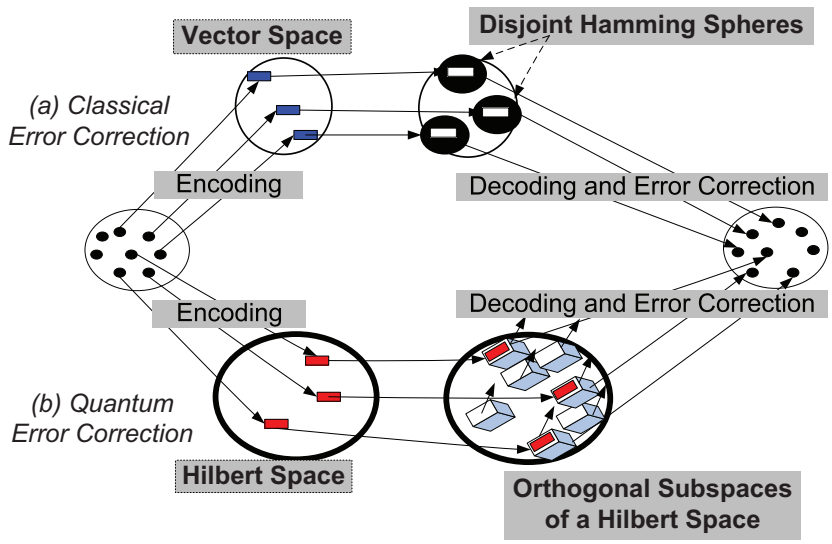



Figure 95: (a) Classical error correction - messages are encoded as codewords in a sparsely populated n-dimensional space; codewords are surrounded by disjoint Hamming spheres of $n$-tuples. Error correction implies that all $n$-tuples within a Hamming sphere are decoded as the codeword at the center of the sphere. (b) Quantum Error Correction - a logical qubit is encoded as several physical qubits. The states of the physical qubits used to encode distinct logical qubits span orthogonal subspaces; the transformations applied to these subspaces to determine if an error has occurred and to identify the error should preserve the relationships between the physical qubits used to encode a single logical qubit.

As we can see in Figure 95(a) classical information is encoded into codewords sparsely distributed in the input space of the communication channel. The codewords are affected by errors during the transmission through the channel; each codeword is surrounded now by a sphere of possible outputs, and, as long as these spheres are disjoint, the decoding process is able to reconstruct the original codeword. The number of distinct messages is the ratio of the volume of the output space to the volume of a sphere and the effective channel capacity is the logarithm of this ratio.

Inspired by classical error correction, quantum error correction separates the space of encoded states into orthogonal subspaces of the Hilbert space and then performs a measurement of the error state without disturbing the encoded state. Any transformation of quantum information can be regarded as a unitary interaction of the quantum system encoding the information with ancilla qubits initially in a null state ($|0\rangle$).

There is a conceptual difference between coding classical and quantum information; *the objective of classical coding is to increase the redundancy and allow each codeword to "represent" an equivalence class of $n$-tuples that will be decoded as the codeword, while the objective of quantum coding is to increase the distinguishability of the codeword states.*

When we transmit quantum information we encode every basis vector in the quantum message space as a quantum codeword in a Hilbert space Figure 95(b). Quantum noise and decoherence affect the quantum state. As a result, the output for a given input lies with a high probability is some subspace of the original Hilbert space. As long as the overlap of

these subspaces is either null, or minimal, we can reliably decode each quantum codeword. We note that quantum encoding must preserve the phases and the amplitudes of all qubits of the quantum codeword. The vectors describing quantum codewords in each of these subspaces are perturbed differently. The decoding process must preserve superpositions of the basis vectors, therefore the amplitudes and the phases of the perturbations must be the same as the initial ones in order to recreate the original states. Questions such as: How do we carry out a measurement to determine if the quantum codeword is in error without altering the state of the quantum register? or How do we distinguish non-orthogonal states? do not have trivial answers.

## 3.27 History Notes

In his quest to provide a rigorous mathematical foundation for communication Claude Shannon created in late 1940 the discipline of information theory. The first "golden age" of information theory research can be traced back to the work carried out at Bell Labs starting in late 1940s and continued at MIT till the end of 1970s by the group around Claude Shannon.

In his 1995 Shannon Lecture [155] David Forney writes: "Our field has a rather remarkable history, in that it sprang almost fully fledged from the brow of one individual. In his original paper [375] Claude Shannon both posted the fundamental problems of information theory and also, to a large extent, answered them. Shannon's good news was this: for every memoryless channel there exist a parameter $C$, the channel capacity, such that the probability of error $\Pr(E)$ can be made arbitrarily small for any rate $R$ less than $C$ by use of an appropriate code and decoder. Good codes exist and, in fact, a randomly chosen code will with high probability turn out to be good."

The article "Profile: Claude E. Shannon" by John Hogan in the January 1990 issue of Scientific American notes that: "Shannon fits the stereotype of an eccentric genius to a T. At Bell Labs (and later at MIT where he returned in 1958 until his retirement in 1978) he was known for riding in the halls on a unicycle sometimes juggling as well. At other times he hopped along the hallways on a pogo stick. He was a lover of gadgets and among other things built a robotic mouse that solved mazes and also a computer called the Throbac - THrifty ROman-numeral BAckward-looking Computer - that computed in roman numerals.... In the 1990s, in one of the life's tragic ironies, Shannon came down with Alzheimer's disease, which could be described as the insidious loss of information in the brain. The communication channel to one's memory - one's past and one's very personality - is progressively degraded until every effort at error correction is overwhelmed and no meaningful signal can pass through. The bandwidth falls to zero. The extraordinary pattern of information processing that was Claude Shannon finally succumbed to the depredation of thermodynamic entropy in February 2001. But some of the signal generated by Shannon lives on, expressed in the information technology in which our lives are immersed."

The group of Thomas Cover at Stanford continues to this day the tradition of Claude Shannon. Many distinguished researchers in the field of information theory are either disciples or collaborators of this group. Papers describing significant research results as well as textbooks have been published by the members of the group [107].

The roots of quantum information theory can be traced back to John von Neumann who introduced the von Neumann entropy in early 1930s [442]. The relation between information

and thermodynamics was captured in 1961 by Landauer [265]. During the next two decades Charles Bennett provided deeper insights into the logical reversibility [30] and the thermodynamics of computations [31]; he was able to reconcile the Maxwell Demon paradox with the second law of thermodynamics.

In 1973 Alexander Holevo found an upper bound of the accessible information in a quantum measurement [211]; Holevo bound was probably the first major result in quantum information theory.

The last decade of the twentieth century witnessed a spectacular development of quantum information theory: in the late 1980s, Charles Bennett and Gilles Brassard developed a practical system of quantum cryptography then, in 1992, Bennett and Wiesener presented a scheme for superdense coding using EPR pairs. In 1993, Bennett and Brassard, in collaboration with Claude Crépeau, Richard Jozsa, Asher Peres, and William Wootters, discovered quantum teleportation. In mid 1990s Bennett, Smolin, Wootters, DiVincenzo, and others developed a quantitative theory of entanglement and introduced several techniques for faithful transmission of classical and quantum information through noisy channels; during the same period, Benjamin Schumacher, Richard Jozsa, Seth Lloyd, and others had major contributions to the study of noisy memoryless quantum channels.

## 3.28 Summary and Further Readings

Landauer principle [265] relates information with thermodynamic entropy: *when a bit of information is erased then the thermodynamic entropy of the environment increases by at least* $k_B \ln 2$*, with* $k_B$ *- Boltzmann's constant.* Once the thermodynamic effect of information erasure was established, the physical nature of information could be quantified.

It seems so obvious now that *the amount of information that can be stored or transmitted by a physical system is related to the number of states the system can find itself in* that one may wonder why it took so long to fully comprehend the physical nature of information and the role of entropy. The entropy, an important concept in thermodynamics, Shannon's information theory, as well as quantum information theory, is a function of the logarithm of the number of states of a system; this logarithm is equal to the number of bits required to uniquely identify the state of the system, in other words to "label" a state. The label becomes an element of the state and any state transformation will affect the label; to prepare a system in a certain state we have to erase the label of the previous state.

Shannon entropy is used to characterize a source of classical information as well as the properties of a classical communication channel. The properties of joint and conditional entropy are summarized in Table 5. The mutual information between the two random variables is $I(X;Y) = H(X) + H(Y) - H(X,Y)$. If the inputs and the outputs of the channel are the random variables $X$ and $Y$, respectively, then the channel capacity, $C$, defined as the maximum rate of information transmission over the channel is $C = \max_{p(X)} I(X;Y)$. Two important theorems are the cornerstones of Shannon's information theory, the *source coding theorem* and the *channel coding theorem.*

The original papers of Claude Shannon [375, 376, 377, 378] and Robert Fano [145] as well as highly regarded textbooks on classical information theory by Robert Ash [15], Robert Gallagher [168], and Thomas Cover and Joy Thomas [107] published in 1965, 1968, and 1991,

respectively, provide the necessary material to understand the basic concepts of classical information theory.

In spite of the significant progress made during the past half century, there are profound questions that still remain to be answered unequivocally by a unified information theory. Some believe that information has three aspects: (i) a syntactic aspect - the relationship between the symbols of the alphabet used to construct a message, (ii) a semantic aspect - the meaning of the message, and (iii) a pragmatic aspect - the actions taken by the parties involved in the exchange. Shannon's theory ignores the semantic aspect of information and cannot describe quantum and biological information [262] which, most certainly, will play a critical role in the third millennium.

Nowadays, some believe that the focus of information theory should migrate from the communication channel to the recipient of information, and from the process of communication to the consequences of receiving information. Indeed, in many instances the timeliness of information is important; the context of the information and the goals of the recipient of information (whether a human or a machine) affect the usefulness of information.

John von Neumann used thermodynamic arguments to define the entropy of a quantum system with the density matrix $\rho$ as $S(\rho) = -\mathrm{tr}\,[\rho \log \rho]$; the expression $\rho \log \rho$ is well defined as the density matrix is positive semidefinite. It is easy to see that von Neumann entropy is the Shannon entropy of the eigenvalues $p_i$ of a diagonal density matrix $\rho$. Indeed, when $\rho$ is a diagonal matrix with eigenvalues $p_i$ then $\mathrm{tr}(\rho) = \sum_i p_i = 1$ and $\rho \log \rho$ is also diagonal with eigenvalues $p_i \log p_i$. A source of quantum information is not fully characterized by the von Neumann entropy; two different sources of quantum information, may have the same density matrix thus, the same von Neumann entropy, but be very different from the quantum information theoretical point of view.

A quantum channel is a completely positive, trace-preserving linear map from density matrices to density matrices, $\rho \mapsto \sum_i A_i \rho A_i^\dagger$ with $\sum_i A_i^\dagger A_i = I$. An alternative definition of a quantum channel is a partial trace of a unitary transformation on a larger Hilbert space.

A quantum channel has four distinct capacities [47]:

1. $C$, the ordinary capacity for transmitting classical information - the maximum rate at which classical bits can be transmitted reliably through a quantum channel encoded as the state of a quantum system and decoded after a measurement of the quantum state.

2. $Q \leq C$, the ordinary capacity for transmitting quantum states - the maximum asymptotic rate at which qubits can be transmitted reliably through a quantum channel.

3. $Q_2 \geq Q$, the classically-assisted quantum capacity - the asymptotic rate for reliable qubit transmission with the assistance of a two-way classical channel.

4. $C_E \geq C \geq Q$, the entanglement-assisted classical capacity - the maximum asymptotic rate of reliable bit transmission with the assistance of unlimited prior entanglement between the sender and receiver.

Table 9 summarizes the properties of classical and quantum information and the main results in classical and quantum information theory.

Table 9: Side-by-side properties of classical and quantum information and main results in classical and Quantum Information Theory.

| Property | Classical information | Quantum information |
|---|---|---|
| Redundancy | Classical bits can be cloned | Qubits cannot be cloned |
| Distinguishability | Classical states are always distinguishable | Only orthogonal states can be reliably distinguished |
| Source of information | Alphabet $\mathcal{A}\{a_1, a_2, \ldots, a_n\}$ probability distribution $p_i$ | Pure/mixed quantum states density matrix $\rho$ |
| Entropy of a source | Shannon Entropy $H = -\sum p_i \log p_i$ | von Neumann Entropy $S = \mathrm{tr}\left[\rho \log \rho\right]$ |
| Measurements | Do not alter the state | Alter the state; project the state onto a subspace |
| Communication channel | $X, Y$ input/output RVs $X \mapsto Y$ | Input/output states $\rho$, $\mathcal{E}\left(\rho\right)$ $\rho \mapsto \mathcal{E}\left(\rho\right)$ |
| Trace distance input - output | $1/2 \sum_x \mid (p_X(x) - p_Y(x)) \mid$ | $1/2\mathrm{tr} \mid \rho - \mathcal{E}\left(\rho\right) \mid$ |
| Fidelity of the channel output | $F\left(p_X(x), p_Y(y)\right) = \sum_x \sqrt{p_X(x) p_Y(y)}$ | $F(\rho, \sigma) = \left\{\mathrm{tr}\left[\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right]\right\}^2$ |
| Correlation of channel input- output | Mutual Information $I(X;Y)$ | Coherent Information $I\left(\rho, \mathcal{E}\right) = S\left(\mathcal{E}\left(\rho\right)\right) - S\left(\rho, \mathcal{E}\right)$ Entropy exchange $S_e(\rho, \mathcal{E}) = S(\rho^{(E)}_{after})$ Entanglement fidelity $F_e(\rho, \mathcal{E}) = \langle \varphi^{(Q,R)} \mid (\mathcal{I}_R \otimes \mathcal{E}) \left(\mid \varphi^{(Q,R)}\rangle\langle\varphi^{(Q,R)} \mid\right) \mid \varphi^{(Q,R)}\rangle$ |
| Information from a measurement of channel output | $I(X;Y) = H(X) - H(X\|Y)$ | Holevo bound $I(X;Y) \leq S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i)$ |
| Fano inequality | $H(X\|Y) \leq H(p_{err}) + p_{err} \log(\mid \mathcal{X} \mid)$ | $S_e\left(\rho, \mathcal{E}\right) \leq h(p) F_s(\rho, \mathcal{E}) + \left[1 - F_s(\rho, \mathcal{E})\right] \log(d^2 - 1)$ |
| Data processing inequality | $X \mapsto Y \mapsto Z$ $H(X) \geq I(X;Y) \geq I(X;Z)$ | $\rho \mapsto \mathcal{E}_1(\rho) \mapsto \mathcal{E}_2 \circ \mathcal{E}_1(\rho)$ $S(\rho) \geq I(\rho, \mathcal{E}_1) \geq I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$ |
| Noiseless channel coding | Source entropy $H(X)$ $n_{bits} = H(X)$ | Pure states $\{\rho_i, p_i\}$ $n_{qubits} = S\left(\sum_i p_i \rho_i\right)$ |
| Noisy channel capacity | $C = \max I(X;Y)$ | Classical capacity - $C(\mathcal{E})$ $C(\mathcal{E}) = \max_{p_i, \rho_i}\left[S(\sigma) - \sum_i p_i S(\sigma_i)\right]$ $\sigma_i = \mathcal{E}(\rho_i), \sigma = \sum_i p_i \sigma_i$ Quantum capacity - $Q \leq C$ Classically-assisted quantum capacity - $Q_2$, $Q \leq Q_2 \leq C$ Entanglement-assisted classical capacity - $C_E$, $Q \leq C \leq C_E$ |

Table 10: Resources necessary to transmit one unit of information.

| | | | |
|---|---|---|---|
| 1 bit | $\leq$ | 1 qubit | To transmit one bit we need at least one qubit. |
| 1 ebit | $\leq$ | 1 qubit | To transmit one ebit we need at least one qubit |
| 1 qubit | $\leq$ | 1 ebit + 2 bits | To transmit one qubit we need at least one ebit and two bits |
| 2 bits | $\leq$ | 1 ebit + 1 qubit | To transmit two bits we need at least one ebit and one qubit |

We use the term "ebit" to describe a source of a shared pair of maximally entangled particles. Table 10 summarizes the amount of resources necessary to transmit one unit of information [37].

The papers co-authored by Charles Bennett [33, 34, 35, 37, 38, 39, 40, 42, 44, 46, 47, 48] introduce many basic concepts of Quantum Information Theory. Major contribution are also due to Benjamin Schumacher [369, 370, 371, 372, 374], Richard Jozsa [221, 223, 227], Seth Lloyd [274, 276], and Peter Shor [385].

A number of dissertations on quantum information theory were written, by Benjamin Schumacher (under Archibald Wheeler) in 1990, Howard Barnum and Michael Nielsen [307] in 1999 and 2000 respectively. The textbook on Michael Nielsen and Isaac Chuang [308] covers many aspects of quantum information theory.

The last decade of the twentieth century witnessed a very spectacular development of the field. In 1992, Charles Bennett and Steven Wiesener presented a scheme for superdense coding using EPR pairs. An influential paper on teleportation of quantum states was published in 1993 by Charles Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William Wootters [34]. Quantum error correcting codes were introduced during the $1994 - 1998$ years by Richard Calderbank, Peter Shor, Andrew Steane, Emmanuel Knill, Raymond Laflamme, Daniel Gottesmann, and others [81, 174, 245, 381, 383].

Benjamin Schumacher, Richard Jozsa, Seth Lloyd, Howard Barnum, Michael Nielsen contributed to the study of noisy memoryless quantum channels [22, 276, 369, 370, 371]. In 1998, Charles Bennett and Peter Shor published a seminal paper on quantum information theory [44] and in 2002 Charles Bennett, Peter Shor, John Smolin, and Ashish Thapliyal discussed the entanglement-assisted capacity of a quantum channel [47].

In 2009, Matthew Hastings provided an answer to one of the most important open questions in quantum information theory showing that the minimum entropy output of a communication channel is not additive [204]. As Shore notes in [391] "It therefore remains an open question whether it is possible to find a simple expression for the information capacity of a quantum channel - similar to Shannon's formula for classical channels - that would make practical calculation of the quantum channel capacity possible."

## 3.29 Exercises and Problems

**Problem 1.** Prove that binary entropy is a concave function of $p$.

**Problem 2.** Let $p(x)$ be the distribution of a discrete random variable $X$ and $p_i, 1 \leq i \leq n$, the probability of individual outcomes $x_i$ of $X$. Show that

$$H(X) \leq \log n$$

with equality if and only if we have a uniform distribution.

**Problem 3.** Prove the chain rule for the entropy of $n$ random variables $X_1, X_2, \ldots, X_n$:

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i-1}, \ldots, X_1).$$

**Problem 4.** Let $X$ and $Y$ be independent binary variables with probability density functions $p_X(x)$ and $p_Y(y)$. $Z$ is also a random variable and $z = x + y \mod 2$. Depict the relationship among $H(X), H(Y), H(Z)$, the conditional entropy, the joint entropy, and the mutual information using the representations in Figure 73 (a) and (b). Discuss the accuracy of each representation.

**Problem 5.** Consider a classical binary symmetric channel with the error probability $p = 1/3$; show that the capacity of this BSC is 0.0817 bits.

**Problem 6.** Prove Jensen's inequality:

$$\sum_i p_X(x_i) f(x_i) \geq f\left(\sum_i p_X(x_i) x_i\right)$$

where $X$ is a random variable and $f(x)$ is a convex function.

**Problem 7.** Prove Fano's inequality:

$$H(p_{err}) + p_{err} \log(\mid \mathcal{X} \mid) \geq H(X|Y).$$

with $X$ a random variable with the probability density function $p_X(x)$ and $\mid \mathcal{X} \mid$ the number of elements in the range of $X$. $Y$ is another random variable related to $X$, with the conditional probability $p_{Y|X}(y|x)$. $p_{err} = \text{Prob}(\hat{X} \neq X)$ is the error when we estimate $X$ based on the observation of $Y$ as $\hat{X} = f(X)$.

**Problem 8.** Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $\beta_1, \beta_2, \ldots, \beta_n$ be non-negative real numbers. Show that

$$\sum_{i=1}^{n} \alpha_i \log \frac{\alpha_i}{\beta_i} \geq \left(\sum_{i=1}^{n} \alpha_i\right) \log \frac{\sum_{i=1}^{n} \alpha_i}{\sum_{i=1}^{n} \beta_i}.$$

**Problem 9.** Prove the grouping property of Shannon entropy:

$$H(p_1, p_2, \ldots, p_n) = h(p_1) + (1 - p_1) H\left(\frac{p_2}{1 - p_1}, \frac{p_3}{1 - p_1}, \ldots, \frac{p_n}{1 - p_1}\right)$$

with $h(p) = -p \log p - (1 - p) \log(1 - p)$ and $H(p_1, p_2, \ldots, p_n)$ the Shannon entropy. Show also that

$$H\left(\frac{p_2}{1 - p_1}, \frac{p_3}{1 - p_1}, \ldots, \frac{p_n}{1 - p_1}\right) \leq \log(n - 1).$$

**Problem 10.** A multiple access channel can be modelled as is a channel with $n$ sets of inputs and one output. For example the input is $x_1, x_2, \ldots, x_n$ with $x_i = \{0, 1\}$ and the output $y = \{0, 1, \ldots n\}$. What is the capacity of the channel? When $n = 2$ and the two sources transmit with rates $R_1$ and $R_2$, respectively, plot the achievable rate.

A broadcast channel consists of a transmitter and $n$ receivers. Such a channel can be modelled as $n$ binary symmetric channels $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n$ with capacities $C_1 > C_2 > \ldots > C_n$ and the probability of error $p_1 < p_2 < \ldots < p_n < 1/2$, respectively. Call $R$ the rate of reliable communication through the channel; plot $R$ function of the probability of error $p$.

**Problem 11.** Let $X = X_1, X_2, \ldots X_i, \ldots$ be a stationary Markov chain. Show that $H(X_i|X_1)$ is monotonously increasing in spite of the fact that $H(X_i)$ does not change with $i$ by stationarity if the Markov chain. Discuss how this result is related to the second law of thermodynamics.

**Problem 12.** Prove the quantum Fano inequality:

$$S_e\left(\rho, \mathcal{E}\right) \leq h(p) F_s(\rho, \mathcal{E}) + \left[1 - F_s(\rho, \mathcal{E})\right] \log(d^2 - 1).$$

In this expression $d$ is the dimension of the Hilbert space of the quantum system $Q$ and $h(p)$ is the binary entropy associated to the probability $p$, $h(p) = -p \log p - (1 - p) \log(1 - p)$.

**Problem 13.** Prove the upper bound of the entanglement fidelity when the channel input is perturbed. If $\mathcal{E}$ is a trace-preserving quantum operation, $\rho$ is the density matrix, and $\Delta$ is a zero trace Hermitian operator Then,

$$\mid F_e\left(\rho + \Delta, \mathcal{E}\right)) - F_e(\rho, \mathcal{E}) \mid \leq 2\mathrm{tr}\left[\mid \Delta \mid\right] + \mathrm{tr}\left[\mid \Delta \mid\right]^2.$$

**Problem 14.** Let $\rho$ be the density matrix of system $Q$, the input to a quantum channel and $\mathcal{E}$ the transformation performed by the channel to its input. Show that the following three conditions are equivalent:

(i) $\mathcal{E}$ has a unitary representation;

(ii) $\mathcal{E}$ has a normalized operator-sum representation;

(iii) $\mathcal{E}$ is a trace-preserving, completely positive linear map on the density operators of $Q$.

*Hint:* See [257].

**Problem 15.** Show that the second part of the quantum data processing inequality:

$$I(\rho, \mathcal{E}_1) \geq I\left(\rho, \mathcal{E}_2 \circ \mathcal{E}_1\right)$$

holds even when $\mathcal{E}_1$ is trace-preserving but $\mathcal{E}_2$ is not.

**Problem 16.** Prove properties (iv)- (vii) of quantum fidelity discussed in Section 3.11.

**Problem 17.** Show that $\rho_1$ and $\rho_2$ in Section 3.13 commute if:

$$F\left(\bar{\rho}_0, \bar{\rho}_1\right) = F_{\mathcal{A}}\left(\rho_0, \rho_1\right) = F_{\mathcal{B}}\left(\rho_0, \rho_1\right).$$

**Problem 18.** Show that the intermediate ensemble $\mathcal{A}' = \{\mid \xi_{ja}\rangle, p_a\}$ in Section 3.20 has the same entropy as the original one, $\mathcal{A} = \{\mid \varphi_a\rangle, p_a\}$. *Hint:* see [35].

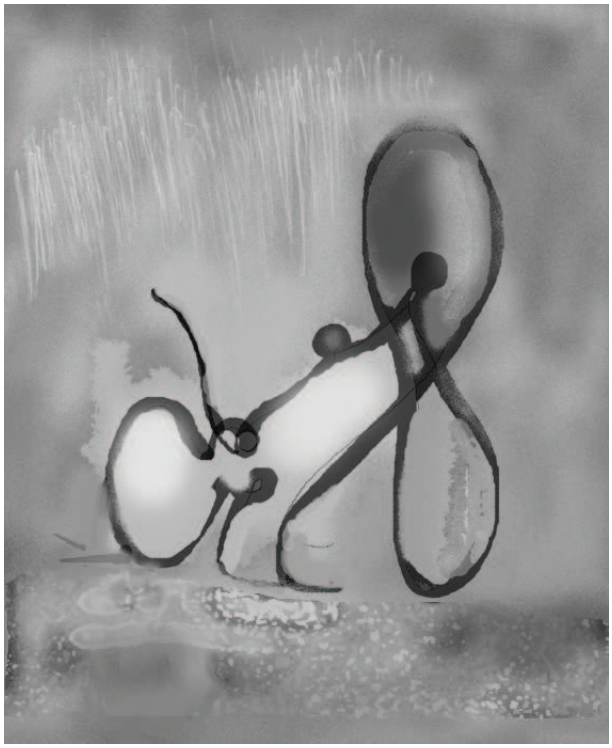
**Problem 19.** Show that the following equality holds

$$\sum_k \mid \mu_k\rangle = 1$$

where $\mid \mu_k\rangle$ are defined in Section 3.17.

**Problem 20.** Prove Schumacher's noiseless coding theorem discussed in Section 3.17. Show that when $N$ is sufficiently large so that $\log N \approx nS(\rho^{\mathcal{A}})$, then the expected probability of error becomes increasingly small, $p_{err} \mapsto 0$.

"Machines should work. People should think." Richard Hamming.



# 4 Classical Error Correcting Codes

Coding theory is an application of information theory critical for reliable communication and fault-tolerant information storage and processing; indeed, Shannon channel coding theorem tells us that we can transmit information on a noisy channel with an arbitrarily low probability of error. A code is designed based on a well-defined set of specifications and *protects the information only for the type and number of errors prescribed in its design.* A good code:

(i) Adds a minimum amount of redundancy to the original message.

(ii) Efficient encoding and decoding schemes for the code exist; this means that information is easily mapped to and extracted from a codeword.

Reliable communication and fault-tolerant computing are intimately related to each other; the concept of a communication channel, an abstraction for a physical system used to transmit information from one place to another and/or from one time to another is at the core of communication, as well as, information storage and processing. It should however be clear

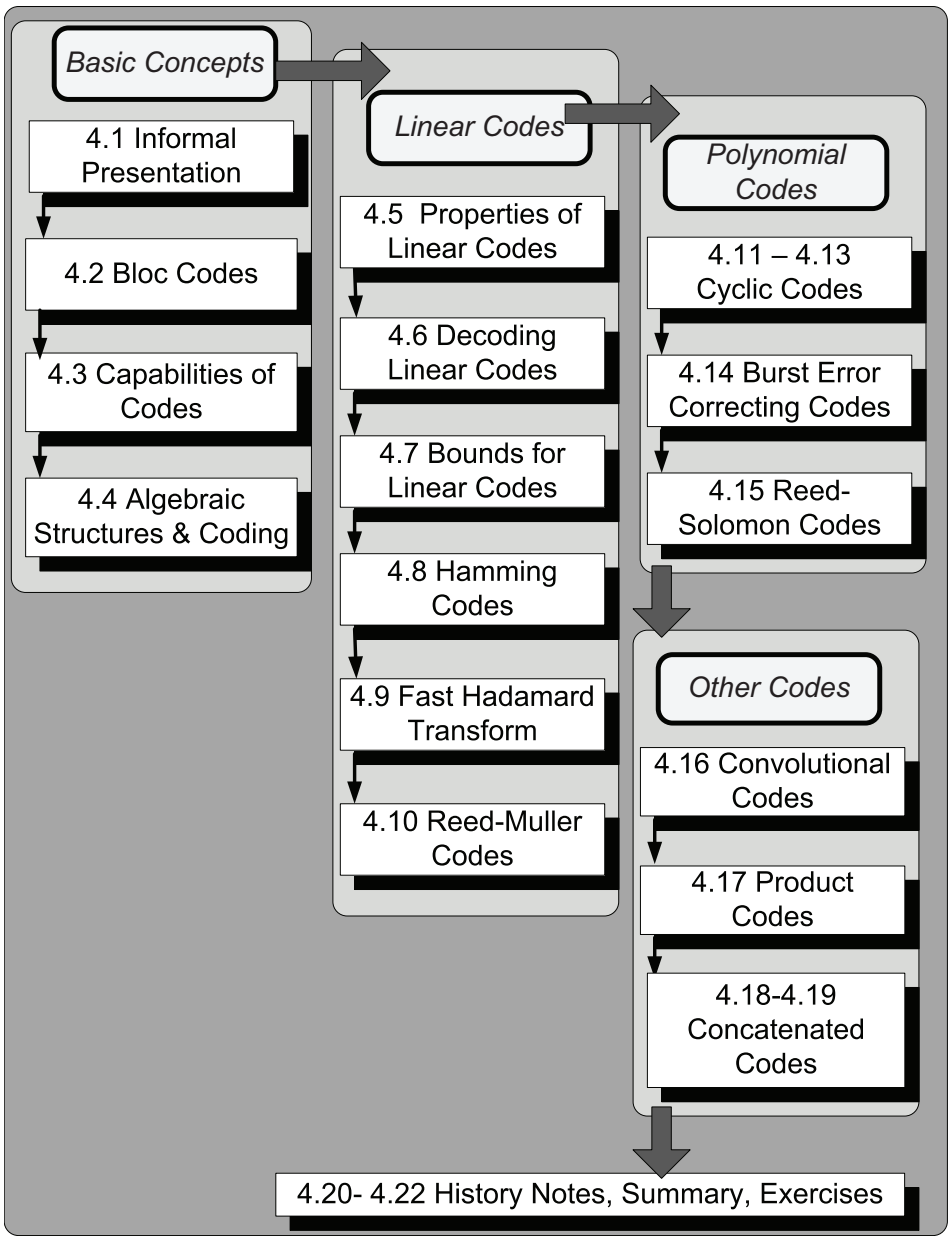



Figure 96: The organization of Chapter 4 at a glance.

that the existence of error-correcting codes does not guarantee that logic operations can be implemented using noisy gates and circuits. The strategies to build reliable computing systems using unreliable components are based on John von Neumann's studies of error detection and error correction techniques for information storage and processing.

This chapter covers topics from the theory of classical error detection and error correction and introduces concepts useful for understanding quantum error correction, the subject of Chapter 5. For simplicity we often restrict our discussion to the binary case. The organization of Chapter 4 is summarized in Figure 96. After introducing the basic concepts we discuss linear, polynomial, and several other classes of codes.

## 4.1 Informal Introduction to Error Detection and Error Correction

*Error detection* is the process of determining if a message is in error; *error correction* is the process of restoring a message in error to its original content. Information is packaged into codewords and the intuition behind error detection and error correction is to artificially increase some "distance" between codewords so that the errors cannot possibly transform one valid codeword into another valid codeword.

Error detection and error correction are based on schemes to increase the redundancy of a message. A crude analogy is to bubble wrap a fragile item and place it into a box to reduce the chance that the item will be damaged during transport. Redundant information plays the role of the packing materials; it increases the amount of data transmitted, but it also increases the chance that we will be able to restore the original contents of a message distorted during communication. Coding corresponds to the selection of both the packing materials and the strategy to optimally pack the fragile item subject to the obvious constraints: use the least amount of packing materials and the least amount of effort to pack and unpack.

**Example.** *A trivial example of an error detection scheme is the addition of a parity check bit to a word of a given length.*

This is a simple scheme but very powerful; it allows us to detect an odd number of errors, but fails if an even number of errors occur. For example, consider a system that enforces even parity for an eight-bit word. Given the string (10111011), we add one more bit to ensure that the total number of 1s is even, in this case a 0, and we transmit the nine-bit string (101110110). The error detection procedure is to count the number of 1s; we decide that the string is in error if this number is odd. When the information symbols are embedded into the codeword, as in this example, the code is called a systematic code; the information symbols are scattered throughout the codeword of a non-systematic code.

This example also hints to the limitations of error detection mechanisms. A code is designed with certain error detection or error correction capabilities and fails to detect, or to correct error patterns not covered by the original design of the code. In the previous example we transmit (101110110) and when two errors occur, in the 4-th and the 7-th bits we receive (101010010). This tuple has even parity (an even number of 1's) and our scheme for error detection fails.

Error detection and error correction have practical applications in our daily life. Error detection is used in cases when an item needs to be uniquely identified by a string of digits and there is the possibility of misidentification. Bar codes, also called Universal Product Codes (UPC), the codes used by the postal service, the American banking system, or by the airlines for paper tickets, all have some error detection capability. For example, in the case of ordinary bank checks and travellers checks there are means to detect if the check numbers are entered correctly during various financial transactions.

**Example.** *The International Standard Book Number (ISBN) coding scheme.* The ISBN code is designed to detect any single digit in error and any transposition of adjacent digits. The ISBN number consists of 10 digits:

$$ISBN \rightarrow d_1 \;\; d_2 \;\; d_3 \;\; d_4 \;\; d_5 \;\; d_6 \;\; d_7 \;\; d_8 \;\; d_9 \;\; d_{10}$$

Here $d_{10}$ is computed as follows:

$$A = w_1 d_1 + w_2 d_2 + w_3 d_3 + w_4 d_4 + w_5 d_5 + w_6 d_6 + w_7 d_7 + w_8 d_8 + w_9 d_9 \mod 11$$

with $w_1 = 10,\ w_2 = 9,\ w_3 = 8,\ w_4 = 7,\ w_5 = 6,\ w_6 = 5,\ w_7 = 4,\ w_8 = 3,\ w_9 = 2$. Then:

$$d_{10} = \begin{cases} 11 - A & \text{if} \quad 2 \leq A \leq 10 \\ X & \text{if} \quad A = 1 \end{cases}$$

For example, consider the book with ISBN $0 - 471 - 43962 - 2$. In this case:

$A = 10 \times 0 + 9 \times 4 + 8 \times 7 + 7 \times 1 + 6 \times 4 + 5 \times 3 + 4 \times 9 + 3 \times 6 + 2 \times 2 \mod 11$

$A = 36 + 56 + 7 + 24 + 15 + 36 + 18 + 4 \mod 11 = 196 \mod 11 = 9.$

$d_{10} = 11 - A = 2.$

The code presented above detects any single errors because the weights are relatively prime with 11. Indeed, if digit $d_i$ is affected by an error and becomes $d_i + e_i$ then

$$\forall e_i \qquad w_i(d_i + e_i) \mod 11 \neq w_i d_i \mod 11$$

because $w_i e_i \neq 0 \mod 11$, $w_i$ is relatively prime to 11. The code also detects any transpositions of adjacent digits. Indeed, the difference of two adjacent weights $w_i - w_{i+1} = 1$ and $d_i - d_{i+1} < 11$, thus it cannot be a multiple of 11. We conclude that the ISBN code detects any single digit in error and any transposition of adjacent digits.

Assume that we wish to transmit through a binary channel a text from a source $A$ with an alphabet with 32 letters. We can encode each one of the $32 = 2^5$ letters of the alphabet as a 5-bit binary string and we can rewrite the entire text in the binary alphabet accepted by the channel. This strategy does not support either detection, or correction of transmission errors because all possible 5-bit combinations are exhausted to encode the 32 letters of the alphabet; any single bit error will transform a valid letter into another valid letter. However, if we use more than five bits to represent each letter of the alphabet, we have a chance to correct a number of single-bit errors.

Adding redundancy to a message increases the amount of information transmitted through a communication channel. The ratio useful information versus total information packed in each codeword defines the rate of a code, see Section 4.2.

**Example.** *We now consider a very simple error correcting code, a "repetitive" code.* Instead of transmitting a single binary symbol, 0 or 1, we map each bit into a string of length $n$ of binary symbols:

$$0 \mapsto (00\ldots0) \qquad \text{and} \qquad 1 \mapsto (11\ldots1).$$

If $n = 2k + 1$ we use a *majority voting scheme* to decode a codeword of $n$ bits. We count the number of 0s; if this number is larger than $k$ we decode the codeword as 0, else we decode it as 1.

Consider a binary symmetric channel with random errors occurring with probability $p$ and call $p^n_{err}$ the probability that an $n$ bit codeword of this repetitive code is in error; then

$$p_{err}^n = 1 - (1-p)^{k+1}\left[(1-p)^k + \binom{2k+1}{1}p(1-p)^{k-1} + \cdots + \binom{2k+1}{k}p^k(1-p)^0\right].$$

Indeed, the probability of error is the probability that $(k+1)$ or more bits are in error which in turn is equal to one minus the probability that at most $k$ bits are in error.

As we can see the larger is $k$, the smaller is the probability of error; this encoding scheme is better than sending a single bit as long as $p_{err}^n < p$. For example, when $n = 3$ and $k = 1$ we encode a single bit as follows: $0 \mapsto (000)$ and $1 \mapsto (111)$. Then the probability of error is

$$p_{err}^3 = 1 - (1-p)^2[(1-p)^1 + \binom{3}{1}p(1-p)^0] = 3p^2 - 2p^3.$$

The encoding scheme increases the reliability of the transmission when $p < 1/2$.

## 4.2 Block Codes. Decoding Policies

In this section we discuss information transmission through a classical channel using codewords of fixed length. The *channel alphabet*, $\mathcal{A}$, is a set of $q$ symbols that could be transmitted through the communication channel; when $q = 2$ we have the *binary alphabet* $\mathcal{A} = \{0, 1\}$. The source encoder maps a message into blocks of $k$ symbols from the code alphabet; then the channel encoder adds $r$ symbols to increase redundancy and transmits <u>$n$-tuples</u> or codewords of length $n = k + r$, as shown in Figure 74.

At the destination, the channel decoder maps $n$-tuples into $k$-tuples, and the decoder reconstructs the message. The quantity $r = n - k > 0$ is called the *redundancy* of the code. The channel encoder adds redundancy, and this leads to the expansion of a message. While the added redundancy is desirable from the point of view of error control, it decreases the efficiency of the communication channel by reducing its effective capacity. The ratio $\frac{k}{n}$ measures the efficiency of a code. First, we give several definitions related to block codes and then discuss decoding policies.

A <u>block code</u> $\mathcal{C}$ of length $n$ over the alphabet $\mathcal{A}$ with $q$ symbols is a set of $M$ $n$-tuples, where each $n$-tuple (codeword) uses symbols from $\mathcal{A}$. Such a block code is also called an $[n, M]$ code over $\mathcal{A}$. The codewords of an $[n, M]$ code $\mathcal{C}$ over $\mathcal{A}$ are a subset, $M$, of all $n$-tuples over $\mathcal{A}$, with $M < q^n$. Given an $[n, M]$ code $\mathcal{C}$ that encodes $k$-tuples into $n$-tuples the <u>rate of the code</u> is $R = k/n$.

We now introduce several metrics necessary to characterize the error detecting and the error correcting properties of a code $\mathcal{C}$ with symbols from the alphabet $\mathcal{A}$.

Given two $n$-tuples $v_i, v_j$, the <u>Hamming distance $d(v_i, v_j)$</u> is the number of coordinate positions in which the two $n$-tuples differ. The Hamming distance is a metric; indeed $\forall v_i, v_j, v_k \in \mathcal{C}$ we have:

1. $d(v_i, v_j) \geq 0$, with equality if and only if $v_i = v_j$.

2. $d(v_i, v_j) = d(v_j, v_i)$.

3. $d(v_i, v_j) + d(v_j, v_k) \geq d(v_i, v_k)$ (triangle inequality).

The proof of this proposition is left as an exercise.

The Hamming distance of a code is the minimum distance between any pairs of codewords

$$d = \min\{d(c_i, c_j) : \quad \forall c_i, c_j \in \mathcal{C}, c_i \neq c_j\}.$$

To compute the Hamming distance for an $[n, M]$ code $\mathcal{C}$, it is necessary to compute the distance between $\binom{M}{2}$ pairs of codewords and then to find the pair with the minimum distance. A block code with the block length $n$, with $k$ information symbols and with distance $d$ will be denoted as $[n, k, d]$; if the alphabet $\mathcal{A}$ has $q$ symbols then $M = q^k$.

Given a codeword $c \in \mathcal{C}$ the Hamming weight, $w(c)$, is the number of non-zero coordinates of the codeword $c$. Let $c_i$, $1 \leq i \leq M$, be the codewords of $\mathcal{C}$, an $[n, M]$ code, and let $S$ be the set of all $n$-tuples over the alphabet of $\mathcal{C}$; the Hamming sphere of radius d around the codeword $c_i$ consists of all $n$-tuples $c$ within distance $d$ of the codeword $c_i$, the *center* of the Hamming sphere

$$S_{c_i} = \{c \in S : \quad d(c, c_i) \leq d\}$$

Figure 97 shows the two Hamming spheres of radius 1 around the 3-tuples 000 and 111.

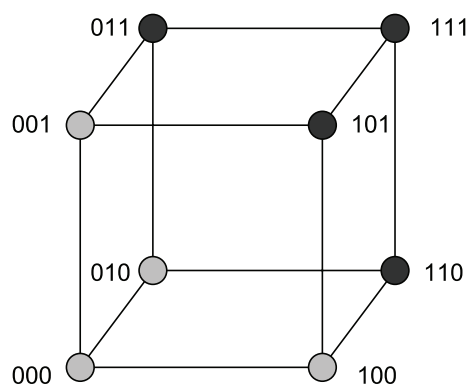



Figure 97: The Hamming sphere of radius 1 around 000 contains $\{000, 100, 010, 001\}$ and the Hamming sphere of radius 1 around 111 contains $\{111, 011, 101, 110\}$.

**Example.** *The Hamming distance of two codewords.* Consider the binary alphabet $\{0, 1\}$ and let the two codewords be $v_i = (010110)$ and $v_j = (011011)$. The Hamming distance between the two codewords is $d(v_i, v_j) = 3$. Indeed, if we number the bit position in each $n$-tuple from left to right as 1 to 6, the two $n$-tuples differ in bit positions 3, 4, and 6.

**Example.** *The Hamming distance of a code.* Consider the code $\mathcal{C} = \{c_0, c_1, c_2, c_3\}$ where $c_0 = (000000)$, $c_1 = (101101)$, $c_2 = (010110)$, $c_3 = (111011)$. This code has distance $d = 3$. Indeed $d(c_0, c_1) = 4$, $d(c_0, c_2) = 3$, $d(c_0, c_3) = 5$, $d(c_1, c_2) = 5$, $d(c_1, c_3) = 3$, $d(c_2, c_3) = 4$.

**Example.** *The Hamming sphere of radius* $d = 1$ *around the codeword* (000000) consists of the center, (0000000), and all binary six-tuples, that differ from (000000) in exactly one bit position, $\{(000000), (100000), (010000), (001000), (000100), (000010), (000001)\}$. In this case one can construct a diagram similar with the one in Figure 97 where the 6-tuples are the vertices of a hypercube in a space with 6 dimensions.

**Decoding policy.** Given an $[n, M]$ code $\mathcal{C}$ with distance $d$, the decoding policy tells the channel decoder what actions to take when it receives an $n$-tuple $v$. This policy consists of two phases:

(1) The recognition phase when the received $n$-tuple $v$ is compared with all the codewords of the code until either a match is found, or we decide that $v$ is not a codeword.

(2) The error correction phase, when the received $n$-tuple $v$ is mapped into a codeword.

The actions taken by the channel decoder are:

(i) If $v$ is a codeword, conclude that no errors have occurred and decide that the codeword sent was $c = v$.

(ii) If $v$ is not a codeword, conclude that errors have occurred and either correct $v$ to a codeword $c$, or declare that correction is not possible.

We observe that any decoding policy fails when $c \in \mathcal{C}$, the codeword sent, is affected by errors and transformed into another valid codeword, $c' \in \mathcal{C}$. This is a fundamental problem with error detection and we shall return to it later. Once we accept the possibility that the channel decoder may fail to properly decode in some cases, our goal is to take the course of action with the highest probability of being correct. Now we discuss three decoding policies.

I. *Minimum distance or nearest neighbor decoding.* If an $n$-tuple $v$ is received, and there is a unique codeword $c \in \mathcal{C}$ such that $d(v, c)$ is the minimum over all $c \in \mathcal{C}$, then correct $v$ as the codeword $c$. If no such $c$ exists, report that errors have been detected, but no correction is possible. If multiple codewords are at the same minimum distance from the received codeword select at random one of them and decode $v$ as that codeword.

II. *Bounded distance decoding.* This decoding policy requires that all patterns of at most $e$ errors and no other errors are corrected. When bounded distance decoding is used and more than $e$ errors occur, an $n$-tuple $v$ is either not decoded, or decoded to a wrong codeword; the first case will be called a <u>decoding failure</u> and the second one a <u>decoding error</u>. It is easy to see that if

$$e = \lfloor \frac{d-1}{2} \rfloor$$

the bounded distance decoding is identical to minimum distance decoding.

Recall from Section 3.7 that in case of a binary symmetric channel the errors occur with probability $p$ and are mutually independent. Given an $[n, k]$ linear code operating on a binary symmetric channel the probabilities of decoding failure, $P_{fail}$, and of decoding error, $P_{err}$, when at most $e$ errors occur are

$$P_{fail} = 1 - \sum_{i=0}^{e} \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=e+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

and

$$P_{err} = \sum_{w>0} \sum_{i=w-e}^{w+e} \sum_{j=0}^{e} A_w T(i,j,w) p^i (1-p)^{n-i}.$$

In this expression $A_w$ is the number of codewords of code $\mathcal{C}$ of weight equal to $w$ and $T(i,j,w)$ is the number of $n$-tuples at distance $i$ from a codeword with weight $w$ and at distance $j$ from the decoded codeword; $e$ is the number of errors.

It is easy to prove that given a binary $[n,k]$ linear code $\mathcal{C}$, $T(i,j,w)$ is given by the following expression:

$$T(i,j,w) = \binom{w}{i-k} \binom{n-w}{k}$$

with $k = 1/2(i+j-w)$ when $(i+k-w)$ is even and when $w-j \leq i \leq w+j$; otherwise $T = 0$.

III. *Maximum likelihood decoding.* Under this decoding policy, of all possible codewords $c \in \mathcal{C}$ the $n$-tuple $v$ is decoded to that codeword $c$ which maximizes the probability $P(v,c)$ that $v$ is received, given that $c$ is sent.

**Proposition.** *If $v$ is received at the output of a $q$-ary symmetric channel and if $d_1 \leq d_2$ with $d_1 = d(v,c_1)$ and $d_2 = d(v,c_2)$, then*

$$P(v,c_1) \geq P(v,c_2) \quad \text{if and only if} \quad p < \frac{(q-1)}{q}.$$

In other words, if $p < (q-1)/q$ the maximum likelihood decoding policy is equivalent to the nearest neighbor decoding; a received vector is decoded to the codeword "closest" to it, in terms of Hamming distance.

Without loss of generality we assume that $d_1 \leq d_2$ and that

$$P(v,c_1) > P(v,c_2).$$

It follows that

$$(1-p)^{n-d_1} \left(\frac{p}{q-1}\right)^{d_1} > (1-p)^{n-d_2} \left(\frac{p}{q-1}\right)^{d_2}$$

$$(1-p)^{d_2-d_1} > \left(\frac{p}{q-1}\right)^{d_2-d_1}$$

and

$$\left(\frac{p}{(1-p)(q-1)}\right)^{d_2-d_1} < 1.$$

If $d_1 = d_2$, this is false, and in fact, $P(v, c_1) = P(v, c_2)$. Otherwise, $d_2 - d_1 \geq 1$ and the inequality is true if and only if

$$\frac{p}{(1-p)(q-1)} < 1, \quad \text{i.e.,} \quad p < \frac{(q-1)}{q}.$$

In conclusion, when the probability of error is $p < (q-1)/q$ and we receive the $n$-tuple $v$, we decide that codeword $c$ at the minimum distance from $v$ is the one sent by the source.

**Example.** *Maximum likelihood decoding.* Assume that the probability of error is $p = 0.15$. Let the binary code $\mathcal{C}$ be $\mathcal{C} = \{(000000), (101100), (010111), (111011)\}$. When we receive the 6-tuple $v = (111111)$ we decode it as $(111011)$ because the probability $P((v), (111011))$ is the largest. Indeed,

$$\begin{aligned}
P(v, (000000)) &= (0.15)^6 = 0.000011 \\
P(v, (101100)) &= (0.15)^3 \times (0.85)^3 = 0.002076 \\
P(v, (010011)) &= (0.15)^3 \times (0.85)^3 = 0.002076 \\
P(v, (111011)) &= (0.15)^1 \times (0.85)^5 = \underline{0.066555}
\end{aligned}$$

A code is capable of *correcting $e$ errors* if the channel decoder is capable of correcting any pattern of $e$ or fewer errors, using the algorithm above.

## 4.3 Error Correcting and Detecting Capabilities of a Block Code

The intuition behind error detection is to use a small subset of the set of $n$-tuples over an alphabet $\mathcal{A}$ as codewords of a code $\mathcal{C}$ and detect an error whenever the $n$-tuple received is not a codeword. To detect errors we choose the codewords of $\mathcal{C}$ so that the Hamming distance between any pair of codewords is large enough and one valid codeword cannot be transformed by errors into another valid codeword. The Hamming distance of a code allows us to accurately express the capabilities of a code to detect and to correct errors. Several propositions quantify the error detection and error correction properties of a code.

**Proposition.** *Let $\mathcal{C}$ be an $[n, k, d]$ code with an odd distance, $d = 2e + 1$. Then $\mathcal{C}$ can correct $e$ errors and can detect $2e$ errors.*

*Proof.* We first show that $S_{c_i}^{(e)}$ and $S_{c_j}^{(e)}$, the Hamming spheres of radius $e$ around two distinct codewords $c_i \neq c_j \in \mathcal{C}$ are disjoint, $S_{c_i}^{(e)} \cap S_{c_j}^{(e)} = \emptyset$. Assume by contradiction that there exists a codeword $v \in \mathcal{C}$ with the property that $v \in S_{c_i}^{(e)} \cap S_{c_j}^{(e)}$ and that $d(v, c_i) \leq e$ and $d(v, c_j) \leq e$. The triangle inequality requires that

$$d(c_i, v) + d(v, c_j) \geq d(c_i, c_j) \quad \Longrightarrow \quad 2e \geq d(c_i, c_j).$$

But the distance of the code is $d = 2e + 1$ so $d(c_i, c_j) \geq 2e + 1$. This is a contradiction and we conclude that $S_{c_i}^{(e)} \cap S_{c_j}^{(e)} = \emptyset$.

Therefore, when the codeword $c_i$ is transmitted and $t \leq e$ errors are introduced, the received $n$-tuple, $v$ is inside the Hamming sphere $S_{c_i}^{(e)}$ and $c_i$ is the unique codeword closest to $v$. The decoder can always correct any error pattern of this type.

If we use the code $\mathcal{C}$ for error detection, then at least $2e+1$ errors must occur in any codeword $c \in \mathcal{C}$ to transform it into another codeword $c' \in \mathcal{C}$. If at least 1 and at most $2e$ errors are introduced, the received $n$-tuple will never be a codeword and error detection is always possible. $\square$

The distance of a code may be even or odd. The case for even distance is proved in a similar manner. Throughout this chapter we use the traditional notation $\lfloor a \rfloor$ for the largest integer smaller than or equal to $a$.

**Proposition.** *Let $\mathcal{C}$ be an $[n,k,d]$ code. Then $\mathcal{C}$ can correct $\lfloor \frac{(d-1)}{2} \rfloor$ errors and can detect $d-1$ errors.*

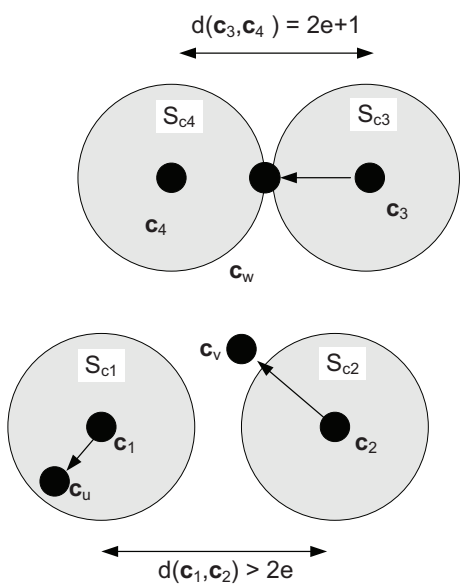



Figure 98: A code $\mathcal{C} = \{c_1, c_2, c_3, c_4, ..\}$ with minimum distance $d = 2e+1$ is used for error detection as well as for error correction. The Hamming spheres $S_{c_1}, S_{c_2}, S_{c_3}$, and $S_{c_4}$ are centered around the codewords $c_1, c_2, c_3$, and $c_4$ respectively. The decoder identifies $c_u, c_v$, and $c_w$ as errors; one of them is corrected, the other two are not. $c_u$, received when the $c_1$ was sent, is correctly decoded as $c_1$ because it is located inside the Hamming sphere around codeword $c_1$, $d(c_1, c_u) < e$. The decoder fails to decode $c_v$, received when $c_2$ was sent, because $d(c_2, c_v) > e$ and $c_v$ is not inside any of the Hamming spheres. The decoder incorrectly decodes $c_w$, received when $c_3$ was sent, as $c_4$ because $d(c_3, c_w) = e+1$ and $d(c_4, c_w) \leq e$.

A geometric interpretation of this proposition is illustrated in Figure 98. The $[n,k,d]$ code $\mathcal{C} = \{c_1, c_2, c_3, c_4, ..\}$ with distance $d = 2e+1$ is used for error correction as well as for error detection. The Hamming spheres of radius $e$ around all codewords do not intersect because the minimum distance between any pair of codewords is $2e+1$. According to the channel

decoding rules, any $n$-tuple $t$ in the Hamming sphere of radius $e$ around codeword $c_1$ is decoded as the center of the Hamming sphere $c_1$. If at most $2e$ errors occur when the codeword $c_i$ is transmitted, the received $n$-tuple, $v$ cannot be masquerading as a valid codeword $c_j$ since the distance between $c_i$ and $c_j$ is at least $2e+1$.

From Figure 98 we see that there are $n$-tuples that are not contained in any of the Hamming spheres around the codewords of a code $\mathcal{C}$. If an $n$-tuple $v$ is received and the decoder cannot place it in any of the Hamming spheres, then the decoder knows that at least $e+1$ errors have occurred and no correction is possible.

If a code $\mathcal{C}$ with distance $2e+1$ is used *for error detection as well as error correction* some patterns of less than $2e$ errors could escape detection. For example, patterns of $e+1$ errors that transform a codeword $c_3$ into an $n$-tuple, $c_w$, inside the Hamming sphere $S_{c_4}$, force the decoder to correct the received $n$-tuple to $c_4$. In this case the $(e+1)$-error pattern causes a faulty decoding.

In conclusion, when $d = 2e+1$ the code $\mathcal{C}$ can correct $e$ errors in general, but is unable to detect additional errors at the same time. If the distance of the code is an even number, the situation changes slightly according to the following proposition.

**Proposition.** *Let $\mathcal{C}$ be an $[n,k,d]$ code with $d = 2e$. Then $\mathcal{C}$ can correct $e-1$ errors and, simultaneously, detect $e$ errors.*

*Proof.* According to a previous proposition the code $\mathcal{C}$ can correct up to

$$\lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{2e-1}{2} \rfloor = \lfloor e - \frac{1}{2} \rfloor = e-1$$

errors. Since the Hamming spheres around codewords have radius $e-1$, any pattern of $e$ errors cannot take a codeword into a an $n$-tuple contained in some Hamming sphere around another codeword. Otherwise, the codewords at the centers of these two Hamming spheres would have distance at most $e+e-1 = 2e-1$, which is impossible since $d = 2e$. Hence, a received word obtained from a codeword by introducing $k$ errors cannot lie in any codeword Hamming sphere and the decoder can detect the occurrence of errors. The decoder cannot detect $e+1$ errors in general, since a vector at distance $e+1$ from a given codeword may be in the Hamming sphere of another codeword, and the decoder would erroneously correct such a vector to the codeword at the center of the second Hamming sphere. □

To evaluate the reliability of the nearest neighbor decoding strategy we have to compute the probability that a codeword $c$ sent over the channel is decoded correctly. Let us consider an $[n,k,d]$ code $\mathcal{C}$ over an alphabet $\mathcal{A}$ with $q$ symbols, $|A| = q$. Assume a probability of error $p$, such that $0 \leq p \leq (q-1)/q$.

When codeword $c$ is sent, the received binary $n$-tuple, $v$, will be decoded correctly if it is in the Hamming sphere $S_c$ of radius $e = \lfloor (d-1)/2 \rfloor$ about $c$. The probability of this event is

$$\sum_{v \in S_c} P(v,c) = \sum_{i=0}^{e} \binom{n}{i} p^i (1-p)^{n-i}.$$

Indeed, the probability of receiving an $n$-tuple with $i$ positions in error is $\binom{n}{i} p^i(1-p)^{n-i}$ and $i$ takes values in the range from 0, no errors, to a maximum of $e$ errors. This expression gives a lower bound on the probability that a transmitted codeword is correctly decoded.

**Example.** *Decoding rule.* Consider the code $\mathcal{C} = \{c_1, c_2, c_3\}$ where

$c_1 = (101100) \quad c_2 = (010111) \quad c_3 = (111011).$

The code $\mathcal{C}$ has distance $d = 3$, and hence can correct 1 error. The set $S$ of all possible words over the alphabet {0,1} consists of all possible binary 6-tuples, hence, $|S| = 64$. Let us now construct the four Hamming spheres of radius 1 about each codeword:

$S_{c_1} = \{(101100), (001100), (111100), (100100), (101000), (101110), (101101)\},$
$S_{c_2} = \{(010111), (110111), (000111), (011111), (010011), (010101), (010110)\},$
$S_{c_3} = \{(111011), (011011), (101011), (110011), (111111), (111001), (111010)\}.$

The Hamming spheres cover 21 of the 64 possible 6-tuples in $S$. Let $S^*$ be the set of 6-tuples that are not included in any Hamming sphere; then $|S^*| = 64 - 21 = 43$. When the decoder receives the binary 6-tuple $v = (000111)$ the distance to each codeword is computed as follows:

$$d(c_1, v) = 4, \quad d(c_2, v) = 1, \quad d(c_3, v) = 4.$$

According to the minimum distance rule $v$ is decoded as $c_2$; $v$ lies in the Hamming sphere $S_{c_2}$.

## 4.4 Algebraic Structures and Coding Theory

A brief excursion into algebraic structures used by coding theory will help us follow the presentation of the codes discussed in this chapter, as well as the quantum error correcting codes presented in Chapter 5. First, we review basic concepts regarding the algebraic structures used in this chapter and in the next one.

A group consists of a set $G$ equipped with a map $\top : (G \times G) \mapsto G$ and $\top^{-1} : G \mapsto G$ and an identity element $e \in G$ so that the following axioms hold:

1. $\forall (g_i, g_j) \in G \quad (g_i \top g_j) \in G.$
2. $\forall (g_i, g_j, g_k) \in G \quad (g_i \top g_j) \top g_k = g_i \top (g_j \top g_k).$
3. $\forall g_i \in G \quad g_i \top e = g_i.$
4. $\forall g_i \in G \quad \exists g_i^{-1}$ such that $g_i \top g_i^{-1} = e.$

$G$ is an Abelian or commutative group if an additional axiom is satisfied: $\forall (g_i, g_j) \in F \quad g_i \top g_j = g_j \top g_i$. The group is either an additive group when the operation ($\top$) is addition denoted as (+), or a multiplicative group when the operation is multiplication denoted as ($\cdot$). A subset $G'$ is called a subgroup of $G$ if $g_1, g_2 \in G'$ implies $g_1 \top g_2^{-1} \in G'$. The number of distinct elements of a group is called the order of the group.

**Example.** *The set of rotational symmetries of an ordinary sphere.* This symmetry group has an infinite number of elements since we can rotate a sphere through any angle about any direction in the 3-dimensional space. The group is called $SO(3)$ [324].

A subgroup $G' \subset G$ is a normal subgroup of $G$ if each element of the group $G$ commutes with the normal subgroup $G'$:

$$G' \top g = g \top G', \quad \forall g \in G.$$

A set $R$ with two binary operations, addition, $(+)$, and multiplication, $(\cdot)$, and a distinct element, $0 \in G$, is a commutative ring with unity, or simply a ring, if the following axioms are satisfied:

1. (R,+) is an Abelian group with the identity element 0.
2. Multiplication is commutative: $\forall (a, b) \in R \quad a \cdot b = b \cdot a$.
3. Identity/unity element for multiplication: $\exists\ 1 \in R$ such that $\forall a \in R \ \ 1 \cdot a = a \cdot 1 = a$.
4. Multiplication is associative: $\forall (a, b, c) \in R \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$.
5. Multiplication is distributive: $\forall (a, b, c) \in R \quad (a + b) \cdot c = a \cdot c + b \cdot c$.

Note that there is no multiplicative inverse of the elements of a ring.

**Example.** *The set of integers with integer addition and integer multiplication,* $(Z, +, \cdot)$. This set forms a ring, denoted as $\mathbb{Z}$, with an infinite number of elements.

An algebraic structure $(F, +, \cdot)$, consisting of a set $F$, two binary operation, addition $(+)$ and multiplication $(\cdot)$, and two distinct elements, 0 and 1 is a field if the following axioms are satisfied:

1. $\forall (a, b) \in F \quad a + b = b + a$.
2. $\forall (a, b, c) \in F \quad (a + b) + c = a + (b + c)$.
3. $\forall a \in F \quad a + 0 = a$.
4. $\forall a \in F \quad \exists (-a) \quad a + (-a) = 0$.
5. $\forall (a, b) \in F \quad a \cdot b = b \cdot a$.
6. $\forall (a, b, c) \in F \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$.
7. $\forall a \in F \quad a \cdot 1 = a$.
8. $\forall a \in F \quad \exists a^{-1} \quad a \cdot a^{-1} = 1$.
9. $\forall (a, b, c) \in F \quad a \cdot (b + c) = a \cdot b + a \cdot c$.

A field $F$ is at the same time an Abelian additive group, $(F, +)$, and an Abelian multiplicative group, $(F, \cdot)$.

A field $F$ with a finite number of elements $q$ is called a <u>finite or Galois field</u> denoted as $GF(q)$, it is also named after the French mathematician Évariste Galois.

We now discuss the congruence relationship and show that it allows us to construct finite algebraic structures consisting of equivalence classes. Given a positive integer, $p \in Z$, the <u>congruence</u> relation among integers is defined as

$$\forall (m, n) \in Z \quad m \equiv n \mod p \quad \Longrightarrow (m - \lfloor m/p \rfloor \cdot p) = (n - \lfloor n/p \rfloor \cdot p).$$

The congruence is an <u>equivalence relation</u> $\mathcal{R}$: it is reflexive, $a\mathcal{R}a$; symmetric, $a\mathcal{R}b \Rightarrow b\mathcal{R}a$; and transitive $[a\mathcal{R}b \text{ and } b\mathcal{R}c] \Rightarrow a\mathcal{R}c$. It partitions a set into equivalence classes. For example, the set of integers modulo 3 is partitioned into three equivalence classes:
$[0] = 0, 3, 6, 9 \ldots$, the integers divisible by 3;
$[1] = 1, 4, 7, 10 \ldots$, the integers with a remainder of 1 when divided by 3; and
$[2] = 2, 5, 8, 11 \ldots$, the integers with a remainder of 2 when divided by 3.

**Proposition.** *The set of integers modulo $p$, $\mathbb{Z}_p$, with the standard integer addition and integer multiplication is a finite field if and only if $p$ is a prime number.*

*Proof:* To prove that the condition is necessary let us assume that $p$ is prime. First, we observe that $\mathbb{Z}_p$ has a finite number of elements, $0, 1, 2, \ldots (p-1)$. Then we verify all the properties or axioms of a field. Here we only show that given $0 < a < p$ the multiplicative inverse, $a^{-1}$ exists. The $\gcd(p, a) = 1$ because $p$ is prime; then according to Euclid's algorithm there exist integers $s$ and $t$ such that $s \cdot p + t \cdot a = 1$. But for all integers $s$ we have $s \cdot p \equiv 0 \mod p$; this implies that $t \cdot a = 1 \mod p$ or $t = a^{-1} \mod p$. Thus, if $p$ is prime, then $\mathbb{Z}_p$ is a finite field.

We start with the observation that if $p$ is not a prime, then it is a product of two terms $p = a \cdot b$ with $a, b \neq 0 \mod p$. To prove that the condition is sufficient we have to show that $b$ has no multiplicative inverse in $\mathbb{Z}_p$. In other words that there is no $c$ such that $b = c^{-1}$ or $b \cdot c \equiv 1 \mod p$. If such $c$ exists then

$$p \cdot c = a \cdot b \cdot c = a \cdot (b \cdot c) = a \cdot 1 \equiv a \mod p = 0 \mod p.$$

But, $a \equiv 0 \mod p$ contradicts the assumption that $a \neq 0 \mod p$. □

**Example.** *The finite field of integers modulo 5 denoted as $\mathbb{Z}_5$ or $GF(5)$:* $\mathbb{Z}_5$ consists of five equivalence classes of integers modulo 5, namely [0], [1], [2], [3], and [4] corresponding to the remainders $0, 1, 2, 3$ and 4 respectively:

$$\begin{array}{lllllllll}
[0] = \{ & 0 & 5 & 10 & 15 & 20 & 25 & 30 & \ldots\} \\
[1] = \{ & 1 & 6 & 11 & 16 & 21 & 26 & 31 & \ldots\} \\
[2] = \{ & 2 & 7 & 12 & 17 & 22 & 27 & 32 & \ldots\} \\
[3] = \{ & 3 & 8 & 13 & 18 & 23 & 28 & 33 & \ldots\} \\
[4] = \{ & 4 & 9 & 14 & 19 & 24 & 29 & 34 & \ldots\}
\end{array}$$

The addition and multiplication tables of $\mathbb{Z}_5$ are

| $+$ | [0] | [1] | [2] | [3] | [4] |
|---|---|---|---|---|---|
| [0] | [0] | [1] | [2] | [3] | [4] |
| [1] | [1] | [2] | [3] | [4] | [0] |
| [2] | [2] | [3] | [4] | [0] | [1] |
| [3] | [3] | [4] | [0] | [1] | [2] |
| [4] | [4] | [0] | [1] | [2] | [3] |

and

| $\cdot$ | [0] | [1] | [2] | [3] | [4] |
|---|---|---|---|---|---|
| [0] | [0] | [0] | [0] | [0] | [0] |
| [1] | [0] | [1] | [2] | [3] | [4] |
| [2] | [0] | [2] | [4] | [1] | [3] |
| [3] | [0] | [3] | [1] | [4] | [2] |
| [4] | [0] | [4] | [3] | [2] | [1] |

.

From these tables we see that [3] is the multiplicative inverse of [2] (indeed, $[3] \cdot [2] = [1]$) and that [1] and [4] are their own multiplicative inverses ($[1] \cdot [1] = [1]$ and $[4] \cdot [4] = [1]$).

$\mathbb{Z}_5, \mathbb{Z}_7, \mathbb{Z}_{11}$, and $\mathbb{Z}_{13}$ are finite fields but $\mathbb{Z}_9$ and $\mathbb{Z}_{15}$ are not because 9 and 15 are not prime numbers. Indeed, the set $\mathbb{Z}_9$ is finite and consists of the following elements: $[0], [1], [2], [3], [4], [5], [6], [7]$ and $[8]$. The multiplication table of $\mathbb{Z}_9$ is

| $\cdot$ | [0] | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] |
|---|---|---|---|---|---|---|---|---|---|
| [0] | [0] | [0] | [0] | [0] | [0] | [0] | [0] | [0] | [0] |
| [1] | [0] | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] |
| [2] | [0] | [2] | [4] | [6] | [8] | [1] | [3] | [5] | [7] |
| [3] | [0] | [3] | [6] | [0] | [3] | [6] | [0] | [3] | [6] |
| [4] | [0] | [4] | [8] | [3] | [7] | [2] | [6] | [1] | [5] |
| [5] | [0] | [5] | [1] | [6] | [2] | [7] | [3] | [8] | [4] |
| [6] | [0] | [6] | [3] | [0] | [6] | [3] | [0] | [6] | [3] |
| [7] | [0] | [7] | [5] | [3] | [1] | [8] | [6] | [4] | [2] |
| [8] | [0] | [8] | [7] | [6] | [5] | [4] | [3] | [2] | [1] |

.

We see immediately that $\mathbb{Z}_9$ is not a field because some non-zero elements, e.g., [3] and [6] do not have a multiplicative inverse. The multiplicative inverse of [2] is [5] (indeed, $[2] \cdot [5] = [1]$), the one of [7] is [4] ( indeed, $[7] \cdot [4] = [1]$), and the one of [1] is [1].

It is easy to prove that a finite field $GF(q)$ with $q = p^n$ and $p$ a prime number has $p^n$ elements. The finite field $GF(q)$ can be considered a vector space $V$ over $\mathbb{Z}_p$. $GF(q)$ is a finite field thus, the vector space $V$ is finite dimensional. If $n$ represents the number of dimensions of $V$ there are $n$ elements $v_1, v_2, \ldots v_n \in V$ which form a basis for $V$:

$$V = \sum_{i=1}^{n} \lambda_i v_i \qquad \lambda_i \in Z_p.$$

**Example.** *Consider $p = 2$, the finite field of binary numbers, $\mathbb{Z}_2$.* For $n \geq 2$ we construct the set of all $2^n$ polynomials with binary coefficients:

$$\mathbb{Z}_2[x] = \{0, 1, x, 1+x, 1+x^2, x^2, x+x^2, 1+x+x^2, x^3, 1+x^3, 1+x+x^3, 1+x^2+x^3, 1+x+x^2+x^3, x+x^3, x^2+x^3, \ldots\}.$$

Given a finite field with $q$ elements, there is an element $\beta \in GF(q)$ called <u>primitive element</u>, or characteristic element of the finite field such that the set of nonzero elements of $GF(q)$ is

$$GF(q) - \{0\} = \{\beta^1, \beta^2, \ldots, \beta^{q-2}, \beta^{q-1}\} \quad \text{with} \quad \beta^{q-1} = 1.$$

Multiplication in a finite field with $q$ elements, $GF(q)$, can be carried out as addition modulo $(q-1)$. Indeed, if $\alpha$ is the primitive element of $GF(q)$ then any two elements can be expressed as powers of $\beta$, $a = \beta^i$ and $b = \beta^j$. Then

$$a \cdot b = \beta^i \cdot \beta^j = \beta^{(i+j) \mod (q-1)}.$$

**Example.** *Table 11 displays the 4 non-zero elements of $GF(5)$ as powers of the primitive element $\beta = [3]$.*

Table 11: The non-zero elements of $GF(5)$ expressed as a power of $\beta = [3]$.

| Element of $GF(5)$ | Element of $GF(5)$ as a power of $\beta$ | Justification |
|---|---|---|
| $[3]$ | $\beta^1$ | $3 \mod 5 = 3$ |
| $[4]$ | $\beta^2$ | $3^2 \mod 5 = 9 \mod 5 = 4$ |
| $[2]$ | $\beta^3$ | $3^3 \mod 5 = 27 \mod 5 = 2$ |
| $[1]$ | $\beta^4$ | $3^4 \mod 5 = 81 \mod 5 = 1$ |

The product of two elements of $GF(5)$ can be expressed as powers of the primitive element. For example, $[2] \cdot [4] = [3]^3 \cdot [3]^2 = [3]^{5 \mod 4} = [3]$; indeed, $(2 \cdot 4 \mod 5) = (8 \mod 5) = 3$.

Let $GF(q)$ be a finite field and $a \in GF(q)$; the <u>order of element</u> $a$, $\text{ord}(a)$ is the smallest integer $s$ such that $a^s = 1$. The set $\{\beta^1, \beta^2, \ldots\}$ is finite thus there are two positive integers $i$ and $j$ such that $\beta^i = \beta^j$. This implies that $\beta^{i-j} = 1$; in turn, this means that there is a smallest $s$ such that $a^s = 1$, so the order is well defined.

By analogy with the ring of integers modulo a positive integer $p$, we can use the congruence relation to construct equivalence classes of polynomials. The set of polynomials $q(x)$ in $x$ and with coefficients from a field $F$ form a ring with an infinite number of elements denoted as $F[x]$. If $f(x) \in F[x]$ is a non-zero polynomial, we define the equivalence class $[g(x)]$, containing the polynomial $g(x) \in F[x]$ as the set of polynomials which produce the same remainder as $g(x)$ when divided by $f(x)$.

Let $\mathbb{F}[x]$ denote the set of polynomials in variable $x$ with coefficients from a finite field $F$ and let $f(x), g(x) \in \mathbb{F}[x]$ be non-zero polynomials; the <u>equivalence class of polynomials</u> $[g(x)]$ relative to $f(x)$ is defined as

$$[g(x)] = \{q(x) : \ q(x) \in\equiv g(x) \mod f(x)\}.$$

Addition and multiplication of equivalence classes of polynomials are defined as

$$[g(x)] + [h(x)] = [g(x) + h(x)] \qquad \text{and} \qquad [g(x)] \cdot [h(x)] = [g(x) \cdot h(x)].$$

We can use the congruence relation to construct a finite ring of polynomials over $F$ modulo the non-zero polynomial $f(x)$. A <u>finite ring of polynomials</u> over $F$ modulo the non-zero polynomial $f(x)$, denoted by $R = \mathbb{F}[x]/(f(x))$ is

$$R = \mathbb{F}[x]/(f(x)) = \{[g(x)] : \ g(x) \in F[x]\}.$$

We are particulary interested in the special case $f(x) = x^n - 1$ when:

$$x^n - 1 \equiv 0 \mod f(x) \implies x^n \equiv 1 \mod f(x).$$

**Example.** *Addition and multiplications in* $\mathbb{Z}_2[x]/(x^4+1)$

$$[1 + x^2 + x^3] + [1 + x + x^2 + x^3] = [1 + x^2 + x^3 + 1 + x + x^2 + x^3] = [x]$$

and

$$[1 + x] \cdot [1 + x + x^2 + x^3] = [1 + x + x^2 + x^3 + x + x^2 + x^3 + x^4] = [1 + x^4] = [0].$$

An <u>ideal $I$ of a ring</u> $(R, +, \cdot)$ is a non-empty subset of $R$, $I \subset R, I \neq \emptyset$, with two properties:

1. $(I, +)$ is a group.
2. $i \cdot r \in I \quad \forall i \in I$ and $\forall r \in R$.

The <u>ideal generated by the polynomial $g(x) \in R$</u> is the set of all polynomials multiple of $g(x)$

$$I = \{g(x) \cdot p(x), \quad \forall p(x) \in R\}.$$

A <u>principal ideal ring</u> $(R, +, \cdot)$ is one when $\forall I \subset R$ there exits $g(x) \in I$ such that $I$ is the ideal generated by $g(x)$.

It is easy to prove that $\mathbb{F}[x]$, the set of polynomials $f(x)$ with coefficients from a field $F$ and $\mathbb{F}(x)/q(x)$, the finite ring of polynomials over the field $F$ modulo the non-zero polynomial $q(x)$ are principal ideal rings.

In summary, the congruence relationship allows us to construct finite algebraic structures consisting of equivalence classes; this process is illustrated in Figure 99. Given the set of integers we can construct the ring $\mathbb{Z}_p$ of integers modulo $p$. Similarly, given a non-zero polynomial with coefficients in $\mathbb{Z}_2$ we construct 16 equivalence classes of polynomials modulo the polynomial $f(x) = x^4 - 1$ :

$$\begin{aligned}
[0] &= \text{ the set of polynomials with remainder 0 when divided by f(x)} \\
[1] &= \text{ the set of polynomials with remainder 1 when divided by f(x)} \\
[\text{x}] &= \text{ the set of polynomials with remainder x when divided by f(x)} \\
[1+\text{x}] &= \text{ the set of polynomials with remainder 1+x when divided by f(x)} \\
&\vdots \\
[x^2 + x^3] &= \text{ the set of polynomials with remainder } x^2 + x^3 \text{ when divided by f(x)} \\
[x + x^2 + x^3] &= \text{ the set of polynomials with remainder } x + x^2 + x^3 \text{ when divided by f(x)}
\end{aligned}$$

We now discuss briefly <u>extension Galois fields</u> $GF(q)$ with $q = p^m$ and $p$ a prime number. The extension field $GF(p^m)$ consists of $p^m$ vectors with $m$ components; each component is

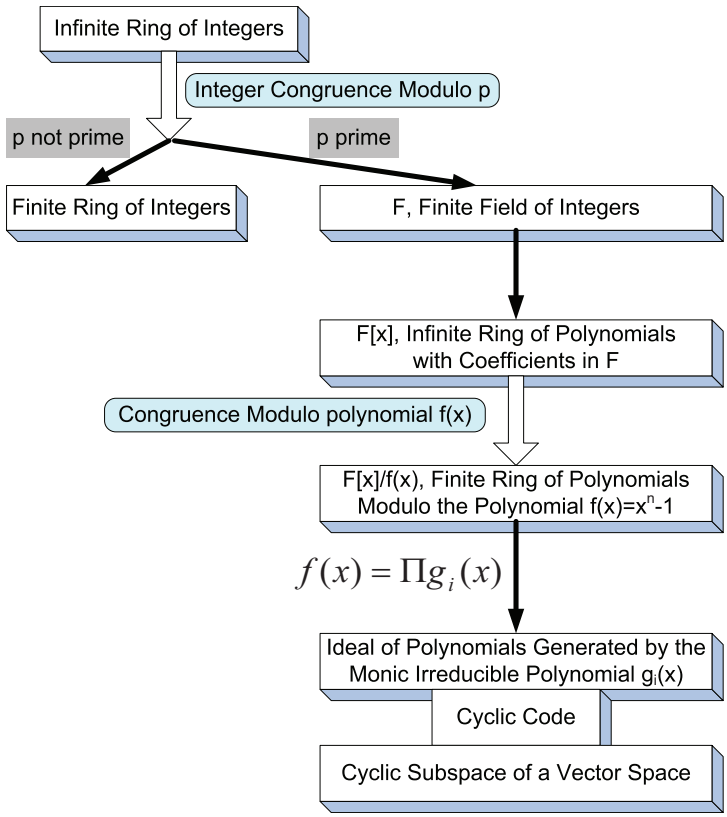



Figure 99: The congruence relationship allows us to construct finite algebraic structures. The integer congruence modulo $p$ leads to a finite ring of integers when $p$ is not a prime number, or to a finite field $F$ when p is prime. The infinite ring of polynomials over the finite field $F$ and the congruence modulo the polynomial $f(x)$ allow us to construct a finite ring of polynomials, $\mathbb{F}[x]/f(x)$. When $f(x) = x^n - 1 = \Pi_{i=1}^{q} g_i(x)$ we can construct ideals of polynomials generated by $g_i(x)$. Such ideals of polynomials are equivalent to cyclic subspaces of a vector space over F, $V_n(F)$, or cyclic codes discussed in Section 4.11.

an element of $GF(p)$. The vector $a = (a_1, a_2, \ldots, a_{m-2}, a_m)$, with $a_i \in GF(p)$, can also be represented as the polynomial $a(x) = a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \ldots + a_1x + a_0$.

The addition operation in $GF(p^m)$ is defined componentwise. The multiplication is more intricate: if $f(x) \in GF(p^m)$ is an irreducible polynomial of degree $m$ with coefficients in $GF(p)$ then multiplication is defined as $a(x) \cdot b(x) \mod f(x)$; thus, if $a(x) \cdot b(x) = q(x) \cdot f(x) + r(x)$ then $r(x) = r_{m-1}x^{m-1} + r_{m-2}x^{m-2} + \ldots + r_1x + r_0$ is a polynomial of degree at most $m-1$. The identity element is $I(x) = 1$ or the $m$-tuple $(1, 0, \ldots 0, 0)$. It is left as an exercise to prove that the axioms for a finite field are satisfied. For example, the $2^m$ elements of $GF(2^m)$ are binary $m$-tuples. The addition is componentwise, modulo 2. If $a, b \in GF(2^m)$ and

$$a = (a_0, a_1, \ldots, a_{m-2}, a_{m-1}), \quad b = (b_0, b_1, \ldots, b_{m-2}, b_{m-1}), \quad a_i, b_i \in GF(2),\ 0 \leq i \leq m-1$$

then

$$a + b = [(a_0 + b_0) \mod 2, (a_1 + b_1) \mod 2, \ldots, (a_{m-2} + b_{m-2}) \mod 2, (a_{m-1} + b_{m-1}) \mod 2].$$

An important property of an extensions field $GF(q^m)$: a polynomial $g(x)$ irreducible over

$GF(q)$ could be factored in $GF(q^m)$. For example, the polynomial $g(x) = x^2 + x + 1$ is irreducible over $GF(2)$ but can be factored over $GF(2^2)$:

$$g(x) = (x + \alpha)(x + \alpha^2)$$

with $\alpha$ a primitive element of $GF(2^2) = \{0, \alpha^0, \alpha^1, \alpha^2\}$ and $g(\alpha) = 0$.

**Example.** *Construct $GF(2^4)$ using the irreducible polynomial $f(x) = x^4 + x + 1$.*

First, we identify the primitive element $\alpha$ which must satisfy the equation $\alpha^{q-1} = 1$. In this case $q = 2^4 = 16$ and $\alpha = x$ and $\alpha^{15} = x^{15} = (x^4+x+1)(x^{11}+x^8+x^7+x^5+x^3+x^2+x)+1$ thus, $\alpha^{15} = 1 \mod (x^4 + x + 1)$.

We construct all powers of $\alpha = x$, or the binary 4-tuple $(0, 0, 1, 0)$. The first elements are $\alpha^2 = x^2, \alpha^3 = x^3$. Then $\alpha^4 \mod (x^4 + x + 1) = x + 1$. Indeed, the remainder when dividing $x^4$ by $(x^4 + x + 1)$ is $(x + 1)$ as $x^4 = 1 \cdot (x^4 + x + 1) + (x + 1)$. The next two elements are $\alpha^5 = x \cdot x^4 = x(x + 1) = x^2 + x$ and $\alpha^6 = x \cdot x^5 = x \cdot (x^2 + x) = x^3 + x^2$.

The other elements are: $\alpha^7 = x^3 + x + 1$, $\alpha^8 = x^2 + 1$, $\alpha^9 = x^3 + x$, $\alpha^{10} = x^2 + x + 1$, $\alpha^{11} = x^3 + x^2 + x$, $\alpha^{12} = x^3 + x^2 + x + 1$, $\alpha^{13} = x^3 + x^2 + 1$, $\alpha^{14} = x^3 + 1$, and $\alpha^{15} = 1$.

Table 12: $GF(2^4)$ construction when the irreducible polynomial and the primitive element are: (Left) $f(x) = x^4 + x + 1$ and $\alpha = x$; (Right) $h(x) = x^4 + x^3 + x^2 + x + 1$ and $\beta = x + 1$.

| Vector | Primitive element | Polynomial | Vector | Primitive element | Polynomial |
|---|---|---|---|---|---|
| 0000 | | $0$ | 0000 | | $0$ |
| 0010 | $\alpha^1$ | $x$ | 0011 | $\beta^1$ | $x + 1$ |
| 0100 | $\alpha^2$ | $x^2$ | 0101 | $\beta^2$ | $x^2 + 1$ |
| 1000 | $\alpha^3$ | $x^3$ | 1111 | $\beta^3$ | $x^3 + x^2 + x + 1$ |
| 0011 | $\alpha^4$ | $x + 1$ | 1110 | $\beta^4$ | $x^3 + x^2 + x$ |
| 0110 | $\alpha^5$ | $x^2 + x$ | 1101 | $\beta^5$ | $x^3 + x^2 + 1$ |
| 1100 | $\alpha^6$ | $x^3 + x^2$ | 1000 | $\beta^6$ | $x^3$ |
| 1011 | $\alpha^7$ | $x^3 + x + 1$ | 0111 | $\beta^7$ | $x^2 + x + 1$ |
| 0101 | $\alpha^8$ | $x^2 + 1$ | 1001 | $\beta^8$ | $x^3 + 1$ |
| 1010 | $\alpha^9$ | $x^3 + x$ | 0100 | $\beta^9$ | $x^2$ |
| 0111 | $\alpha^{10}$ | $x^2 + x + 1$ | 1100 | $\beta^{10}$ | $x^3 + x^2$ |
| 1110 | $\alpha^{11}$ | $x^3 + x^2 + x$ | 1011 | $\beta^{11}$ | $x^3 + x + 1$ |
| 1111 | $\alpha^{12}$ | $x^3 + x^2 + x + 1$ | 0010 | $\beta^{12}$ | $x$ |
| 1101 | $\alpha^{13}$ | $x^3 + x^2 + 1$ | 0110 | $\beta^{13}$ | $x^2 + x$ |
| 1001 | $\alpha^{14}$ | $x^3 + 1$ | 1010 | $\beta^{14}$ | $x^3 + x$ |
| 0001 | $\alpha^{15}$ | $1$ | 0001 | $\beta^{15}$ | $1$ |

**Example.** *Construct $GF(2^4)$ using the irreducible polynomial $h(x) = x^4 + x^3 + x^2 + x + 1$.*

The primitive element is $\beta = x+1$. To show that $\beta^{15} = 1$ we carry out a binomial expansion and a polynomial division and conclude that $(x+1)^{15} = 1 \mod (x^4 + x^3 + x^2 + x + 1)$.

Now $\beta^2 = (x+1)\cdot(x+1) = x^2+1$. Then $\beta^{k+1} = \beta^k(x+1)$. For example, after computing $\beta^5 = x^3 + x^2 + 1$ we can compute $\beta^6 = (x^3 + x^2 + 1)\cdot(x+1)$; as $x^4 = (x^3 + x^2 + x + 1)$ it follows that $\beta^6 = x^3$.

Table 12 shows the correspondence between the 4-tuples, the powers of the primitive element, and polynomials for the two examples.

**Example.** *Identify the primitive element of $GF(3^2)$ constructed as class of remainders of the irreducible polynomial $f(x) = x^2 + 1$.* This is another example when the primitive element is not $x$ but $\alpha = x + 1$. Indeed, the non-zero elements of the field are:

$$\begin{array}{llll} 1 = \alpha^0 & 2 = \alpha^4 & x = \alpha^6 & 2x = \alpha^2 \\ 1 + x = \alpha^1 & 2 + x = \alpha^7 & 1 + 2x = \alpha^3 & 2 + 2x = \alpha^5 \end{array}$$

The order of the each of the eight elements of $GF(3^2)$ is given by the power of the primitive element. For example, the order of $2 + x$ is 7, the order of $x$ is 6.

It is easy to see if $\beta \in \{GF(q) - 0\}$ then $\beta^{q-1} = 1$. Observe that the multiplicative inverse of $\beta$, is in $GF(q)$ and let the $(q-1)$ distinct non-zero elements $GF(q)$ be $a_1, a_2, \ldots a_{q-1}$. Then all the products $\beta\cdot a_1, \beta\cdot a_2, \ldots, \beta\cdot a_{q-1}$ must be distinct. Indeed, if $\beta\cdot a_i = \beta\cdot a_j, \;\; i \neq j$ then $(\beta^{-1}\cdot\beta)\cdot a_i = (\beta^{-1}\cdot\beta)\cdot a_j \;\; i \neq j$ or $a_i = a_j$ which contradicts that $a_i$ and $a_j$ are distinct. It follows immediately that the product of $(q-1)$ distinct non-zero elements of $GF(q)$ must be the same

$$(\beta\cdot a_1)\cdot(\beta\cdot a_2)\ldots\cdot(\beta\cdot a_{q-1}) = a_1\cdot a_2\cdot\ldots a_{q-1} \quad \text{or} \quad \beta^{q-1}\cdot(a_1\cdot a_2\cdot\ldots a_{q-1}) = a_1\cdot a_2\cdot\ldots a_{q-1}$$

and this implies that $\beta^{q-1} = 1$. Another way of reading this equation is that any element $\beta \in GF(q)$ satisfies the equation $\beta^q - \beta = 0$; this leads us to the conclusion that any element of $GF(q)$, must be a root of the equation $x^q - x = 0$.

It is also easy to show that any element $\beta \in GF(q^m)$ is in $GF(q)$ if and only if $\beta^q = \beta$. A polynomial of degree $n$ in $x$ with coefficients in $GF(q)$ has at most $n$ roots in $GF(q)$. The condition $\beta^q = \beta$ shows that $\beta$ is a root of the polynomial equation $x^q - x = 0$ which has as roots all the elements of $GF(q)$. We conclude that $\beta$ is in $GF(q)$ if and only if $\beta^q = \beta$.

We now show that the orders of all non-zero elements of $GF(q)$ divide $(q-1)$. Consider an element $\beta \in GF(q)$ of order $a = \text{ord}(\beta)$ thus $\beta^a = 1$. Let us assume that $a$ does not divide $(q-1)$ so $q - 1 = a\cdot b + c \quad \text{with} \quad 1 \leq c < a$. We have shown earlier that $\forall\beta \in GF(q)$ we have the equality $\beta^{q-1} = 1$. Thus,

$$\beta^{q-1} = \beta^{a\cdot b+c} = \beta^{a\cdot b}\cdot\beta^c = (\beta^a)^b\cdot\beta^c = 1.$$

But $(\beta^a)^b = 1$ thus, $\beta^c = 1$. This is possible only if $c = 0$ thus $q - 1 = a\cdot b$ and $a = \text{ord}(\beta)$ divides $q - 1$.

Finally, we show that there exists a primitive element in any finite field $GF(q)$. Assume that $\beta$ is the element of the largest order, $b$ in $GF(q)$, $\beta^b = 1$. In the trivial case $b = q - 1$ and we have concluded the proof, $\beta$ is the primitive element of $GF(q)$.

Let us now concentrate on the case when $b < q-1$. Assume that the orders of all elements in $GF(q)$ divide $b$; this implies that every non-zero element of $GF(q)$ satisfies the polynomial equation $y^b - 1 = 0$. This equation has at most $b$ roots and can generate at most $b$ elements of $GF(q)$; we have reached a contradiction with the assumption $b < q-1$.

Let $\gamma$ be the element of $GF(q)$ whose order, $c$, does not divide $b$ thus,

$$\gcd(b, c) = d > 1, \quad \Longrightarrow \quad c = d \cdot g \quad \text{with} \;\; \gcd(g, b) = 1.$$

But $\gamma^c = 1$ implies that there exists an element $\delta$ of order $d$ in $GF(q)$; indeed, $\gamma^c = \gamma^{c \cdot g} = (\gamma^g)^d = 1$. Thus $\delta = \gamma^g \in GF(q)$.

Let us now consider another element of $GF(q)$ constructed as the product of two elements $\beta \cdot \delta$; it is easy to see that this is an element of order $b \cdot d$ as $\beta^b \cdot \delta^d = 1$. We have thus found an element in $GF(q)$ of order $b \cdot d$ and this contradicts the assumption $b$ is the order of the largest element.

The following proposition summarizes some of the results discussed above.

**Proposition.** *Consider a finite field with $q$ elements, $GF(q)$ and an element $\alpha \in GF(q)$; the following properties are true:*

1. $\alpha^j = 1 \quad \Longleftrightarrow \quad \text{ord}(\alpha) = j$.
2. $\text{ord}(\alpha) = s \quad \Longrightarrow \quad \alpha^1 \neq \alpha^2 \neq \ldots \neq \alpha^s$.
3. $\text{ord}(\alpha) = s, \; \text{ord}(\beta) = r, \; \gcd(s, r) = 1 \quad \Longrightarrow \quad \text{ord}(\alpha \cdot \beta) = s \cdot r$.
4. $\text{ord}(\alpha^j) = \frac{\text{ord}(\alpha)}{\gcd(\text{ord}(\alpha), j)}$.
5. $GF(q)$ has an element of order $(q-1)$.
6. The order of any element in $GF(q)$ divides $(q-1)$.
7. $\gamma^q - \gamma = 0, \quad \forall \gamma \in GF(q)$.

After this brief review of algebras we start our discussion of error correcting codes with an informal introduction.

## 4.5 Linear Codes

Linear algebra allows an alternative description of a class of block codes called *linear codes.* In this section we show that an $[n, k]$ linear code $\mathcal{C}$ is a $k$-dimensional subspace $V_n^{(k)}$ of an $n$-dimensional vector space $V_n$. Once we select a set of $k$ linearly independent vectors in $V_n$ we can construct a generator matrix $G$ for $\mathcal{C}$. Then the code $\mathcal{C}$ can be constructed by multiplying the generator matrix $G$ with message vectors in $V_k$. Several classes of linear codes, codes defined as subspaces of a vector space, are known, including the cyclic codes and the BCH codes discussed in Section 4.11. Throughout the presentation of linear codes we use the term $n$-tuple instead of vector in an $n$-dimensional vector space.

Recall that an $[n, M]$ block code $\mathcal{C}$ over the alphabet $\mathcal{A}$ with $q$ symbols encodes $k = \log_q M$ information symbols into codewords with $n = k + r$ symbols. Two alternative descriptions of a linear code $\mathcal{C}$ are used: $[n, k]$ code and $[n, k, d]$ code with $d$ the Hamming distance of

the code. Decoding is the reverse process, extracting the message from a codeword. Both processes can be simplified when the code has an algebraic structure.

We consider an alphabet with $q$ symbols from a finite field $F = GF(q)$ and vectors with $n$ or $k$ components also called $n$-tuples and, respectively, $k$-tuples; the corresponding vector spaces are denoted as $V_n(F)$ and $V_k(F)$. When there is no confusion about the alphabet used instead of $V_n(F)$ and $V_k(F)$ we shall use the notation $V_n$ and $V_k$, respectively. These two vector spaces can also be regarded as the finite fields $GF(p^n)$ and $GF(p^k)$, respectively. A subspace of dimension $k$ of the vector space $V_n$ is denoted as $V_n^{(k)}$. Most of our examples cover binary codes, the case when $q = 2$.

We start this section with formal definitions and then discuss relevant properties of linear codes. In our presentation we refer to the linear code $\mathcal{C}$ as an $[n, k]$ code.

A <u>linear $[n, k]$ code</u> over $F$ is a $k$-dimensional subspace of the vector space $V_n$. A linear code is an one-to-one mapping $f$ of $k$-tuples from the *message space* to $n$-tuples

$$f : V_k \mapsto V_n$$

with $n > k$. The $n$-tuples selected as codewords form a *subspace* of $V_n^{(k)} \subset V_n$ spanned by the $k$ linearly independent vectors. Given a set of messages

$$M = \{m_1, m_2, \ldots, m_{q^k}\} \text{ where } m_i = (m_{i,1}, m_{i,2}, \ldots, m_{i,n}) \in V_k \text{ and } m_{i,j} \in F$$

and a basis $B_k$ for the $k$-dimensional subspace $V_n^{(k)} \subset V_n$, $B_k = \{v_1, v_2, \ldots, v_k\}$ with $v_i \in V_n$, <u>encoding</u> can be defined as the product of vectors in $M$ with vectors in $B_k$; the code $\mathcal{C}$ is then

$$\mathcal{C} = \{c_i = \sum_{j=1}^{k} m_{i,j} \cdot v_j \quad \forall m_{i,j} \in M \text{ with } v_j \in B_k\}.$$

**Example.** *Consider a 3-dimensional subspace of a vector space over the finite field with $q = 2$ and $n = 6$.* We choose the following three vectors as a basis

$$B_k = \{v_1 = (111000) \;\; v_2 = (100110) \;\; v_3 = (110011)\}.$$

The message space consists of all 3-tuples

$$\begin{aligned} M = \{&m_1 = (000), \quad m_2 = (001), \quad m_3 = (010), \quad m_4 = (011), \\ &m_5 = (100), \quad m_6 = (101), \quad m_7 = (110), \quad m_8 = (111)\} \end{aligned}.$$

Each codeword of the code $\mathcal{C}$ is a 6-tuple obtained as a mapping of the products of individual message bits with the basis vectors:

$$\begin{aligned}
m_1 = (000) \quad &\rightarrow \quad c_1 = 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 = (000000) \\
m_2 = (001) \quad &\rightarrow \quad c_2 = 0 \cdot v_1 + 0 \cdot v_2 + 1 \cdot v_3 = (110011) \\
m_3 = (010) \quad &\rightarrow \quad c_3 = 0 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 = (100110) \\
m_4 = (011) \quad &\rightarrow \quad c_4 = 0 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3 = (010101) \\
m_5 = (100) \quad &\rightarrow \quad c_5 = 1 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 = (111000) \\
m_6 = (101) \quad &\rightarrow \quad c_6 = 1 \cdot v_1 + 0 \cdot v_2 + 1 \cdot v_3 = (001011)
\end{aligned}$$

$$
\begin{aligned}
m_7 = (110) \quad &\rightarrow \quad c_7 \;=\; 1 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 = (011110) \\
m_8 = (111) \quad &\rightarrow \quad c_8 \;=\; 1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3 = (101101)
\end{aligned}
$$

An $[n, k]$ linear code $\mathcal{C}$ is characterized by a $k \times n$ matrix whose rows form a basis in $V_n$; this matrix is called the <u>generator matrix</u> of the linear code $\mathcal{C}$.

**Example.** *The generator matrix of the code $\mathcal{C}$ with $q = 2, n = 6, k = 3$ is*

$$
\mathrm{G} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.
$$

The code $\mathcal{C} = \{c_1, c_2 \ldots c_8\}$ generated by matrix G is obtained as products of the vectors in the message space $M$ with G. For example:

$$
c_3 = m_3 \cdot \mathrm{G} = (010) \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} = (100110)
$$

$$
c_7 = m_7 \cdot \mathrm{G} = (110) \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} = (011110)
$$

Given two $[n, k]$ linear codes, $\mathcal{C}_1$ and $\mathcal{C}_2$, over the filed $F$ with generator matrices $\mathrm{G}_1$ and $\mathrm{G}_2$, respectively, we say that they are <u>equivalent codes</u> if

$$
\mathrm{G}_2 = \mathrm{G}_1 \cdot P
$$

with $P$ a permutation matrix and $(\cdot)$ denoting matrix multiplication.

**Example.** *Consider the generator matrix* G *of the code $\mathcal{C}$*

$$
\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}
$$

*and a permutation matrix $P$ given by*

$$
P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.
$$

We construct the generator matrix $\mathrm{G}_p = \mathrm{G}P$ as

$$\mathrm{G}_p = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Recall that the code $\mathcal{C}$ is

$$\mathcal{C} = \{(000000)\ (110011)\ (100110)\ (010101)\ (111000)\ (001011)\ (011110)\ (101101)\}.$$

It is easy to see that $\mathcal{C}_p$, the code generated by $\mathrm{G}_p$ is

$$\mathcal{C}_p = \{(000000)\ (111001)\ (011010)\ (100011)\ (101100)\ (010101)\ (110110)\ (001111)\}.$$

The permutation matrix $P$ maps columns of matrix G to columns of $\mathrm{G}_p$ as follows: $2 \mapsto 1$, $5 \mapsto 2$, $1 \mapsto 3$, $3 \mapsto 4$, $4 \mapsto 5$, and $6 \mapsto 6$. To map column $i$ of the original matrix G into column $j$ of matrix $\mathrm{G}_p$ then the element $p_{ij}$ of the permutation matrix must be $p_{ij} = 1$.

This example shows that being equivalent does not mean that the two codes are identical. Since a $k$-dimensional subspace of a vector space is not unique, the generator matrix of a code is not unique but one of them may prove to be more useful than the others. We shall see shortly that among the set of equivalent codes there is one whose generator matrix is of the form $\mathrm{G} = [I_k A]$. Indeed, given the $k \times n$ generator matrix G of code $\mathcal{C}$ with a subset of columns being the binary representation of the integers 1 to $k$ we can easily construct the permutation matrix that exchanges the columns of G to obtain the generator matrix $\mathrm{G}_I = [I_k A]$ of the equivalent code $\mathcal{C}_I$.

Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$ the orthogonal complement, or the <u>dual of the code $\mathcal{C}$</u> denoted as $\mathcal{C}^\perp$ consists of vectors (or $n$-tuples) orthogonal to every vector in $\mathcal{C}$

$$\mathcal{C}^\perp \ = \ \{c \in V_n : \quad c \cdot w = 0, \quad \forall\ w \in \mathcal{C}\}$$

Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$ let H be the generator matrix of the dual code $\mathcal{C}^\perp$. Then $\forall c \in V_n, \ \ \mathrm{H}c^T = 0 \ \ \Leftrightarrow \ \ c \in \mathcal{C}$; the matrix H is called the <u>parity-check matrix of $\mathcal{C}$</u>.

Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$ with the parity check matrix H, the <u>error syndrome</u> $s$ of the $n$-tuple $v$ is defined as $s = \mathrm{H}v^T$. There is a one-to-one correspondence between the error syndrome and the bits in error, thus, the syndrome is used to determine if an error has occurred and to identify the bit in error. We wish to have an efficient algorithm to identify the bit(s) in error once the syndrome is calculated, a topic addressed in Section 4.6. A linear code $\mathcal{C}$ with the property that $\mathcal{C}^\perp \subset \mathcal{C}$ is called <u>weakly self-dual</u>. When $\mathcal{C} = \mathcal{C}^\perp$ the code is called <u>strictly self-dual</u>.

**Some properties of linear codes**. We now discuss several properties of linear codes over a finite field. There are many $k$-dimensional subspaces of the vector space $V_n$ thus, many $[n, k]$ codes; their exact number is given by the following proposition.

**Proposition.** *The number of $k$-dimensional subspaces $V_n^{(k)}$ of the vector space $V_n$ over $GF(q)$, with $n > k$ is*

$$s_{n,k} = \frac{(q^n-1)(q^{n-1}-1)\dots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\dots(q-1)}$$

*Proof:* First, we show that there are

$$(q^n-1)(q^n-q)\dots(q^n-q^{k-1})$$

ordered sets of $k$ linearly independent vectors in $V_n$. The process of selecting one by one the $k$ linearly independent vectors is now described. To select the first one we observe that there are $q^n-1$ non-zero vectors in $V_n$. This explains the first term. There are $q$ vectors that are linearly dependent on a given vector. Thus, for every vector in the first group there are $(q^n-q)$ linearly independent vectors. Therefore, there are $(q^n-1)(q^n-q)$ different choices for the first two vectors. The process continues until we have selected all $k$ linearly independent vectors of the $k$-dimensional subspace.

A similar reasoning shows that a $k$-dimensional subspace consists of

$$(q^k-1)(q^k-q)\dots(q^k-q^{k-1})$$

ordered sets of $k$ linearly independent vectors. This is the number of the ordered sets of $k$ linearly independent vectors generating the same $k$-dimensional subspace. To determine the number of distinct $k$-dimensional subspaces we divide the two numbers. □

**Example.** *If $q=2, n=6, k=3$ the number of 3-dimensional subspaces of $V_6$ is*

$$s_{n,k} = \frac{(2^6-1)(2^5-1)(2^4-1)}{(2^3-1)(2^2-1)(2^1-1)} = \frac{63\times 31\times 15}{7\times 3\times 1} = 1395.$$

The weight of the minimum weight non-zero vector, or $n$-tuple, $w \in \mathcal{C}$, is called the Hamming weight of the code $\mathcal{C}$, $w(\mathcal{C}) = min\{w(c) : \forall c \in \mathcal{C}, c \neq 0\}$.

**Proposition.** *The Hamming distance of a $[n,k]$-linear code $\mathcal{C}$ over $GF(2)$ is equal to the Hamming weight of the code*

$$d(\mathcal{C}) = w(\mathcal{C}).$$

*Proof:* $\forall c_i, c_j \in \mathcal{C}, c_i \neq c_j, c_i \oplus c_j = c_k \in \mathcal{C}$ and $d(c_i,c_j) = w(c_i \oplus c_j)$ with $\oplus$ the bitwise addition modulo 2 in $GF(2)$. Therefore, $d(\mathcal{C}) = min\{w(c_k) : \forall c_k \in \mathcal{C}, c_k \neq 0\} = w(\mathcal{C})$. □

**Proposition.** *Given an $[n,k]$ linear code $\mathcal{C}$ over the field $F$, among the set of generator matrices of $\mathcal{C}$, or of codes equivalent to $\mathcal{C}$, there is one, where the information symbols appear in the first $k$ positions of a codeword. Thus, the generator matrix is of the form*

$$\mathrm{G} = [I_k A].$$

Here $I_k$ is the $k \times k$ identity matrix and $A$ is a $k \times (n-k)$ matrix.

**Proposition.** *Given an $[n,k]$ linear code $\mathcal{C}$ over the field $F$, its dual, $\mathcal{C}^\perp$, is an [n,n-k] linear code over the same field, $F$. If $\mathrm{G} = [I_k A]$ is a generator matrix for $\mathcal{C}$, or for a code equivalent to $\mathcal{C}$, then $\mathrm{H} = [-A^T I_{n-k}]$ is a generator matrix for its dual code, $\mathcal{C}^\perp$.*

*Proof:* It is relatively easy to prove that $\mathcal{C}^\perp$ is a subspace of $V_n$. From the definition of H

$$\mathrm{H} = [-A^T I_{n-k}] = \begin{pmatrix} -a_{1,1} & -a_{2,1} & -a_{3,1} & \dots & -a_{n,1} & 1 & 0 & 0 & \dots & 0 \\ -a_{1,2} & -a_{2,2} & -a_{3,3} & \dots & -a_{n,2} & 0 & 1 & 0 & \dots & 0 \\ -a_{1,3} & -a_{2,3} & -a_{3,3} & \dots & -a_{n,3} & 0 & 0 & 1 & \dots & 0 \\ \dots & & & & & & & & & \\ -a_{1,n-k} & -a_{2,n-k} & -a_{3,n-k} & \dots & -a_{n,n-k} & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

it follows that

$$\mathrm{H}^T = \begin{pmatrix} -a_{1,1} & -a_{1,2} & -a_{1,3} & \dots & -a_{1,n-k} \\ -a_{2,1} & -a_{2,2} & -a_{2,3} & \dots & -a_{2,n-k} \\ -a_{3,1} & -a_{3,2} & -a_{3,3} & \dots & -a_{3,n-k} \\ \dots & & & & \\ -a_{n,1} & -a_{n,2} & -a_{n,3} & \dots & -a_{n,n-k} \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix}.$$

Then,

$$\mathrm{G}\mathrm{H}^T = [I_k A] \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix} = 0.$$

It is easy to see that the element $c_{ij}$ $1 \le i \le k$, $1 \le j \le n-k$ of the product is $c_{ij} = -a_{ij} + a_{ij} = 0$. Note also that in $GF(q)$ the additive inverse of $a_{ij}$ is an element such that $a_{ij} + (-a_{ij}) = 0$; for example, $(-1) = 1$ in $GF(2)$ and $(-1) = 2$ in $GF(3)$.

Call span(H) the subspace generated by the rows of H. The rows of H are linearly independent so the row space of H spans an $(n-k)$ subspace of $V_n$. If the rows of H are orthogonal to the rows of G it follows that

$$\mathrm{span}(\mathrm{H}) \subseteq \mathcal{C}^\perp.$$

To complete the proof we have to show that $\mathcal{C}^\perp \subseteq \mathrm{span}(\mathrm{H})$. We leave this as an exercise to the reader. Thus, $\mathcal{C}^\perp = \mathrm{span}(\mathrm{H})$ and $\dim(\mathcal{C}^\perp) = n-k$.

**Example.** *The previous proposition suggests an algorithm to construct the generator matrix of the dual code, $\mathcal{C}^{\perp}$. Let the generator matrix of the code*

$$\mathcal{C} = \{(000000)\ (101001)\ (011010)\ (110011)\ (101100)\ (000101)\ (110110)\ (011111)\}$$

be

$$\mathrm{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

This matrix can be transformed by permuting the columns of G. We map columns $1 \mapsto 4$, $2 \mapsto 5$, $3 \mapsto 6$, $4 \mapsto 1$, $5 \mapsto 2$, and $6 \mapsto 3$ to construct $\mathrm{G}_I$

$$\mathrm{G}_I = \mathrm{G}P = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

or,

$$\mathrm{G}_I = [I_3 A] \quad \text{with} \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

It is easy to observe that in this case $A^T = A$ and $\mathrm{H} = [-A\ I_3]$, or:

$$\mathrm{H} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} = \mathrm{G}.$$

**Proposition.** *The necessary and sufficient condition for the linear code $\mathcal{C}$ with generator matrix* G *to be weakly self-dual is that:*

$$\mathrm{G}^T\mathrm{G} = 0.$$

The proof of this proposition is left as an exercise for the reader.

**Proposition.** *Given an $[n, k]$ linear code $\mathcal{C}$ over the field $F$ with the parity check matrix* H*, every set of $(d-1)$ columns of* H *are linearly independent if and only if $\mathcal{C}$ has a distance at least equal to $d$.*

**Example.** *The parity-check matrix of a code capable of correcting a single error.* A single-error correcting code must have a distance $d \geq 3$. According to the previous proposition we need a matrix H such that no 2 or fewer than 2 columns are linearly dependent. Let us first examine the set of single columns. If we eliminate the all-zero column, then no single

column can be linearly dependent. Two columns $h_i$ and $h_j$ are linearly dependent if there are non-zero scalars, $\lambda_i, \lambda_j$ such that

$$\lambda_i h_i + \lambda_j h_j = 0.$$

This implies that the two columns are scalar multiple of each other. Indeed, the previous equation allows us to express $h_i = -(\lambda_j/\lambda_i)h_j$.

In summary, if we want a code capable of correcting any single error we have to construct a matrix H to be used as the parity-check matrix of the code such that:

1. The matrix H does not contain an all-zero column.
2. No two columns of H are scalar multiples of each other.

**Example.** *Generate the code with the parity-check matrix* H.

First, we transform the parity-check matrix by column operations as follows

$$\mathrm{H} \quad \Rightarrow \quad \mathrm{H}' = [I_k A].$$

Then we construct the generator matrix of the code $\mathrm{G} = [-A^T I_{n-k}]$. Finally, we multiply all vectors in the message space with the generator matrix and obtain the set of codewords.

**Example.** *Correct a single error in the $n$-tuple $v$ using the parity check matrix* H *of the* $[n, k]$ *linear code* $\mathcal{C}$. Apply the following algorithm:

1. Compute $s^T = \mathrm{H}v^T$. If $s^T = 0$ then $v \in \mathcal{C}$.
2. If $s^T \neq 0$ compare $s^T$ with $h_1, h_2, \ldots, h_n$, the columns of H.
   - If there is an index $j$ such that $s^T = h_j$, then the error is an $n$-tuple $e$ with a 1 in the $j$-th bit position and 0s elsewhere. The codeword is then $c = v + e$.
   - Else, there are two or more errors.

For example, assume that the codeword $c = (0111100)$ of the code $\mathcal{C}$ with the parity-check matrix H is affected by an error in the last bit position; the $n$-tuple received is $v = (0111101)$. Then

$$s^T = \mathrm{H}v^T = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

The vector $s^T$ is identical to the last column of H thus $e = (0000001)$ and the codeword is $c = v + e = (0111101) + (0000001) = (0111100)$. In this example H is the parity-check matrix of a Hamming code; Hamming codes are discussed in Section 4.8.

**Proposition.** *If $\mathcal{C}$ is an $[n, k]$ code with a generator matrix* G *whose columns are the $2^k - 1$ non-zero binary $k$-tuples then all non-zero codewords have weight $2^{k-1}$.*

*Hint:* About half of the $(2^k - 1)$ $k$-tuples that are columns of G have an even weight and half have an odd weight. More precisely $2^{k-1}$ have an odd weight $2^{k-1} - 1$ have an even weight. It follows that half of the components of a codeword obtained by linear combinations of the rows of G are 1s.

The weight distribution of a code $\mathcal{C}$ is a vector $A_{\mathcal{C}} = (A_0, A_1, \ldots, A_n)$ with $A_w$ the number of codewords of weight $w$; the weight enumerator of $\mathcal{C}$ is the polynomial

$$A(z) = \sum_{w=0}^{n} A_w z^w.$$

If $\mathcal{C}$ is an $[n, k, d]$ linear code then $A_w$ represents also the number of codewords at distance $d = w$ from a given codeword.

**Proposition.** *If $A(z)$ is the weight enumerator of the $[n, k, d]$ linear code $\mathcal{C}$, then the weight enumerator of its dual, $\mathcal{C}^{\perp}$, is*

$$B(z) = 2^k (1 + z)^n A\left(\frac{1 - z}{1 + z}\right).$$

In the next section we discuss algorithms for efficient decoding of linear codes.

## 4.6 Syndrome and Standard Array Decoding of Linear Codes

Now we discuss the partitioning of a vector space $V_n$ over the finite field $F$ into equivalence classes such that all $n$-tuples in the same class have the same syndrome; then we present a decoding procedure that exploits this partitioning. We start with a few definition of algebraic structures.

Let $G$ be an additive Abelian group and $S$ be a subgroup of $G$, $S \subseteq G$. Two elements $g_1, g_2 \in G$ are congruent modulo the subgroup $S$, if $g_1 - g_2 \in S$[32], or

$$g_1 \equiv g_2 \bmod S \quad \Longleftrightarrow \quad g_1 - g_2 \in S.$$

Let $G$ be an additive Abelian group and $S$ be a subgroup of $G$, $S \subseteq G$; the equivalence classes for congruence modulo $S$ are called the cosets of the subgroup $S$.[33]

[32] In the general case the congruence relationship implies that $g_1 \top g_2^{-1} \in S$ where "$\top$" is the group operation and $g_2^{-1}$ is the inverse of $g_2$ under "$\top$". In other words $g_2 \top g_2^{-1} = e$ with $e$ the neutral element of $G$ under the group operation. In case of an additive group the operation is $+$, "addition" and the "inverse" $g^{-1} = -g$.

[33] In the general case the elements defined here are called the *right cosets*, but for an Abelian group they are simply called cosets.

**Proposition.** *The order $m$ of a subgroup $S \subseteq G$, divides $n$, the order of the group (Lagrange Theorem).*

*Proof:* The elements of the subgroup are $S = \{s_1, s_2, \ldots s_m\}$. Coset $i$ consists of the elements congruent with $s_i$ and is denoted as $Ss_i$ and coset $j$ consists of the elements congruent with $s_j$ and is denoted as $Ss_j$. Let $f$ be the mapping

$$f : Ss_i \quad \rightarrow \quad Ss_j \quad \Longrightarrow \quad f(g_k s_i) = g_k s_j$$

The function $f$ maps distinct elements of the coset $Ss_i$ into distinct elements of the coset $Ss_j$. All elements of $Ss_j$ are mapped onto. Thus $f$ is an one-to-one map. This implies that the two cosets, and then obviously all cosets, have the same cardinality, $m$. But congruence does partition the set $G$ of $n$ elements into a number of disjoint cosets and $n = (number\, of cosets\, of\, S) \times m$. □

We now present an alternative definition of the cosets useful for syndrome calculation and decoding of linear codes. We discuss only binary codes, codes over $GF(2^n)$, but similar definitions and properties apply to codes over $GF(q^n)$. Throughout this section $V_n$ is $GF(2^n)$, we consider binary $n$-tuples; the code $\mathcal{C}$ is a subset of $GF(2^n)$.

Let $\mathcal{C}$ be an $[n, k]$ code, $\mathcal{C} \subset GF(2^n)$. The <u>coset of the code</u> $\mathcal{C}$ containing $v \in GF(2^n)$ is the set $v + \mathcal{C} = \{v + c,\ c \in \mathcal{C}\}$. The binary $n$-tuple of minimum weight in $\mathcal{C}_i$ is called the <u>leader of the coset</u> $\mathcal{C}_i$.

The decoding procedure we present later is based on the following two propositions.

**Proposition.** *Two binary $n$-tuples are in the same coset if and only if they have the same syndrome.*

*Proof:* Recall that if H is the parity check matrix of the code $\mathcal{C}$ and $c \in \mathcal{C}$ then $\mathrm{H}c^T = 0$. The syndrome corresponding to an $n$-tuple $v \in GF(2^n)$ is $\sigma_v = \mathrm{H}v^T$.

If the two binary $n$-tuples are in the same coset they can be written as $v_1 = c_1 + e$ and $v_2 = c_2 + e$ with $c_1, c_2 \in \mathcal{C}$, $\ e \in GF(2^n)$. The two syndromes are equal

$$\sigma_{v_1} = \mathrm{H}v_1^T = \mathrm{H}(c_1 + e)^T = \mathrm{H}e^T \ \text{ and } \ \sigma_{v_2} = \mathrm{H}v_2^T = \mathrm{H}(c_2 + e)^T = \mathrm{H}e^T \quad \Longrightarrow \quad \sigma_{v_1} = \sigma_{v_2}.$$

If the two $n$-binary tuples $v_1$ and $v_2$ have the same syndrome then

$$\mathrm{H}v_1^T = \mathrm{H}v_2^T \quad \Longrightarrow \quad \mathrm{H}(v_1 - v_2)^T = 0 \quad \Longrightarrow \quad (v_1 - v_2) \in \mathcal{C}.$$

$(v_1 - v_2)$ is a codeword thus $v_1$ and $v_2$ must be in the same coset. □

It is easy to see that there is a one-to-one correspondence between the coset leaders and the syndromes.

**Proposition.** *Let $\mathcal{C}$ be an $[n, k]$ binary linear code. $\mathcal{C}$ can correct $t$ errors if and only if all binary $n$-tuples of weight $t$ or less are coset leaders.*

*Proof:* To prove that this condition is necessary assume that $\mathcal{C}$ has the distance $d = 2t + 1$, thus it can correct $t$ errors. Let $v_1, v_2 \in GF(2^n)$; if $v_1$ and $v_2$ belong to the same coset, then

Table 13: The correspondence between the coset leaders and the syndromes

| Coset leader | Syndrome |
|---|---|
| 0000000 | $(000)^T$ |
| 0000001 | $(111)^T$ |
| 0000010 | $(011)^T$ |
| 0000100 | $(110)^T$ |
| 0001000 | $(101)^T$ |
| 0010000 | $(001)^T$ |
| 0100000 | $(010)^T$ |
| 1000000 | $(100)^T$ |

the binary $n$-tuple $(v_1 - v_2)$ is a codeword, $(v_1 - v_2) \in \mathcal{C}$, according to the definition of the coset and to the previous proposition. If their weight is at most $t$, $w(v_1), w(v_2) \leq t$, then the weight of the codeword $(v_1 - v_2)$ is $w(v_1 - v_2) \leq 2t$ which contradicts the fact that the distance of the code is $d = 2t + 1$. Thus, if the code can correct $t$ errors, all binary $n$-tuples of weight less or equal to $t$ must be in distinct cosets and each one can be elected as coset leader.

To prove that this condition is sufficient assume that all binary $n$-tuples of weight less or equal to $t$ are coset leaders but there are two codewords $c_1$ and $c_2$ such that $d(c_1, c_2) \leq 2t$. There exist two coset leaders, $v_1$ and $v_2$, such that $w(v_1), w(v_2) \leq t$ and $c_1 + v_1 = c_2 + v_2$. Indeed we could choose them such that $c_1 - c_2 = v_2 - v_1$. This is possible because $d(c_1, c_2) \leq 2t$ and also $d(v_1, v_2) \leq 2t$ as each coset leader has a weight of at most $t$. But $c_1 - c_2 = v_2 - v_1$ implies that the two syndromes are equal, $\sigma_{c_1+v_1} = \sigma_{c_2+v_2}$. This is a contradiction because we assumed that two coset leaders, $v_1$ and $v_2$ are in different cosets. Thus, the assumption that $d(c_1, c_2) \leq 2t$ is incorrect and the code is capable to correct $t$ errors. □

**Example.** *Let $\mathcal{C}$ be a $[7, 4, 3]$ code with the parity check matrix:*

$$\mathrm{H} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Table 13 shows the correspondence between the coset leaders and the syndromes for code $\mathcal{C}$.

Next we discuss an efficient decoding procedure for a linear code $\mathcal{C}$.

**Standard array decoding.** Consider an $[n, k]$ code $\mathcal{C} = \{c_1, c_2, \ldots c_{2^k}\}$ over $GF(2)$. Let $V_{2^n}$ be the vector space of all $n$-tuples over $GF(2)$. $\mathcal{C}$ is a subgroup of order $2^k$ of $V_{2^n}$ with $2^n$ elements. According to the Lagrange theorem there are $2^n/2^k = 2^{n-k}$ cosets of $\mathcal{C}$ denoted as $\{\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_i, \ldots, \mathcal{C}_{2^{n-k-1}}\}$.

We start with the coset leader $\mathcal{L}_0 = 00\ldots0$, the all-zero $n$-tuple and then construct the coset $\mathcal{C}_0 = \{\mathcal{L}_0, c_1, c_2, \ldots c_{2^k}\}$. Then we select as coset leaders $n$-tuples of weight 1. After we select all $n$-tuples of weight 1 as coset leaders we move to select those $n$-tuples of weight 2 that do not appear in any previously constructed coset and continue the process until all $2^{n-k}$

cosets have been constructed. Once we select the leader of a coset $j$, $\mathcal{L}_j$, the $i$-th element of the coset is $\mathcal{L}_j + c_i$.

The cosets could be arranged in order as rows of a table with $2^{n-k}$ rows and with $2^k$ columns. This table is called a *standard array* for code $\mathcal{C}$.

**Example.** *Consider the binary* $[6,3]$ *code with distance* 3 *and the generator matrix*

$$\mathrm{G} = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

The standard array has as a first row the codewords $c_1, c_2, c_3, c_1+c_2, c_1+c_3, c_2+c_3, c_1+c_2+c_3$ and as the first column the coset leaders.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 000000 | 001011 | 010110 | 100101 | 011101 | 101100 | 110011 | 111000 |
| 000001 | 001010 | 010111 | 100100 | 011100 | 101101 | 110010 | 111001 |
| 000010 | 001001 | 010100 | 100111 | 011111 | 101110 | 110001 | 111010 |
| 000100 | 001111 | 010010 | 100001 | 011001 | 101000 | 110111 | 111100 |
| 001000 | 000011 | 011110 | 101101 | 010101 | 100100 | 111011 | 110000 |
| 010000 | 011011 | 000110 | 110101 | 001101 | 111100 | 100011 | 101000 |
| 100000 | 101011 | 110110 | 000101 | 111101 | 001100 | 010011 | 011000 |

**Proposition.** *Consider an* $[n,k]$ *code* $\mathcal{C} = \{c_1, c_2, \ldots, c_{q^k-1}\}$ *over* $GF(2)$. *Let* $SA$ *be a standard array with cosets* $\mathcal{C}_i :\ i \in \{1, 2^{n-k}\}$ *and coset leaders* $\mathcal{L}_i :\ i \in \{1, 2^{n-k}\}$. *The* $(i+1, j+1)$ *entry is* $\mathcal{L}_i + c_j$, *as described earlier. Then* $\forall h \in \{0, 2^k\}$

$$d(\mathcal{L}_i + c_j, c_j) \leq d(\mathcal{L}_i + c_j, c_h).$$

In other words, the $j+1$ entry in column $i$ is closer to the codeword at the head of its column, $c_j$, than to any other codeword of $\mathcal{C}$.

*Proof:* Recall that $w(c_i)$ is the weight of the codeword $c_i$. Then from the definition of the entries of the standard table

$d(\mathcal{L}_i + c_j, c_j) = w(\mathcal{L}_i)$

$d(\mathcal{L}_i + c_j, c_h) = w(\mathcal{L}_i + c_j - c_h)$.

Coset $\mathcal{C}_i$ is obtained by adding the coset leader $\mathcal{L}_i$ to the codewords of $\mathcal{C}$, $c_i \in \mathcal{C}$. Clearly, the $n$-tuple $(\mathcal{L}_i + c_j - c_h)$, is a member of the same coset, $\mathcal{C}_i$ because $c_j - c_h \in \mathcal{C}$. Yet, by construction, the coset leader $\mathcal{L}_i$ is the $n$-tuple with the minimum weight in $\mathcal{C}_i$

$$w(\mathcal{L}_i) \leq w(\mathcal{L}_i + c_j - c_h).$$

□

**Proposition.** *The following procedure based on a standard table (ST) can be used to decode a received* $n$*-tuple* $v$ *of a linear code*

(i) Locate $v$ in ST.

(ii) Correct $v$ as the codeword at the top of its column.

**Example.** *If we receive $v = 111101$ we identify it at the fifth element of the last row of the standard table (ST) constructed for the previous example and decode it as the codeword at the top of the fifth column, $c = 011101$.*

In the next section we discuss the limitations imposed on our ability to construct "good" linear codes.

## 4.7 Hamming, Singleton, Gilbert-Varshamov, and Plotkin Bounds

Ideally, we want codes with a large number of codewords (thus, large value of $k$) which can correct as large a number of errors, $e$, as possible and with the lowest possible number of parity check symbols $r = n - k$. These contradictory requirements cannot be satisfied concurrently. To illustrate these limitations Table 14[34] lists the number of binary codewords for the case when the length $n$ of a codeword is in the range $5 \leq n \leq 15$ and the number of errors the code is able to correct is in the range $1 \leq e \leq 3$.

Several bounds on the parameters of a linear code have been established and in this section we discuss Hamming, Singleton, Gilbert-Varshamov, and Plotkin bounds. When $n$ and $k$ are given, the Hamming bound shows the limit of the efficiency of an $[n, k]$ code, it specifies the minimum number of redundancy symbols for the code, while the Singleton bound provides an upper limit to the distance $d$ of an $[n, k, d]$ code. The Gilbert-Varshamov and the Plotkin bounds limit the number of codewords on an $[n, k, d]$ code when $n$ and $d$ are specified.

We consider $[n, k, d]$ linear codes $\mathcal{C}$ over an alphabet with $q$ symbols; recall that an $[n, k, d]$ code encodes $k$ information symbols into codewords consisting of $n$ symbols and guarantees a minimum distance $d$ between any pair of codewords. The codes are vector spaces over the field $GF(q)$ and the codewords $c \in \mathcal{C}$ are vectors in $GF(q^n)$.

**Hamming bound.** Let $\mathcal{C}$ be an $[n, k]$ code over $GF(q)$ capable of correcting $e$ errors; then $\mathcal{C}$ satisfies the Hamming bound

$$q^n \geq q^k \sum_{i=0}^{e} \binom{n}{i} (q-1)^i.$$

*Proof:* The number of $n$-tuples within a Hamming sphere of radius $e$ around a codeword is

$$\sum_{i=0}^{e} \binom{n}{i} (q-1)^i$$

There are $q^k$ such Hamming spheres and the total number of $n$-tuples is $q^n$.

**Example.** *Determine the minimum number $r$ of redundancy bits required for a binary code able to correct any single error.* Our analysis is based on the following observations:

[34] The table is from **www.ams.org**

Table 14: The number of binary codewords for $5 \leq n \leq 15$ and $e = 1, 2, 3$.

| n | e=1 | e=2 | e=3 |
|---|---|---|---|
| 5 | 4 | 2 | - |
| 6 | 8 | 2 | - |
| 7 | 16 | 2 | 2 |
| 8 | 20 | 4 | 2 |
| 9 | 40 | 6 | 2 |
| 10 | 72 | 12 | 2 |
| 11 | 144 | 24 | 4 |
| 12 | 256 | 32 | 4 |
| 13 | 512 | 64 | 8 |
| 14 | 1024 | 128 | 16 |
| 15 | 2048 | 256 | 32 |

1. To construct an $[n, M]$ block code we select a subset of cardinality $M = 2^k$ as codewords out of the set of all $2^n$ $n$-tuples.

2. To correct all 1-bit errors, the Hamming spheres of radius one centered at all $M$ codewords must be disjoint.

3. A Hamming sphere of radius one centered around an $n$-tuple consists of $(n+1)$ $n$-tuples; indeed, the number of $n$-tuples at distance one from the center of the Hamming sphere is equal to $n$, one for each bit position and we should add one for the center.

It is easy to see that $M \times (n+1) \leq 2^n$. Indeed, the number of $n$-tuples of all the M Hamming spheres of radius one cannot exceed the total number of $n$-tuples. But $n = k + r$ thus this inequality becomes

$$2^k \times (k + r + 1) \;\leq\; 2^{(k+r)} \quad \text{or} \quad 2^r - r \;\geq\; k + 1.$$

We conclude that the minimum number of redundancy bits has a lower bound:

$$2^r \;\geq\; n + 1.$$

For example, if $k = 15$, the minimum value of $r$ is 5.

A binary code is able to correct at most $e$ errors if

$$2^n \geq 2^k \sum_{i=0}^{e} \binom{n}{i}$$

or, using the fact that $d = 2e + 1$ and $r = n - k$

$$2^r \geq \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i}.$$

When $d = 3$ this becomes $2^r \geq n + 1$ as before.

**Singleton bound.** Every $[n, k, d]$ linear code $\mathcal{C}$ over the field $GF(q)$ satisfies the inequality

$$k + d \leq n + 1.$$

*Proof:* Let $\{e_1, e_2, \ldots\}$ be a basis in the vector space $GF(q^n)$. Express every $c \in \mathcal{C}$ in this basis; then construct a new code $\mathcal{C}_s$ consisting of codewords of length $n - d + 1$ obtained by deleting the first $d - 1$ symbols of every $c \in \mathcal{C}$. The codewords of $\mathcal{C}$ are at a distance at least $d$ from one another. When we remove the first $d-1$ symbols from every $c \in \mathcal{C}$, the remaining $(n - d + 1)$-tuples $c_s \in \mathcal{C}_s$ are at least at distance 1, thus they are distinct. The original code $\mathcal{C}$ consists of $\mid \mathcal{C} \mid = q^k$ distinct codewords, thus the cardinality of the new code is

$$\mid \mathcal{C}_s \mid = q^k.$$

There cannot be more than $q^{n-d+1}$ tuples of length $(n - d + 1)$ in $GF(q^{n-d+1})$. Thus,

$$q^k \leq q^{n-k+1} \quad \Longrightarrow \quad k \leq n - d + 1$$

or

$$k + d \leq n + 1.$$

□

A <u>maximum distance separable (MDS)</u> $[n, k, d]$ linear code is a code $\mathcal{C}$ with the property

$$k + d = n + 1.$$

MDS codes are optimal in the following sense: for a given length of a codeword, $n$, and a given number of information symbols, $k$, the distance $d$ of the code reaches its maximum.

**Corollary.** *Consider an infinite sequence of linear codes $[n_i, k_i, d_i]$ such that when $i \to \infty$ then $n_i \to \infty$ and:*

$$R = \lim_{i \to \infty} \frac{k_i}{n_i} \qquad \delta = \lim_{i \to \infty} \frac{d_i}{n_i},$$

*Then the Singleton bound requires that*

$$\frac{k_i}{n_i} + \frac{d_i}{n_i} \leq 1 + \frac{1}{n_i}.$$

*When $i \to \infty$ then*

$$R + \delta \leq 1.$$

Figure 100(a) shows that in the limit any sequence of codes must have the rate and the relative minimum distance ($\delta = d/n$) within the triangle $R + \delta \leq 1$.

**Gilbert-Varshamov bound.** If $\mathcal{C}$ is an $[n, k, d]$ linear code over $GF(q)$ and if $M(n, d) = \mid \mathcal{C} \mid$ is the largest possible number of codewords for given $n$ and $d$ values, then the code satisfies the Gilbert-Varshamov bound

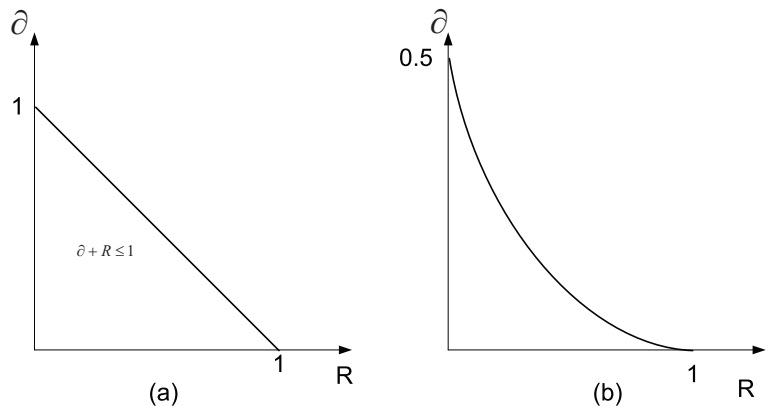



Figure 100: (a) The Singleton bound. (b) The Gilbert-Varshamov bound.

$$M(n,d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i}.$$

*Proof:* Let $c \in \mathcal{C}$. Then there is an $n$-tuple $v \in GF(q^n)$ such that $d(c,v)$, the distance between $c$ and $v$, satisfies the inequality

$$d(c,v) \leq d-1.$$

The existence of such an $n$-tuple is guaranteed because otherwise $v$ could be added to the code $\mathcal{C}$, while maintaining the minimum distance of the code as $d$. Recall that $M(n,d)$ is the largest possible number of codewords for given $n$ and $d$ values. A Hamming sphere of radius $(d-1)$ about a codeword $c \in \mathcal{C}$ contains

$$\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i$$

$n$-tuples. We assume that there are $M(n,d)$ codewords in $\mathcal{C}$ thus an equal number of Hamming spheres of radius $(d-1)$, one about each codeword, $S_{c_i},\ c_i \in \mathcal{C}$. Thus, the entire space $GF(q^n)$ is contained in the union of all Hamming spheres of radius $d-1$ around individual codewords:

$$GF(q^n) \subset \cup_{c_j \in \mathcal{C}} S_{c_j}(d-1)$$

or

$$\mid GF(q^n) \mid \leq M(n,d) \sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i.$$

It follows that

$$q^n \leq M(n,d) \sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i$$

or

$$M(n,d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i}.$$

□

The Gilbert-Varshamov bound applies not only to linear codes but also to any code of length $n$ and distance $d$. If the code has distance $d$ the Hamming spheres of radius $\lceil (d-1)/2 \rceil$ are disjoint but those of radius $d$ are not disjoint; two spheres of radius $d$ have many $n$-tuples in common thus, the sum of the volumes of all these spheres has $q^n$ as a lower bound, as we are guaranteed to include all $n$-tuple and some of them are counted multiple times.

**Plotkin bound.** If $\mathcal{C}$ is an $[n,k,d]$ linear code and if the condition $(2d > n)$ is satisfied, then the code satisfies the Plotkin bound

$$M \leq 2\lfloor \frac{d}{2d-n} \rfloor$$

with $M = \mid \mathcal{C} \mid$.

*Proof:* Let $c_x, c_y \in \mathcal{C}$. We find first a lower bound for $\sum_{c_x,c_y \in \mathcal{C}} d(c_x,c_y)$; we observe that

$$\sum_{c_x,c_y \in \mathcal{C}} d(c_x,c_y) \geq dM(M-1).$$

Indeed, there are $M(M-1)$ pairs of codewords in $\mathcal{C}$ and $d$ is the minimum distance between any pair of codewords.

We can arrange the codewords of $\mathcal{C}$ in a table with $M$ rows and $n$ columns. Consider column $i$ of this table and call $z_i$ the number of 0s and $(M - z_i)$ the number of 1s in that column. Then,

$$\sum_{c_x,c_y \in \mathcal{C}} d(c_x,c_y) = \sum_{i=1}^{n} 2z_i(M-z_i)$$

because each choice of a 0 and an 1 in the same column of the table contributes 2 to sum of distances. Now we consider two cases:

(1) $M$ is an even integer. In that case the sum on the right hand side of the equality is maximized when $z_i = M/2$. Thus we obtain an upper bound for the sum of distances:

$$\sum_{c_x,c_y \in \mathcal{C}} d(c_x,c_y) \leq \frac{1}{2}nM^2.$$

Now we combine the lower and the upper bound to get

$$M(M-1)d \leq \frac{1}{2}nM^2,$$

or

$$M \leq \frac{2d}{2d-n}$$

because we assumed that $2d - n > 0$. Since $M$ is an even integer it follows that

$$M \leq 2\lfloor \frac{d}{2d-n} \rfloor.$$

(2) $M$ is an odd integer. In that case the sum on the right hand side of the equality is maximized when $z_i = (M \pm 1)/2$. Thus we obtain an upper bound for the sum of distances:

$$\sum_{c_x, c_y \in \mathcal{C}} d(c_x, c_y) \leq \frac{1}{2} n(M^2 - 1).$$

Now,

$$M(M-1)d \leq \frac{1}{2} n(M^2 - 1),$$

or,

$$M \leq \frac{2d}{2d-n} - 1.$$

Again we use the fact that $M$ is an integer:

$$M \leq \lfloor \frac{2d}{2d-n} - 1 \rfloor = \lfloor \frac{2d}{2d-n} \rfloor - 1 \leq 2\lfloor \frac{d}{2d-n} \rfloor. \qquad \square$$

It is now the time to present several families of codes and, not unexpectedly, we start our journey with the codes discovered by Richard Hamming.

## 4.8 Hamming Codes

Hamming codes, introduced in 1950 by Richard Hamming, are the first error correcting codes ever invented. Hamming codes are linear codes; they can detect up to two simultaneous bit errors, and correct single-bit errors. At the time of their discovery Hamming codes were considered a great improvement over simple parity codes which cannot correct errors and can only detect an odd number of errors. Throughout this section we refer to the linear code $\mathcal{C}$ consisting of $n$-tuples with $k$ information symbols in each codeword and with distance $d$ as an $[n, k, d]$ code.

A <u>$q$-ary Hamming code of order $r$</u> over the field $GF(q)$ is an $[n, k, 3]$-linear code where

$$n = \frac{q^r - 1}{q - 1} \quad \text{and} \quad k = n - r.$$

The parity-check matrix of the code $\mathrm{H}_r$ is an $r \times n$ matrix such that it contains no all-zero columns and no two columns are scalar multiples of each other.

A code $\mathcal{C}$ such that the Hamming spheres of radius $e$ about each codeword are disjoint and exhaust the entire space of $n$-tuples is called a <u>perfect code</u> capable of correcting $e$ errors.

The last statement means that every $n$-tuple is contained in the Hamming sphere of radius $e$ about some codeword.

**Proposition.** *The Hamming code of order $r$ over $GF(q)$ is a perfect code $\mathcal{C}_r$.*

*Proof:* If $c \in \mathcal{C}_r$ then the Hamming sphere of radius $e$ about $c$ contains: (i) one $n$-tuple at distance 0 from $c$ ($c$ itself), (ii) $n(q-1)$ $n$-tuples at distance 1 from $c$, because there are $n$ symbols in an $n$-tuple and each symbol can be altered to any of the other $q-1$ letters of the alphabet. Thus, the total number of $n$-tuples in a Hamming sphere about a codeword is $(1+n(q-1))$.

There are $k$ information symbols, thus the total number of messages is $q^k$ and the total number of such Hamming spheres is also $q^k$. The total number of $n$-tuples in all the Hamming spheres about codewords is equal to the number of Hamming spheres times the number of $n$-tuples in each Hamming sphere

$$q^k(1+n(q-1)) = q^{n-r}(1+\frac{(q^r-1)}{(q-1)}(q-1)) = q^{n-r}(1+(q^r-1)) = q^n.$$

But $q^n$ is precisely the total number of $n$-tuples with elements from an alphabet with $q$ symbols. This proves that the Hamming spheres of radius $e=1$ about all codewords contain all possible $n$-tuples. The spheres of radius 1 about codewords are disjoint because the minimum distance of the code is 3.

Assume that we transmit a codeword $c \in \mathcal{C}$ where $\mathcal{C}$ is a linear code with distance $d=3$ and parity check matrix H. A single error occurs and the error vector $e$ has a Hamming weight of 1; what we receive is an $n$-tuple $v = c+e$. From the definition of the parity check matrix it follows that the syndrome associated with the $n$-tuple $v$ is

$$\mathrm{H}v^T = \mathrm{H}(c+e)^T = \mathrm{H}e^T.$$

Since $e$ has only one non-zero coordinate, say coordinate $j$ which has value $\beta$, then $\mathrm{H}e^T = \beta h_j$ where $h_j$ is the $j$-th column of the parity-check matrix H. This justifies the following proposition.

**Proposition.** *The following procedure can be used to decode any single error linear code, and in particular, Hamming codes:*

(i) Compute the syndrome $\mathrm{H}v^T$.

(ii) If the syndrome is zero no error has occurred.

(iii) If $\mathrm{H}e^T = s^T \neq 0$ then compare the syndrome with the columns of the parity-check matrix H.

(iv) If there is some integer $j$ such that $s^T = \beta h_j$ then the error vector $e$ is an $n$-tuple with $\beta$ in position $j$ and zero elsewhere. Then $c = v - e$.

(v) If there is no integer $j$ such that $s^T = \beta h_j$ then more than one error has occurred.

An $[n, n-r, d]$ code whose parity check matrix has all nonzero binary vectors of length $r$ as columns is a <u>binary Hamming code</u>.

There are $2^r - 1$ nonzero vectors of length $r$, thus $n = 2^r - 1$. The number of information symbols is $k = n - r = 2^r - 1 - r$. Recall that the minimum distance of the code is equal to the minimum number of linearly dependent columns of the parity check matrix; by definition

any two rows are linearly independent, thus $d = 3$. It follows that a binary Hamming code is a $[(2^r - 1), (2^r - 1 - r), 3]$ linear code.

**Example.** *The binary Hamming code of order $r = 3$ is a $[7, 4, 3]$-linear code.*

Indeed, $(2^3 - 1)/(2 - 1) = 7$, thus $n = 7$ and $k = n - r = 7 - 3 = 4$. The parity-check matrix of the code is

$$\mathrm{H}_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

A code obtained by adding one all-0s column and one all-1s row to the generator matrix of a binary Hamming code is called an extended binary Hamming code. An extended binary Hamming code is a $[2^r, (2^r - r - 1), 4]$ linear code.

**Example.** *The binary Hamming code and the extended binary Hamming code of order $r = 4$.* They are $[15, 11, 3]$ and, respectively, $[16, 11, 4]$ linear codes with parity check matrices, $H_4$ and $HE_4$ given by:

$$H_4 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

and

$$HE_4 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

## 4.9 Proper Ordering, and the Fast Walsh-Hadamard Transform

In Section 1.21 we have introduced Hadamard matrices and the Walsh-Hadamard transform. In this section we continue our discussion, provide an alternative definition of the Walsh-Hadamard transform, and introduce the fast Walsh-Hadamard transform.

Consider all binary $q$-tuples, $b_1, b_2, \ldots, b_{2^q}$. The proper ordering, $\pi_q$, of binary $q$-tuples is defined recursively:

$$\begin{aligned} \pi_1 &= [0, 1] \\ &\vdots \\ \pi_i &= [b_1, b_2, \ldots b_{2^i}] \\ \pi_{i+1} &= [b_1 0, b_2 0, \ldots b_{2^i} 0, b_1 1, b_2 1, \ldots, b_{2^i} 1] \quad \forall i \in \{2, q-1\}. \end{aligned}$$

**Example.** *Given that* $\pi_1 = [0, 1]$ *it follows that:*

$$\begin{aligned}
\pi_2 &= [00, 10, 01, 11] \\
\pi_3 &= [000, 100, 010, 110, 001, 101, 011, 111] \\
\pi_4 &= [0000, 1000, 0100, 1100, 0010, 1010, 0110, 1110, 0001, 1001, 0101, 1101, 0011, 1011, 0111, 1111]
\end{aligned}$$

Let $n = 2^q$ and let the $n$ $q$-tuples under the proper order be

$$\pi_q = [u_0, u_1, \ldots, u_{n-1}] \,.$$

Let us enforce the convention that the leftmost bit of $u_i$ is the least significant bit of the integer represented by $u_i$ as a $q$-tuple. For example, consider the following mappings of 4-tuples to integers:

$$0000 \mapsto 0,\ 1000 \mapsto 1,\ 0100 \mapsto 2,\ 1100 \mapsto 3,\ 0010 \mapsto 4,\ 1010 \mapsto 5,\ \ldots,\ 0111 \mapsto 14,\ 1111 \mapsto 15.$$

**Proposition.** *The matrix* $H = [h_{ij}]$ *with*

$$h_{ij} = (-1)^{u_i \cdot u_j} \qquad \forall (i, j) \in \{0, n-1\},$$

*where* $u_i$ *and* $u_j$ *are members of*

$$\pi_q = [u_0, u_1, \ldots u_{n-1}] \,,$$

*is the Hadamard matrix of order* $n = 2^q$*. Observe that the rows and the columns of the matrix* $H$ *are numbered from* $0$ *to* $n - 1$*.*

As an example, consider the case $q = 3$. It is easy to see that the matrix elements of the first row, $h_{0i}$, and the first column, $h_{i0}$, are all $+1$ because $u_0 \cdot u_i = u_i \cdot u_0 = 0, \forall i \in \{1, 7\}$ and $(-1)^0 = +1$. It is also easy to see that the diagonal element $h_{ii} = -1$ when the weight of $i$ (the number of 1's in the binary representation of integer $i$) is odd and $h_{ii} = +1$ when the weight of $i$ is even. Individual calculations of $h_{ij}, \quad i \neq j$ are also trivial. For example, $h_{32} = (-1)^{(110)\cdot(010)} = (-1)^1 = -1$. Indeed $(110) \cdot (010) = 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 = 1$. This shows that $H = H_3$.

Let $c = (c_0 c_1 c_2 \ldots c_i \ldots c_{2^q-1})$ be a binary $2^q$-tuple, $c_i = \{0, 1\}$ and let $B_q$ be a $q \times 2^q$ matrix whose columns are all $2^q$ possible $q$-tuples $b_i$. We define $c(b)$ to be the component of $c$ selected by the $q$-vector $b$ according to the matrix $B$. The binary value $c(b)$ can be either 0 or 1, depending on the value of the corresponding element of the binary $2^q$-tuple $c$.

For example, let $q = 3$ and $B_3$ be given by

$$B_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

Let $c = (0\ 1\ 1\ 1\ 1\ 0\ 0\ 1)$. Then $c(1\ 1\ 0) = 1$ because $(1\ 1\ 0)$ is the 4-th column of $B$ and the 4-th element of $c$ is 1. Similarly $c(1\ 1\ 1) = 0$ because $(1\ 1\ 1)$ is the 7-th column of $B$ and the 7-th element of $c$ is 0.

We can define a vector $R(c)$ whose components $R_i(c)$ are $(-1)^{c(b)}$ and their value can be either $+1$, or $-1$. In our example

$$R(c) = (+1,\ -1,\ -1,\ -1,\ -1,\ +1,\ +1,\ -1).$$

Indeed,

$$\begin{aligned}
R_1(c) &= (-1)^{c(100)} = (-1)^0 = +1\\
R_2(c) &= (-1)^{c(010)} = (-1)^1 = -1\\
R_3(c) &= (-1)^{c(001)} = (-1)^1 = -1\\
R_4(c) &= (-1)^{c(110)} = (-1)^1 = -1\\
R_5(c) &= (-1)^{c(011)} = (-1)^1 = -1\\
R_6(c) &= (-1)^{c(101)} = (-1)^0 = +1\\
R_7(c) &= (-1)^{c(111)} = (-1)^0 = +1\\
R_8(c) &= (-1)^{c(000)} = (-1)^1 = -1.
\end{aligned}$$

Let $d$ be a binary $q$-tuple and let $c$ be a binary $2^q$-tuple. Let $R(c) = (-1)^{c(b)}$ be a $2^q$-tuple with entries either $+1$, or $-1$ as defined earlier; then the <u>Walsh-Hadamard transform of $R(c)$</u> is defined as

$$\hat{R}(d) = \sum_{b \in B_q} (-1)^{d \cdot b} R(c))$$

or

$$\hat{R}(d) = \sum_{b \in B_q} (-1)^{d \cdot b + c(b)}.$$

**Example.** *Let $q = 3$, $d = (\ 1\ 1\ 1)^T$ and $c = (0\ 1\ 1\ 1\ 1\ 0\ 0\ 1)$. Then,*

$$\hat{R}(1\ 1\ 1) = \sum_{b \in B_3} (-1)^{(\ 1\ 1\ 1) \cdot b + c(b)}$$

$$\begin{aligned}
\hat{R}(1\ 1\ 1) &= (-1)^{(\ 1\ 1\ 1)\cdot(1\ 0\ 0)+c(1\ 0\ 0)} + (-1)^{(1\ 1\ 1)\cdot(0\ 1\ 0)+c(0\ 1\ 0)} + (-1)^{(1\ 1\ 1)\cdot(0\ 0\ 1)+c(0\ 0\ 1)}\\
&\quad +(-1)^{(1\ 1\ 1)\cdot(1\ 1\ 0)+c(1\ 1\ 0)} + (-1)^{(1\ 1\ 1)\cdot(0\ 1\ 1)+c(0\ 1\ 1)} + (-1)^{(1\ 1\ 1)\cdot(1\ 0\ 1)+c(1\ 0\ 1)}\\
&\quad +(-1)^{(1\ 1\ 1)\cdot(1\ 1\ 1)+c(1\ 1\ 1)} + (-1)^{(1\ 1\ 1)\cdot(0\ 0\ 0)+c(0\ 0\ 0)}\\
&= (-1)^{1+0} + (-1)^{1+1} + (-1)^{1+1} + (-1)^{0+1} + (-1)^{0+1} + (-1)^{0+0} + (-1)^{1+0} + (-1)^{0+1}\\
&= -2.
\end{aligned}$$

**Proposition.** *Given the binary vector*

$$t = c + \sum_{i=1}^{q} d_i v_i$$

*with* $d = (d_1 d_2 \ldots d_q)^T$ *a binary q-tuple and* $v_i$ *the i-th row of* $B_q$*, then* $\hat{R}(d)$ *is the number of* $0s$ *minus the number of* $1s$ *in* $t$*.* This proposition is the basis of a decoding scheme for first-order Reed-Muller codes $R(1, r)$ discussed in Section 4.10.

There is a faster method to compute the Walsh-Hadamard transform by using matrices $M_{2^q}^{(i)} = I_{2^{q-i}} \otimes H(2) \otimes I_{2^{i-1}}$. If $q$ is a positive integer the <u>fast Walsh-Hadamard transform</u> allows a speedup of $(2^{q+1} - 1)/3q$ over the classical Walsh-Hadamard transform.

**Proposition.** *If* $q$ *is a positive integer and* $M_{2^q}^{(i)} = I_{2^{q-i}} \otimes H(2) \otimes I_{2^{i-1}}$ *with* $H(2)$ *the Hadamard matrix,* $I_{2^{q-i}}$ *the identity matrix of size* $2^{q-i} \times 2^{q-i}$*, and* $I_{2^{i-1}}$ *the identity matrix of size* $2^{i-1} \times 2^{i-1}$*, then:*

$$H(2^q) = M_{2^q}^{(1)} M_{2^q}^{(2)} \ldots M_{2^q}^{(q)}.$$

We can prove this equality by induction. For $q = 1$ we need to prove that $H(2) = M_2^{(1)}$. We denote by $I_{n \times n}$ or simply $I_n$ the $n \times n$ identity matrix. By definition:

$$M_2^{(1)} = I_{2^{1-1}} \otimes H(2) \otimes I_{2^{1-1}} = H(2).$$

Assume that this is true for $q = k$ and consider the case $q = k + 1$. For $q = k$ we have:

$$H(2^k) = M_{2^k}^{(1)} M_{2^k}^{(2)} \ldots M_{2^k}^{(k)}.$$

Now for $q = k + 1$ we have to prove that

$$H(2^{k+1}) = M_{2^{k+1}}^{(1)} M_{2^{k+1}}^{(2)} \ldots M_{2^{k+1}}^{(k)} M_{2^{k+1}}^{(k+1)}.$$

Then

$$M_{2^{k+1}}^{(i)} = I_{2^{k+1-i}} \otimes H(2) \otimes I_{2^{i-1}} = I_2 \otimes I_{2^{k-i}} \otimes H(2) \otimes I_{2^{i-1}} = I_2 \otimes M_{2^k}^{(i)}$$

Thus,

$$H(2^{k+1}) = (I_2 \otimes M_{2^k}^{(1)})(I_2 \otimes M_{2^k}^{(2)}) \ldots (I_2 \otimes M_{2^k}^{(k)}) M_{2^{k+1}}^{(k+1)}.$$

We know that the tensor product of square matrices $V, W, X, Y$ has the following property $(V \otimes W)(X \otimes Y) = VX \otimes WY$. Applying this property repeatedly after substituting $H(2^k)$ for $M_{2^k}^{(1)} M_{2^k}^{(2)} \ldots M_{2^k}^{(k)}$ we get

$$H(2^{k+1}) = (I_2 \otimes M_{2^k}^{(1)} M_{2^k}^{(2)} \ldots M_{2^k}^{(k)})(M_{2^{k+1}}^{(k+1)}) = (I_2 \otimes H(2^k)(M_{2^{k+1}}^{(k+1)}).$$

But from the definition of $M_{2^{k+1}}^{(k+1)}$ we see that

$$M_{2^{k+1}}^{(k+1)} = H(2) \otimes I_{2^k}.$$

Thus,

$$H(2^{k+1}) = (I_2 \otimes H(2^k))(H(2) \otimes I_{2^k}) = (I_2 H(2)) \otimes (H(2^k) I_{2^k}) = H(2) \otimes H(2^k) = H(2^{k+1}).$$

Compared with the traditional Walsh-Hadamard transform, the fast Walsh-Hadamard transform allows a speedup $S_{FHT/HT}$ given by:

$$S_{FHT/HT} = \frac{2^{q+1} - 1}{3q}.$$

For example, for $q = 16$ we have $S_{FHT/HT} = 2^{17}/48 \approx 128,000/48 \approx 4,200$.

**Example.** *Given* $R(c) = (+1, -1, -1, -1, -1, +1, +1, -1)$ *let us compute* $\hat{R} = RH$ *with* $H = H(2^3) = M_8^1 M_8^2 M_8^3$. First, we calculate $M_8^1$, $M_8^2$, and $M_8^3$ as follows:

$$M_8^1 = I_4 \otimes H(2) \otimes I_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Thus,

$$M_8^1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

Then,

$$M_8^2 = I_2 \otimes H(2) \otimes I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus,

$$M_8^2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix}.$$

Finally,

$$M_8^3 = I_1 \otimes H(2) \otimes I_4 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus,

$$M_8^3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Then,

$$RM_8^1 = (+1, -1, -1, -1, -1, +1, +1, -1) \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix},$$

or,

$$RM_8^1 = (0, +2, -2, 0, 0, -2, 0, +2).$$

Next,

$$(RM_8^1)M_8^2 = (0, +2, -2, 0, 0, -2, 0, +2) \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix},$$

or,

$$(RM_8^1)M_8^1 = (-2, +2, +2, +2, 0, 0, 0, -4).$$

Finally,

$$(RM_8^1 M_8^2)M_8^3 = (-2,+2,+2,+2,0,0,0,-4)\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{pmatrix},$$

or

$$\hat{R} = (RM_8^1 M_8^2)M_8^3 = (-2,+2,+2,-2,-2,+2,+2,+6).$$

## 4.10 Reed-Muller Codes

Reed-Muller codes are linear codes with numerous applications in communication; they were introduced in 1954 by Irving S. Reed and D. E. Muller. Muller invented the codes [303] and then Reed proposed the majority logic decoding scheme [353]. In Section 4.8 we defined $\mathrm{H}_r$, the parity check matrix of a binary $(2^r-1, 2^r-1-r)$ Hamming code. $\mathrm{H}_r$ has $2^r-1$ columns, consisting of all non-zero binary $r$-tuples.

The <u>first-order Reed-Muller code</u> $R(1,r)$ is a binary code with the generator matrix

$$\mathrm{G} = \begin{pmatrix} \mathbf{1} & 1 \\ \mathrm{H}_r & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ B_r \end{pmatrix}$$

where $\mathbf{1}$ is an $(2^r-1)$-tuple of 1s. The generator matrix G of the code is constructed by augmenting $\mathrm{H}_r$ with a zero column to obtain a matrix $B_r$, and then adding as a first row a $2^r$-tuple of 1s.

**Example.** *For* $r = 3$*, given* $\mathrm{H}_3$

$$\mathrm{H}_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

we construct first

$$B_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$

and then

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

**Proposition.** *The code with the generator matrix* G*, a first-order Reed-Muller code, has codewords of length* $n = 2^r$*, the number of information symbols is* $k = r + 1$*. The code has distance* $d = 2^{r-1}$*.*

*Proof:* Recall that $\mathrm{H}_r$ is an $r \times n'$ matrix with $n' = 2^r - 1$. Then $B_r$ is an $r \times n$ matrix with $n = n' + 1 = 2^r$ ; its first $n-1$ columns, $b_i,\ 1 \leq i \leq n-1$ are all non-zero binary $r$-tuples and its last column is the all-zero $r$-tuple, $B_r = (b_1 b_2 \ldots b_{n-1} 0)$. It follows that half of the $n = 2^r$ columns of $B_r$ have even parity and half have odd parity.

The generator matrix G of the code $\mathcal{C}$ is an $(r+1) \times n$ matrix with the first row an $n$-tuple of all 1s and the other $r$ rows the rows of $B_r$. Half of the columns of G have an even parity and half have an odd parity as every column of G adds a 1 as its first element to a column of $B_r$. Every codeword $c \in \mathcal{C}$ is obtained as a product of a message $(r+1)$-tuple with the generator matrix G thus, every codeword will have half of its $2^r$ components 1s and half 0s. Thus, the minimum weight of any codeword is exactly $2^{r-1}$.

**Proposition.** *Assume that the columns of* $B_r$ *are in the proper order* $\pi_r$ *and that* $H$ *is the Hadamard matrix* $H = H(2^r)$*. Let* $w$ *be a received* $n$*-tuple when a first order Reed-Muller code is used for transmission. The following procedure is used to decode* $w$*:*

(i) Compute $R$ and $\hat{R}$.

(ii) Find a component $\hat{R}(u)$ whose magnitude is maximum. Let $u = (u_1, u_2, \ldots, u_r)^T$ be a binary $r$-tuple and let $b_i$ be the $i$-th row of $B_r$.

(iii) If $\hat{R}(u) > 0$ then decode $w$ as $\sum_{i=1}^{r} u_i b_i$.

(iv) If $\hat{R}(u) \leq 0$ then decode $w$ as $I + \sum_{i=1}^{r} u_i b_i$ with $I$ an $n$-tuple of 1s.

A justification of this decoding procedure can be found in most standard texts on error correcting codes such as [436].

**Example.** *Consider the* $B_3$ *matrix whose columns are* $\pi_3$*, in other words they are the binary expression of all* 3*-tuples in proper order, e.g., the* 4*-th column is* (0 1 1) *the binary expression of integer* 3*, the* 7*-th column is* 1 1 0 *the binary expression of integer* 6*, and so on. We assume that the leftmost bit is the least significant one.*

$$B_3 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

The corresponding generator matrix of the Reed-Muller code is

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Assume that we receive $w = (0\ 1\ 1\ 1\ 1\ 0\ 0\ 1)$. Then,

$$R = (+1,\ -1,\ -1,\ -1,\ -1,\ +1,\ +1,\ -1).$$

The Hadamard transform of $R$ is (see for example [284])

$$\hat{R} = RH_8 = (-2, \ +2, \ +2, \ -2, \ -2, \ +2, \ +2, \ -4)$$

with

$$H_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

We see that the largest component of $\hat{R}$ occurs in position 8, so we set $u$ equal to the last column of $B_r$, $u = (1\ 1\ 1)^T$. Since the largest component is negative (it is equal to $-4$) we decode $w$ as

$$v = I + \sum_{i=1}^{3} u_i b_i = (1\ 1\ 1\ 1\ 1\ 1\ 1\ 1) + 1 \cdot b_1 + 1 \cdot b_2 + 1 \cdot b_3$$

$$v = (1\ 1\ 1\ 1\ 1\ 1\ 1\ 1) + 1 \cdot (0\ 1\ 0\ 1\ 0\ 1\ 0\ 1) + 1 \cdot (0\ 0\ 1\ 1\ 0\ 0\ 1\ 1) + 1 \cdot (0\ 0\ 0\ 0\ 1\ 1\ 1\ 1) =$$
$$(1\ 1\ 1\ 1\ 1\ 1\ 1\ 1) + (0\ 1\ 1\ 0\ 1\ 0\ 0\ 1) = (1\ 0\ 0\ 1\ 0\ 1\ 1\ 0).$$

An alternative description of Reed-Muller codes is based on manipulation of Boolean monomials. A <u>monomial</u> is a product of powers of variables. For example, given the variables $x, y$, and $z$ then a monomial is of the form $x^a y^b z^c$ with $a, b$, and $c$ nonnegative integers, $a, b, c \in Z^+$. A <u>Boolean monomial</u> in variables $x_1, x_2, \ldots, x_n$ is an expression of the form

$$p = x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n} \quad \text{with} \quad r_i \in Z^+, \quad 1 \leq i \leq n.$$

The *reduced form* of $p$ is the result of applying two rules: (i) $x_i x_j = x_j x_i$; and (ii) $x_i^2 = x_i$.

A <u>Boolean polynomial</u> is a linear combination of Boolean monomials, where $x_i$ are binary vectors of length $2^m$ and the exponents $r_i$ are binary. The degree 0 monomial is a vector $\mathbf{1}$ of $2^m$ ones; e.g., when $m = 3$ then this monomial is (11111111). The degree 1 monomial $x_1$ is a vector of $2^{m-1}$ ones followed by $2^{m-1}$ zeros. The monomial associated with $x_2$ is a vector of $2^{m-2}$ ones followed by $2^{m-2}$ zeros followed by $2^{m-2}$ ones followed by $2^{m-2}$ zeros. In general, the vector associated with the monomial $x_i$ consists of $2^{m-i}$ ones followed by $2^{m-i}$ zeros repeated until a vector of length $2^m$ is obtained. For example, when $m = 3$ the vector associated with the monomial $x_1 x_2 x_3$ is

$$(11110000) \otimes (11001100) \otimes (10101010) = (10000000)$$

with "$\otimes$" the regular bitwise `XOR` operation. The scalar product of two binary vectors is defined as the sum of the product of individual components; for example,

$$(11110000) \cdot (11001100) = 1 + 1 + 0 + 0 + 0 + 0 + 0 + 0 = 0.$$

An $r$-th order Reed-Muller code $R(r, m)$ is the set of all binary strings of length $n = 2^m$ associated with the Boolean polynomials $p = x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n}$ of degree at most $r$. For example, the 0-th order code $R(0, m)$ is a repetition of strings of zeros or ones of length $2^m$; the $m$-th order code $R(m, m)$ consists of all binary strings of length $2^m$.

The generator matrix of an $R(r, m)$ with $r > 1$ code is formed by adding $\binom{m}{r}$ rows to the generator matrix of an $R(r-1, m)$ code. For example, the generator matrix of an $R(1, 3)$ code is

$$\mathrm{G}_{R(1,3)} = \begin{pmatrix} \mathbf{1} \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

We add the rows $x_1 \otimes x_2$, $x_1 \otimes x_3$, and $x_2 \otimes x_3$ to $\mathrm{G}_{R(1,3)}$ to obtain the generator matrix of an $R(2, 3)$ code,

$$\mathrm{G}_{R(2,3)} = \begin{pmatrix} \mathbf{1} \\ x_1 \\ x_2 \\ x_3 \\ x_1 \otimes x_2 \\ x_1 \otimes x_3 \\ x_2 \otimes x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

The generator matrix of an $R(2, 4)$ code is

$$\mathrm{G}_{R(2,4)} = \begin{pmatrix} \mathbf{1} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_1 \otimes x_2 \\ x_1 \otimes x_3 \\ x_1 \otimes x_4 \\ x_2 \otimes x_3 \\ x_2 \otimes x_4 \\ x_3 \otimes x_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

The <u>set of characteristic vectors</u> of any row of the generator matrix is the set of all monomials in $x_i$ and $\bar{x}_i$, the bitwise reverse of $x_i$, that are not in the monomial associated with that row. For example, the characteristic vectors of the row before the last of the generator matrix of the $R(2, 4)$ code, correspond to the vector $x_2 \otimes x_4$ and are $x_1 \otimes x_3$, $\bar{x}_1 \otimes x_3$, $x_1 \otimes \bar{x}_3$, and $\bar{x}_1 \otimes \bar{x}_3$. The scalar product of the vectors in this set with all the rows of the generator matrix of the code, except the row they belong to, is equal to zero.

The decoding algorithm for Reed-Muller codes described below is not very efficient, but is straightforward; it allows us to extract the message $\mu$ from the $n$-tuple received, $v$, and to determine $c$, the original codeword sent.

- Step 1. Compute the $2^{r-m}$ characteristic vectors $\kappa_i^j$ for each row of the generator matrix $\mathrm{G}_{R(m,r)}$ except the first row.
- Step 2. Compute the scalar products of $v$ with all the characteristic vectors $\kappa_i^j$ for rows $g_i, i > 1$ of the generator matrix. Determine the binary indicator $f_i$ for each row, starting with the bottom row; if the majority of the scalar products $v \cdot \kappa_i^j$ are equal to zero then $f_i = 0$, else $f_i = 1$. For example, the rows for the $R(1,3)$ code are: $g_1 = x_1, g_2 = x_2$, and $g_3 = x_3$.
- Step 3. Compute the binary vector $\nu$ as follows: multiply each $f_i$ with the corresponding row $g_i$ of the generator matrix and then add them together, $\nu = \sum_i f_i g_i$. For example, $\nu = f_3 x_3 + f_2 x_2 + f_1 x_1$ for the $R(1,3)$ code.
- Step 4. Add the result of Step 3 to the received $n$-tuple; if $\nu + v$ has a majority of zeros then assign a zero to the binary indicator of the first row of the generator matrix, $f_0 = 0$, else $f_0 = 1$. The message sent is $\mu = (f_0 f_1 f_2 .....)$. For example, $\mu = f_0 f_1 f_2 f_3$ for the $R(1,3)$ code.

**Example.** *Correction of one error for the $R(1,3)$ Reed-Muller code.* The distance of this code is $d = 2^{r-m} = 2^{3-1} = 4$; since $d \geq 3$ the code can correct a single error in $v$. The message sent is $\mu = (1110)$; it is encoded as the codeword $c \in R(1,3)$ by multiplying the message vector with the generator matrix of the code:

$$c = \mu \mathrm{G}_{R(1,3)} = \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

The message is affected by an error in the second bit and the 8-tuple received is $w = 10000011$.

Let us now apply the algorithm outlined above:
We compute first the characteristic set of the vector $x_3 = 10101010$ corresponding to the last row of the generator matrix; this set includes $x_1 \otimes x_2$, $x_1 \otimes \bar{x}_2$, $\bar{x}_1 \otimes x_2$, and $\bar{x}_1 \otimes \bar{x}_2$. Recall that $x_1 = (11110000)$ thus, $\bar{x}_1 = (00001111)$; similarly $x_2 = (11001100)$, so $\bar{x}_2 = (00110011)$. Therefore, $x_1 \otimes x_2 = (11000000)$, $x1 \otimes \bar{x}_2 = (00110000)$, $\bar{x}_1 \otimes x_2 = (00001100)$, and $\bar{x}_1 \otimes \bar{x}_2 = (00000011)$. The scalar products of these vectors with $v$ are:
$(11000000) \cdot (10000011) = 1$; $(00110000) \cdot (10000011) = 0$;
$(00001100) \cdot (10000011) = 0$; $(00000011) \cdot (10000011) = 0$.
The majority logic leads us to conclude that $f_3 = 0$. We continue with the third row of the matrix, $x_2 = (11001100)$. The characteristic vectors are $x_1 \otimes x_3$, $x_1 \otimes \bar{x}_3$, $\bar{x}_1 \otimes x_3$, and $\bar{x}_1 \otimes \bar{x}_3$ These vectors are (10100000), (01010000), (00001010), and (00000101), respectively. The scalar products of these vectors with $v$ are:
$(10100000) \cdot (10000011) = 1$; $(01010000) \cdot (10000011) = 0$;

$(00001010) \cdot (10000011) = 1;\ (00000101) \cdot (10000011) = 1.$
We conclude that $f_2 = 1$. The second row of the matrix is $x_1 = (11110000)$. The characteristic vectors are $x_2 \otimes x_3$, $x_2 \otimes \bar{x}_3$, $\bar{x}_2 \otimes x_3$, and $\bar{x}_2 \otimes \bar{x}_3$. These vectors are (10001000), (00100010), (00001010), and (00010001), respectively. The scalar products of these vectors with $v$ are:
$(10001000) \cdot (10000011) = 1;\ (00100010) \cdot (10000011) = 1;$
$(01000100) \cdot (10000011) = 0;\ (00010001) \cdot (10000011) = 1.$
We conclude that $f_1 = 1$. Now we construct

$$\nu = f_3x_3 + f_2x_2 + f_1x_1 = 0 \cdot (10101010) + 1 \cdot (11001100) + 1 \cdot (11110000) = (00111100)$$

and then compute

$$\nu + w = (00111100) + (10000011) = (10111111).$$

This string has more ones than zeros; this implies that the coefficient of the first row $f_0$ of the generator matrix is a 1, thus the message is

$$\mu = (f_0f_1f_2f_3) = (1110).$$

The zero in the second position of $(\nu + w)$ indicates an error, namely a zero has been received in error instead of an one.

We note that a Reed-Muller code $R(r, m)$ exists for any integers $m \geq 0$ and $0 \leq r \leq m$; the code $R(r, m)$ is a binary $[n, k, d]$ code with codewords of length $n = 2^m$, distance $d = 2^{r-m}$, and $k(r, m) = k(r, m-1) + k(r-1, m-1)$. For example, $R(3, 4)$ and $R(3, 5)$ are $[16, 15, 2]$ and $[32, 26, 4]$ Hamming codes, respectively, while $R(0, 1)$, $R(1, 3)$, and $R(2, 5)$ are $[2, 1, 2]$, $[8, 4, 4]$, and $[32, 16, 8]$ self-dual codes, respectively. The dual code of $R(r, m)$ is the code $R(m - r - 1, m)$. The majority logic decoding used for Reed-Muller codes computes several checksums for each received codeword element. To decode an $r$-th order code, we have to decode iteratively $r + 1$ times before we identify the original codeword.

We continue our discussion with one of the most important classes of linear codes, cyclic codes.

## 4.11 Cyclic Codes

There is a one-to-one mapping between vectors from $V_n$, an $n$-dimensional vector space over the finite field $F$ and polynomials of degree $(n - 1)$ with coefficients from the same field $F$. This observation is the basis of another formalism to describe a class of linear codes, the cyclic codes, a formalism based on algebras of polynomials over finite fields.

A cyclic code is characterized by a generator polynomial; encoding and decoding of cyclic codes reduces to algebraic manipulation of polynomials. We start our discussion of cyclic codes with several definitions and then discuss some properties of polynomials with coefficients from a finite field.

Given the vector $c = (c_0c_1 \ldots c_{n-2}c_{n-1}) \in V_n$, a <u>cyclic shift</u> is a transformation $c \mapsto c'$ with $c' = (c_{n-1}c_0c_1 \ldots c_{n-2}) \in V_n$.

We wish to establish a correspondence between vector spaces and sets of polynomials. Such a correspondence will facilitate the construction of codes with an algebraic structure and will allow us to express properties of linear codes, as well as algorithms for error detection and error correction in terms of the algebra of polynomials. We associate a codeword, a vector $c \in V_n$ with a polynomial of degree $(n-1)$ as follows: the coefficients of the polynomial terms, from low to high order, are the vector elements

$$c = (c_0 c_1 \ldots c_{n-1}) \in V_n \quad \Longleftrightarrow \quad c(x) = c_0 + c_1 x + c_2 x^2 + \ldots c_{n-1} x^{n-1}.$$

Let us now introduce cyclic codes. First, we define the concept of a *cyclic subspace of a vector space* as a subspace closed to cyclic shifts.

A cyclic subspace $\mathcal{C}$ of an $n$-dimensional vector space $V_n$ over the field $F$ is a set of vectors $\{c\} \in \mathcal{C}$ with the property

$$c = (c_0 c_1 c_2 \ldots c_{n-2} c_{n-1}) \in \mathcal{C} \quad \Longrightarrow \quad c' = (c_{n-1} c_0 c_1 \ldots c_{n-2}) \in \mathcal{C}.$$

A linear code $\mathcal{C}$ is a cyclic code if $\mathcal{C}$ is a cyclic subspace of an $n$-dimensional vector space $V_n$ over the field $F$.

Consider again the case $f(x) = x^n - 1$ and the finite ring of polynomials modulo $f(x)$, $\mathbb{F}[x]/f(x)$. It is easy to prove the following proposition.

**Proposition.** *Multiplication by $x$ of a polynomial $c(x)$ in the finite ring of polynomials modulo $f(x)$, $\mathbb{F}[x]/f(x)$, is equivalent with a cyclic shift of the corresponding vector $c = (c_0, c_1, c_2, \ldots, c_{n-2} c_{n-1})$ into $c' = (c_{n-1}, c_0, c_1, \ldots, c_{n-2})$ with $c, c' \in V_n$.*

*Proof:* Let $c(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_{n-2} x^{n-2} + c_{n-1} x^{n-1} \mod f(x)$. Then

$$c'(x) = x \cdot c(x) = c_0 x + c_1 x^2 + c_2 x^3 + \ldots + \ldots + c_{n-2} x^{n-1} + c_{n-1} x^n \mod f(x)$$

But $x^n = 1 \mod f(x)$ thus,

$$c'(x) = c_{n-1} + c_0 x + c_1 x^2 + c_2 x^3 + \ldots + \ldots + c_{n-2} x^{n-1} \mod f(x).$$

**Example.** *Let $\mathbb{Z}_2$ be the binary field and $f(x) = 1 + x^4$. Then*
$\mathbb{Z}_2[x]/(x^4+1) = \{[0], [1], [x], [1+x], [x^2], [1+x^2], [x+x^2], [1+x+x^2],$
$[x^3], [1+x^3], [1+x+x^3], [1+x^2+x^3], [1+x+x^2+x^3], [x+x^3], [x^2+x^3], [x+x^2+x^3]\}$

The corresponding vectors in $V_4(GF(2))$ are:

$$\begin{array}{lllllllll} \mathcal{C} = \{ & 0000, & 1000, & 0100, & 1100, & 0010, & 1010, & 0110, & 1110, \\ & 0001, & 1001, & 1101, & 1011, & 1111, & 0101, & 0011, & 0111 \ \}. \end{array}$$

$\mathcal{C}$ is a cyclic subspace of $V_4(GF(2))$, it is closed to cyclic shifts. Indeed, if we apply a cyclic shift we get:

$0000 \mapsto 0000$, $1000 \mapsto 0100$, $0100 \mapsto 0010$, $1100 \mapsto 0110$, $0010 \mapsto 0001$, $1010 \mapsto 0101$,
$0110 \mapsto 0011$, $1110 \mapsto 0111$, $0001 \mapsto 1000$, $1001 \mapsto 1100$, $1101 \mapsto 1110$, $1011 \mapsto 1101$,
$1111 \mapsto 1111$, $0101 \mapsto 1010$, $0011 \mapsto 1001$, $0111 \mapsto 1011$

We notice that:

$$x \cdot [0] = [0], \quad x \cdot [1] = [x], \quad \ldots \quad , x \cdot [x^2 + x^3] = [1 + x^3], \quad x \cdot [x + x^2 + x^3] = [1 + x^2 + x^3].$$

Therefore, multiplication with $x$ of the polynomials in $\mathbb{Z}_2[x]/(x^4 + 1)$ leads to the set:

$\{[0], [1], [x], [1 + x], [x^2], [1 + x^2], [x + x^2], [1 + x + x^2],$
$[x^3], [1 + x^3], [1 + x + x^3], [1 + x^2 + x^3], [1 + x + x^2 + x^3], [x + x^3], [x^2 + x^3], [x + x^2 + x^3]\}$

We can easily verify that the correspondence between the vectors subject to a cyclic shift and polynomials multiplied by $x$ is preserved.

Now we discuss a very important proposition which relates cyclic subspaces of an $n$-dimensional vector space with ideals of a ring of polynomials. We show that there is a one-to-one correspondence between cyclic subspaces of the vector space $V_n$ and ideals of polynomials generated by a polynomial $g(x)$ in a finite ring of polynomials over $F$ modulo the non-zero polynomial $f(x) = x^n - 1$ when $g(x)$ divides $f(x)$.

**Proposition.** *There is an one-to-one correspondence between the cyclic subspaces of an $n$-dimensional vector space $V_n$ over the field of $F = GF(q)$ and the monic polynomials*[35] $g(x) \in \mathbb{F}[x]$ *which divide $f(x) = x^n - 1$ with $\mathbb{F}[x]$ a ring of polynomials. If*

$$f(x) = \prod_{i=1}^{q} g_i^{a_i}(x)$$

*with $a_i$ positive integers and $g_i^{a_i}(x)$, $1 \leq i \leq q$, distinct irreducible monic polynomials, then $V_n$ contains*

$$Q = \prod_{i=1}^{q} (a_i + 1)$$

*cyclic subspaces. If $g(x)$ is a monic polynomial of degree $(n - k)$ and it divides $f(x) = x^n - 1$, then $g(x)$ is the generator polynomial of a cyclic sub-space of $V_n$. The dimension of this cyclic subspace is equal to $k$.*

This proposition follows from several propositions whose proofs are left to the reader[36]. First, let $S$ be a subspace of an $n$-dimensional vector space $V_n$ over the field F, $S \subset V_n$. Let $R$ be the ring of polynomials associated with $V_n$ and let $I \in R$ be the set of polynomials corresponding to $S$. We can show that $S$ is a cyclic subspace of $V_n$ if and only if $I$ is an ideal in $R$, Figure 101.

[35] The coefficient of the highest order term of a monic polynomial is 1: $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + 1$.
[36] See the exercises at the end of this chapter.

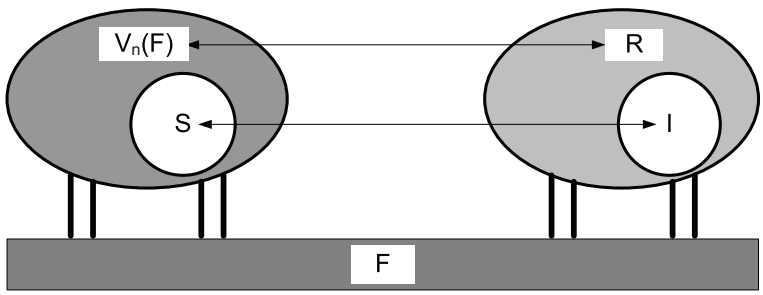



Figure 101: There is a one-to-one correspondence between the vector space $V_n(F)$, an $q^n$-dimensional vector space over the finite field $F$, and $R$, the ring of polynomials with coefficients from the finite field $F$, modulo the polynomial $f(x) = x^n - 1$. There is also a one-to-one correspondence between $S$, a cyclic subspace of $V_n$, and $I$, an ideal of the ring $R$ generated by the polynomial $g(x) \in R$. Moreover, $f(x)$ is a multiple of $g(x)$, $f(x) = g(x) \cdot h(x)$.

If $f(x) = x^n - 1$, $R$ is the ring of equivalence classes of polynomials modulo $f(x)$ with coefficients in the field $F$, $R = \mathbb{F}[x]/(f(x))$, and $g(x)$ is a monic polynomial which divides $f(x)$, then we can show that $g(x)$ is the generator of the ideal of polynomials $I = m(x) \cdot g(x),\ m(x) \in R$. We can also show that there is a one-to-one correspondence between the cyclic subspaces of $V_n$ and the monic polynomials $g(x)$ which divide $f(x) = x^n - 1$.

**Example.** *We consider several values of $n$, namely $5, 10, 15, 20$, and $25$. Show the generator polynomials for the cyclic subspaces of $V_n(GF(2))$ and calculate their dimension.*

- $n = 5$. Then:

  $x^5 - 1 = (1+x) \cdot (1+x+x^2+x^3+x^4)$; $\quad g_1(x) = 1+x \quad$ and $\quad g_2(x) = 1+x+x^2+x^3+x^4$.

  The number of cyclic subspaces is $Q = 2 \times 2 = 4$.

  The cyclic subspace generated by $g_1$ has dimension 4 as

  $n - k_1 = 1 \quad \rightarrow \quad k_1 = n - 1 = 4$.

  The cyclic subspace generated by $g_2$ has dimension 1 as

  $n - k_2 = 4 \quad \rightarrow \quad k_2 = n - 4 = 1$.

- $n = 10$. Then:

  $x^{10} - 1 = (1+x)^2 \cdot (1+x+x^2+x^3+x^4)^2$; $\quad g_1(x) = 1+x \quad$ and $\quad g_2(x) = 1+x+x^2+x^3+x^4$.

  The number of cyclic subspaces is: $Q = 3 \cdot 3 = 9$.

  The cyclic subspace generated by $g_1$ has dimension 9 as

  $n - k_1 = 1 \quad \rightarrow \quad k_1 = n - 1 = 9$.

  The cyclic subspace generated by $g_2$ has dimension 6 as

  $n - k_2 = 4 \quad \rightarrow \quad k_2 = n - 4 = 6$.

- $n = 15$. Then:

  $x^{15} - 1 = (1 + x) \cdot (1 + x + x^2) \cdot (1 + x + x^2 + x^3 + x^4) \cdot (1 + x + x^4) \cdot (1 + x^3 + x^4)$

  $g_1(x) = 1 + x, \quad g_2(x) = 1 + x + x^2, \quad g_3(x) = 1 + x + x^2 + x^3 + x^4$

  $g_4(x) = 1 + x + x^4, \quad g_5(x) = 1 + x^3 + x^4.$

  The number of cyclic subspaces is: $Q = 2 \times 2 \times 2 \times 2 \times 2 = 32$.

  The cyclic subspace generated by $g_1$ has dimension 14 as

  $n - k_1 = 1 \quad \rightarrow \quad k_1 = n - 1 = 14.$

  The cyclic subspace generated by $g_2$ has dimension 13 as

  $n - k_2 = 2 \quad \rightarrow \quad k_2 = n - 2 = 13.$

  The cyclic subspaces generated by $g_3, g_4$ and $g_5$ have dimension 11 as

  $n - k_3 = 4 \quad \rightarrow \quad k_3 = n - 4 = 11.$

- $n = 20$. Then:

  $x^{20} - 1 = (1+x)^4 \cdot (1+x+x^2+x^3+x^4)^4; \quad g_1(x) = 1+x \quad \text{and} \quad g_2(x) = 1+x+x^2+x^3+x^4.$

  The number of cyclic subspaces is: $Q = 5 \times 5 = 25$.

  The cyclic subspace generated by $g_1$ has dimension 19 as

  $n - k_1 = 1 \quad \rightarrow \quad k_1 = n - 1 = 19.$

  The cyclic subspace generated by $g_2$ has dimension 16 as

  $n - k_2 = 4 \quad \rightarrow \quad k_2 = n - 4 = 16.$

- $n = 25$. Then:

  $x^{25} - 1 = (1 + x) \cdot (1 + x + x^2 + x^3 + x^4) \cdot (1 + x^5 + x^{10} + x^{15} + x^{20}).$

  $g_1(x) = 1 + x, \quad g_2(x) = 1 + x + x^2 + x^3 + x^4, \quad \text{and} \quad g_3(x) = 1 + x^5 + x^{10} + x^{15} + x^{20}.$

  The number of cyclic subspaces is: $Q = 2 \times 2 \times 2 = 8$.

  The cyclic subspace generated by $g_1$ has dimension 24 as

  $n - k_1 = 1 \quad \rightarrow \quad k_1 = n - 1 = 24.$

  The cyclic subspace generated by $g_2$ has dimension 21 as

  $n - k_2 = 4 \quad \rightarrow \quad k_2 = n - 4 = 6.$

  The cyclic subspace generated by $g_3$ has dimension 5 as

  $n - k_3 = 20 \quad \rightarrow \quad k_2 = n - 20 = 5.$

In summary, a cyclic code is an ideal $I$ in the ring $R$ of polynomials modulo $f(x) = x^n - 1$ over a finite field, $GF(q)$; $f(x)$ is a product of irreducible polynomials $f(x) = \Pi_i(g_i(x))$. An $[n, k]$ cyclic code is generated by a polynomial $g(x) \in I$ of degree $(n - k)$; all codewords of $\mathcal{C}_g$ are multiples of the generator polynomial $g(x)$. The generator polynomial of a code $\mathcal{C}_g$ may be irreducible, $g(x) = g_i(x)$, or a product of several of several irreducible polynomials $g_i(x)$.

## 4.12 Encoding and Decoding Cyclic Codes

In this section we recast familiar concepts from the theory of linear codes such as generator and parity check matrices and syndrome in terms of algebras of polynomials and discuss encoding and decoding cyclic codes. Without loss of generality we only consider vector spaces over the finite field, $GF(2)$, a fancy way of expressing the fact that the coefficients of polynomials are only zeros and ones. The $[n, k]$ linear code $\mathcal{C}$ is characterized by the generator polynomial $g(x)$ of degree $(n-k)$:

$$g(x) = g_0 + g_1 x + g_2 x^2 + \ldots + g_{n-k-1} x^{n-k-1} + g_{n-k} x^{n-k}, \quad g_i \in \{0, 1\}.$$

All operations are performed in an algebra modulo of the polynomial $f(x)$ with

$$f(x) = x^n - 1$$

subject to the condition that $g(x)$ divides $f(x)$

$$f(x) = x^n - 1 = g(x) \cdot h(x).$$

A message of $k$ bits is now represented by $m(x)$, a polynomial of degree at most $(k-1)$:

$$m(x) = m_0 + m_1 x + \ldots + m_{k-2} x^{k-2} + m_{k-1} x^{k-1}, \quad m_i \in \{0, 1\}.$$

The message $m(x)$ is encoded as the polynomial $c(x)$:

$$c(x) = g(x) \cdot m(x).$$

The degree of the polynomial $c(x)$ cannot be larger than $(n-1)$:

$$\deg[c(x)] = \deg[g(x)] + \deg[m(x)] \leq (n-k) + (k-1) = n-1.$$

Thus, when we encode the message polynomial $m(x)$ into the codeword polynomial $c(x)$ we do not need to perform a reduction modulo $f(x)$. We can anticipate a very important property of cyclic codes, a direct consequence of the way we encode a message into a polynomial: all codewords are represented by polynomials multiple of the generator polynomial of the code. Thus, *to decide if a received polynomial $r(x)$ belongs to the code $\mathcal{C}$ we divide it by the generator polynomial of the code, $g(x)$.*

We now revisit concepts such as the generator and the parity matrix of a code, dual code, and encoding/decoding algorithms using polynomial representation of codewords and the properties of cyclic subspaces of a vector space.

A <u>generator matrix</u> $\mathrm{G}_g$ of the code $\mathcal{C}[n, k]$ is a $k \times n$ matrix having as elements the coefficients of the generator polynomial $g(x)$,

$$\mathrm{G}_g = \begin{pmatrix} g(x) \\ x \cdot g(x) \\ x^2 \cdot g(x) \\ \vdots \\ x^{k-1} \cdot g(x) \end{pmatrix} = \begin{pmatrix} g_0 & g_1 & \ldots & g_{n-k-1} & g_{n-k} & 0 & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_{n-k-1} & g_{n-k} & \ldots & 0 \\ 0 & 0 & g_0 & g_1 & \ldots & g_{n-k-1} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 \ldots & \ldots & g_0 & \ldots & g_{n-k-1} & g_{n-k} \end{pmatrix}.$$

**Example.** *Consider the case $n = 10$ and express*

$$x^{10} - 1 = (1 + x)^2 \cdot (1 + x + x^2 + x^3 + x^4)^2.$$

If we choose $g(x) = 1 + x + x^2 + x^3 + x^4$ as the generator polynomial of the code it follows that $\deg[g(x)] = n - k = 4$, thus $k = 10 - 4 = 6$. A generator matrix of the code is

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The message $m(x) = 1 + x + x^5$ corresponding to the 5-tuple (110001) is encoded as

$$\begin{aligned} c(x) &= m(x) \cdot g(x) \\ &= (1 + x + x^5) \cdot (1 + x + x^2 + x^3 + x^4) \\ &= (1 + x + x^2 + x^3 + x^4) + (x + x^2 + x^3 + x^4 + x^5) + (x^5 + x^6 + x^7 + x^8 + x^9) \\ &= 1 + x^6 + x^7 + x^8 + x^9 \end{aligned}.$$

Recall that $n = 10$ and the codeword $c(x)$ corresponds to the $n$-tuple (1000001111), the same vector obtained by multiplying the message vector with the generator matrix:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

**An alternative method to construct the generator matrix.** Consider a cyclic code with the generator polynomial $g(x)$; when we divide $x^{n-k+i}$, $0 \leq i \leq k-1$ by the generator polynomial we obtain a quotient $q^{(i)}(x)$ and a remainder $r^{(i)}(x)$:

$$x^{n-k+i} = g(x) \cdot q^{(i)}(x) + r^{(i)}(x), \quad 0 \leq i \leq k-1.$$

The degree of the remainder satisfies the following condition:

$$\deg[r^{(i)}(x)] < \deg[g(x)] = n - k, \quad 0 \leq i \leq k-1.$$

Clearly,

$$x^{n-k+i} - r^{(i)}(x) = g(x) \cdot q^{(i)}(x), \quad 0 \leq i \leq k-1.$$

Consider a $k \times (n-k)$ matrix $R_g$ with rows corresponding to $[-r^{(i)}(x)]$:

$$R_g = \begin{pmatrix} -r^0(x) \\ -r^1(x) \\ \vdots \\ -r^{k-1}(x) \end{pmatrix}.$$

Then a generator matrix of the code is $\mathrm{G}_g = [R_g \ I_k]$; we'll show that a parity check matrix of the code is $H_g = [I_{n-k} \ -R_g^T]$ with $I_k$ a $k \times k$ identity matrix.

**Example.** *Alternative construction of the generator matrix of a cyclic code when $n = 10$, $k = 6$, and $g(x) = 1 + x + x^2 + x^3 + x^4$.* The following table gives the polynomials $r^i(x)$ necessary for the construction of the matrix $R$:

| $i$ | $x^{n-k+i}$ | $r^{(i)}(x)$ |
|---|---|---|
| 0 | $x^4 = [(1) \cdot (1 + x + x^2 + x^3 + x^4)] + 1 + x + x^2 + x^3$ | $1 + x + x^2 + x^3$ |
| 1 | $x^5 = [(x + 1) \cdot (1 + x + x^2 + x^3 + x^4)] + 1$ | $1$ |
| 2 | $x^6 = [(x^2 + x) \cdot (1 + x + x^2 + x^3 + x^4)] + x$ | $x$ |
| 3 | $x^7 = [(x^3 + x^2) \cdot (1 + x + x^2 + x^3 + x^4)] + x^2 + x^3$ | $x^2 + x^3$ |
| 4 | $x^8 = [(x^4 + x^3) \cdot (1 + x + x^2 + x^3 + x^4)] + x^3$ | $x^3$ |
| 5 | $x^9 = [(x^5 + x^4 + 1) \cdot (1 + x + x^2 + x^3 + x^4)] + 1 + x + x^2 + x^3$ | $1 + x + x^2 + x^3$ |

Thus,

$$R = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

A generator matrix of the code is $\mathrm{G} = RI_k$

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The message $m(x) = 1 + x + x^5$ which corresponds to the 5-tuple (110001) is encoded as

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that the information symbols appear as the last $k$ bits of the codeword; this is a general property of this method of constructing the generator matrix of a cyclic code.

Given a polynomial of degree $m$, $h(x) = \sum_{k=0}^{m} h_k x^k,\ h_m \neq 0$, its reciprocal, $\bar{h}(x)$, is the polynomial defined by

$$\bar{h}(x) = \sum_{k=0}^{m} h_{m-k} x^k.$$

The vectors corresponding to the two polynomials, $h(x)$ and $\bar{h}(x)$ are respectively:

$$h = (h_0 h_1 \ldots h_{m-1} h_m) \quad \text{and} \quad \bar{h} = (h_m h_{m-1} \ldots h_1 h_0).$$

For example, when $m = 4$:

$$\begin{array}{llll} h(x) = 1 + x + x^4 & \mapsto & \bar{h}(x) = 1 + x^3 + x^4 & \text{or} \quad (11001) \mapsto (10011), \\ h(x) = x^2 + x^3 + x^4 & \mapsto & \bar{h}(x) = 1 + x + x^2 & \text{or} \quad (00111) \mapsto (11100). \end{array}$$

**The equivalent and the dual of a cyclic code with generator polynomial $g(x)$.** Recall that two codes, $\mathcal{C}$ with generator matrix G and $\mathcal{C}'$ with generator matrix G$'$, are said to be equivalent if $\mathrm{G}' = \mathrm{G} \cdot P$ with $P$ a permutation matrix. Recall also that a linear code and its dual are orthogonal subspaces of a vector space and the generator matrix of a code is the parity check matrix of the dual code.

The polynomial $h(x)$ defined by the expression $f(x) = x^n - 1 = g(x) \cdot h(x)$ is called the *parity-check polynomial* of $\mathcal{C}$ and has several properties: (i) it is a monic polynomial of degree $k$, and (ii) it is the generator of a cyclic code $\mathcal{C}_h$. Indeed, any monic polynomial which divides $f(x)$ generates an ideal of polynomials, thus a cyclic code. The codewords of the two cyclic codes generated by $g(x)$ and $h(x)$, $\mathcal{C}_g$ and $\mathcal{C}_h$, respectively, have the following property:

$$c_g(x) \cdot c_h(x) \equiv 0 \mod f(x) \quad \forall c_g(x) \in \mathcal{C}_g, \quad \forall c_h(x) \in \mathcal{C}_h.$$

A codeword of a cyclic code is represented by a polynomial multiple of the generator polynomial of the code. Thus:

$$c_g(x) = g(x) \cdot q_g(x) \mod f(x) \qquad \text{and} \qquad c_h(x) = h(x) \cdot q_h(x) \mod f(x).$$

Thus:

$$c_g(x) \cdot c_h(x) = [g(x) \cdot q_g(x)] \cdot [h(x) \cdot q_h(x)] = [g(x) \cdot h(x)] \cdot [q_g(x) \cdot q_h(x)] \equiv 0 \mod f(x)$$

as $g(x) \cdot h(x) = 0 \mod f(x)$. The equation $c_g(x) \cdot c_h(x) \equiv 0 \mod f(x)$ does not imply that the two codes are orthogonal; indeed, $\mathcal{C}_h \neq \mathcal{C}_g^{\perp}$, the two codes, $\mathcal{C}_h$ and $\mathcal{C}_g^{\perp}$, are not each other's dual.

Given that $f(x) = x^n - 1 = g(x) \cdot h(x)$ we define the codes $\mathcal{C}_g, \mathcal{C}_h, \mathcal{C}_g^{\perp}$, and $\mathcal{C}_{\bar{h}}$ as follows:

- $\mathcal{C}_g$ is the cyclic code generated by the monic polynomial $g(x)$ of degree $k$; $\mathrm{G}_g$, the generator matrix of $\mathcal{C}_g$, is a $(k \times n)$ matrix and $\mathrm{H}_g$, the parity-check matrix of $\mathcal{C}_g$, is a $(n-k) \times n$ matrix.

- $\mathcal{C}_h$ is the cyclic code generated by the monic polynomial $h(x)$ of degree $k$; the generator matrix of $\mathcal{C}_h$ is an $(n-k) \times n$ matrix denoted by $\mathrm{G}_h$.
- $\mathcal{C}_g^{\perp}$ is the dual of $\mathcal{C}_g$; $\mathrm{G}_g^{\perp}$, the generator matrix of $\mathcal{C}_g^{\perp}$, is an $(n-k) \times n$ matrix.
- $\mathcal{C}_{\bar{h}}$ is the code generated by $\bar{h}(x)$, the reciprocal of the polynomial $h(x)$; the generator matrix of $\mathcal{C}_{\bar{h}}$ is an $(n-k) \times n$ matrix denoted by $\mathrm{G}_{\bar{h}}$.

**Proposition.** *The codes $\mathcal{C}_g, \mathcal{C}_h, \mathcal{C}_g^{\perp}$, and $\mathcal{C}_{\bar{h}}$ have the following properties:*
*(i) $\mathcal{C}_g^{\perp}$ and $\mathcal{C}_h$ are equivalent codes; the generator matrices of the two codes satisfy the relation:*

$$\mathrm{G}_h = \mathrm{G}_g^{\perp} \cdot P.$$

*(ii) The dual of the code $\mathcal{C}_g$ is the code generated by $\bar{h}(x)$, the reciprocal of $h(x)$: $\mathcal{C}_g^{\perp} = \mathcal{C}_{\bar{h}}$; the parity check matrix of $\mathcal{C}_g$ is the generator matrix of $\mathcal{C}_{\bar{h}}$:*

$$\mathrm{H}_g = \mathrm{G}_{\bar{h}}.$$

The generator matrices of the cyclic code $\mathcal{C}_h$ and $\mathcal{C}_{\bar{h}}$ are:

$$\mathrm{G}_h = \begin{pmatrix} h(x) \\ x \cdot h(x) \\ x^2 \cdot h(x) \\ \vdots \\ x^{n-k-1} \cdot h(x) \end{pmatrix} \quad \text{and} \quad \mathrm{G}_{\bar{h}} = \begin{pmatrix} \bar{h}(x) \\ x \cdot \bar{h}(x) \\ x^2 \cdot \bar{h}(x) \\ \vdots \\ x^{n-k-1} \cdot \bar{h}(x) \end{pmatrix}.$$

It is easy to see that the parity check matrix of the code $\mathcal{C}_g$ is $\mathrm{H}_g = \mathrm{G}_{\bar{h}} = [I_{n-k} \ -R_g^T]$ with $R_g^T$ an $(n-k) \times k$ matrix with the rows corresponding to $[-r^i(x)]$, the remainder when we divide $x^{n-k+i}$ by $g(x)$.

Let two arbitrary codewords $c_g(x) \in \mathcal{C}_g$ and $c_h(x) \in \mathcal{C}_h$ and the corresponding vectors be:
$c_g(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \quad \mapsto \quad c_g = (a_0 a_1 \ldots a_{n-1})$
$c_h(x) = b_0 + b_1 x + \ldots + b_{n-1} x^{n-1} \quad \mapsto \quad c_h = (b_0 b_1 \ldots b_{n-1})$.
Then,

$$c_g(x) \cdot c_h(x) = (\sum_{i=0}^{n-1} a_i x^i) \cdot (\sum_{i=0}^{n-1} b_i x^i) \equiv \sum_{i=0}^{n-1} d_i x^i \mod (x^n - 1).$$

Call $\bar{c}_h = (b_{n-1} b_{n-2} \ldots b_1 b_0)$ the reciprocal of the vector $c_h = (b_0 b_1 \ldots b_{n-2} b_{n-1})$. Then the coefficient of $x^0$ can be expressed as the inner product $d_0 = c_g \cdot \bar{c}_h$:

$$d_0 = c_g \cdot \bar{c}_h = (a_0 a_1 \ldots a_{n-1}) \cdot (b_{n-1} b_{n-2} \ldots b_1 b_0) = a_0 b_{n-1} + a_1 b_{n-2} + a_2 b_{n-3} + \ldots + a_{n-1} b_0.$$

Now compute $d_k$, $1 \leq k \leq n-1$, and observe that $d_k$ becomes the constant term if we multiply the product $\sum_{i=0}^{n-1} d_i x^i$ by $x^{n-k}$. The multiplication corresponds to the following product of polynomials

$$c_g(x) \cdot \left[x^{n-k} \cdot c_h(x)\right].$$

Let $\bar{c}_h^{(k)}$ be the vector obtained by cyclically shifting $c_h(x)$, $k+1$ positions to the left, and then switching the last component of the vector with the first one. Then,

$$d_k = c_g \cdot \bar{c}_h^{(k)}.$$

**Example.** *The procedure to compute the coefficients $c_k$, $0 \leq k \leq n$.* Consider the case $n = 5$ and $f(x) = x^5 - 1$. Let:

$$c_g(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \quad \text{and} \quad c_h(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4.$$

The corresponding vectors are:

$$c_g = (a_0a_1a_2a_3a_4) \quad \text{and} \quad c_h = (b_0b_1b_2b_3b_4).$$

Then,

$$\begin{aligned} c_g(x) \cdot c_h(x) = (a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4) \cdot (b_0 + b_1x + b_2x^2 + b_3x^3 + b_4x^4) = \\ (a_0b_0 + a_1b_4 + a_2b_3 + a_3b_2 + a_4b_1) + (a_0b_1 + a_1b_0 + a_2b_4 + a_3b_3 + a_4b_2)x + \\ (a_0b_2 + a_1b_1 + a_2b_0 + a_3b_4 + a_4b_3)x^2 + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0 + a_4b_4)x^3 + \\ (a_0b_4 + a_1b_3 + a_2b_2 + a_3b_1 + a_4b_0)x^4 \mod f(x) = x^5 - 1. \end{aligned}$$

The coefficient of the constant term in this expression is given by the inner product of the vectors $c_g$ and $\bar{c}_h^{(1)}$:

$$c_g \cdot \bar{c}_h^{(1)} = (a_0a_1a_2a_3a_4) \cdot (b_0b_2b_3b_4b_1) = a_0b_0 + a_1b_2 + a_2b_3 + a_3b_4 + a_4b_1.$$

Indeed, $\bar{c}_h^{(1)}$ is obtained by a cyclic shift of $c_h = (b_0b_1b_2b_3b_4)$ one position to the left to obtain $(b_1b_2b_3b_4b_0)$ and then by switching the last with the first components leading to $(b_0b_2b_3b_4b_1)$.

Similarly, the coefficient of $x^2$ in this expression is the inner product of $c_g$ and $\bar{c}_h^{(3)}$:

$$c_g \cdot \bar{c}_h^{(3)} = (a_0a_1a_2a_3a_4) \cdot (b_2b_1b_0b_4b_3) = a_0b_2 + a_1b_1 + a_2b_0 + a_3b_4 + a_4b_3.$$

Now, $\bar{c}_h^{(3)}$ is obtained by a cyclic shift of $c_h$ three positions to the left to obtain $(b_3b_4b_0b_1b_2)$ and then by switching the last with the first components to obtain $(b_2b_1b_0b_4b_3)$.

Now we consider the general case:

$$c_g(x) \cdot c_h(x) \equiv 0 \mod (x^n - 1).$$

The coefficients of every power of $x$ in the product of polynomials must be zero. But these coefficients are precisely the scalars $d_k$ computed above. Now consider $d$ to be any cyclic shift of the vector obtained from $c_h$ by reversing its components. Then:

$$c_g \cdot d = 0$$

Thus, the code $\mathcal{C}_{\bar{h}}$ obtained by such cyclic shifts of $c_h$ followed by reversal of the components is the dual of $\mathcal{C}_g$. It is left as an exercise to prove that $\mathcal{C}_{\bar{h}}$ and $\mathcal{C}_h$ are equivalent codes.

**Example.** *Consider the polynomial* $f(x) = x^9 - 1 = (1+x)\cdot(1+x+x^2)\cdot(1+x^3+x^6)$ *and choose* $g(x) = 1+x^3+x^6$. In this case $n = 9$ and $\deg[g(x)] = n-k = 6$ thus, $k = 3$. It follows that $h(x) = (1+x)(1+x+x^2) = (1+x+x^2)+(x+x^2+x^3) = 1+x^3$. The vectors corresponding to $g(x)$ and $h(x)$ are: $g(x) = 1+x^3+x^6 \mapsto (100100100)$ and $h(x) = 1+x^3 \mapsto (100100000)$. The generator and the parity matrices of the code $\mathcal{C}_g$ are $k \times n$ and $(n-k)\times n$ matrices:

$$\mathrm{G}_g = \begin{pmatrix} g(x) \\ x\cdot g(x) \\ x^2\cdot g(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} = [I_3 A],$$

$$\mathrm{H}_g = [-A^T I_{n-k}] = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \mathrm{G}_g^{\perp}.$$

The generator matrix of the code $\mathcal{C}_h$ is an $(n-k)\times n = 6\times 9$ matrix

$$\mathrm{G}_h = \begin{pmatrix} h(x) \\ x\cdot h(x) \\ x^2\cdot h(x) \\ x^3\cdot h(x) \\ x^4\cdot h(x) \\ x^5\cdot h(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

As we know the parity check matrix $\mathrm{G}_h$ is the generator matrix of the dual code $\mathcal{C}^{\perp}$. The vector $h = (100100000)$ thus, $\bar{h} = (000001001)$ and $\mathrm{G}_{\bar{h}}$ is

$$\mathrm{G}_{\bar{h}} = \begin{pmatrix} \bar{h}(x) \\ x\cdot \bar{h}(x) \\ x^2\cdot \bar{h}(x) \\ x^3\cdot \bar{h}(x) \\ x^4\cdot \bar{h}(x) \\ x^5\cdot \bar{h}(x) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

**Example.** *Consider the* $[7,4]$ *cyclic code with generator polynomial* $g(x) = 1+x+x^3$. In this case $n = 7$ and $k = 3$ and $x^n - 1 = x^7 - 1 = g(x)\cdot h(x)$ with $h(x) = 1+x+x+x^4$. A generator matrix of the code is:

$$\mathrm{G}^a = \begin{pmatrix} g(x) \\ x\cdot g(x) \\ x^2\cdot g(x) \\ x^2\cdot g(x) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Now we compute $r^{(i)}(x) = r_0^i + r_1^i x + r_2^i x^2$ for $0 \le i \le k$, the remainders when we divide of $x^{n-k+i}$ by $g(x)$:

$$
\begin{array}{rll}
x^3 = & 1\cdot(1+x+x^3)+(1+x) & \Rightarrow\ r^{(0)}(x)=1+x \\
x^4 = & (x)\cdot(1+x+x^3)+(x+x^2) & \Rightarrow\ r^{(1)}(x)=x+x^2 \\
x^5 = & (1+x^2)\cdot(1+x+x^3)+(1+x+x^2) & \Rightarrow\ r^{(2)}(x)=1+x+x^2 \\
x^6 = & (1+x+x^3)\cdot(1+x+x^3)+(1+x^2) & \Rightarrow\ r^{(3)}(x)=1+x^2
\end{array}
$$

and construct another generator matrix $\mathrm{G}_g^b = [RI_4]$

$$
\mathrm{G}_g^b = \begin{pmatrix} r_0^{(0)} & r_1^{(0)} & r_2^{(0)} & 1 & 0 & 0 & 0 \\ r_0^{(1)} & r_1^{(1)} & r_2^{(1)} & 0 & 1 & 0 & 0 \\ r_0^{(2)} & r_1^{(2)} & r_2^{(2)} & 0 & 0 & 1 & 0 \\ r_0^{(3)} & r_1^{(3)} & r_2^{(3)} & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}
$$

To construct the parity check matrix $\mathrm{H}_g = [-A^T I_{n-k}]$ of the code we transform $\mathrm{G}_g^b$ to the canonical form $\mathrm{G}_g = [I_k A]$. The permutation matrix $P$ transforms the columns of $\mathrm{G}_g^b$ as follows $1 \mapsto 5$, $2 \mapsto 6$, $3 \mapsto 7$, $4 \mapsto 1$, $5 \mapsto 2$, $6 \mapsto 3$, and $7 \mapsto 4$. Then $\mathrm{G}_g = \mathrm{G}_g^b P$

$$
\mathrm{G}_g = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}
$$

Then

$$
\mathrm{G}_g^{\perp} = \mathrm{H}_g = [-A^T I_{n-k}] = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.
$$

Recall that $h(x) = h(x) = 1 + x + x + x^4$ and $\mathrm{G}_h$ is

$$
\mathrm{G}_h = \begin{pmatrix} h(x) \\ x\cdot h(x) \\ x^2\cdot h(x) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.
$$

It is easy to see that $\mathrm{G}_h = \mathrm{G}_g^{\perp} P_1$ with $P_1$ the permutation matrix exchanging the columns of $\mathrm{G}_g^{\perp}$ as follows: $1 \mapsto 2$, $2 \mapsto 4$, $3 \mapsto 3$, $4 \mapsto 5$, $5 \mapsto 1$, $6 \mapsto 6$, and $7 \mapsto 7$:

$$
\mathrm{G}_h = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.
$$

**Decoding cyclic codes.** We recast linear codes decoding strategies in terms of algebras of polynomials. Recall that the parity-check matrix H of an $[n, k]$ linear code $\mathcal{C}$ is used to compute the error syndrome $s$ given the $n$-tuple $v$: $s^T = \mathrm{H}v^T$. Once we compute the syndrome we are able to determine the error pattern $e$; there is a one-to-one correspondence between the two. Indeed, if the codeword $c \in \mathcal{C}$ is affected by the error $e$ and received as $v$ then

$$v = c + e \quad \Longrightarrow s^T = \mathrm{H}v^T = \mathrm{H}(c+e)^T = \mathrm{H}e^T \qquad \text{as} \qquad \mathrm{H}c^T = 0, \ \forall c \in \mathcal{C}.$$

The following proposition gives a simple algorithm for computing the syndrome polynomial. This algorithm requires the determination of the remainder at the division of polynomials and can be implemented very efficiently using a family of linear circuits called *feedback shift registers*.

**Proposition.** *The syndrome corresponding to the polynomial* $v(x) = v_0 + v_1 x + \ldots + v_{n-1}x^{n-1}$ *is the polynomial* $s(x)$ *satisfying the condition* $\deg[s(x)] \leq n-k-1$*, obtained as the remainder when* $v(x)$ *is divided by* $g(x)$*.*

$$v(x) = q(x)g(x) + s(x).$$

*Proof:* We have shown that a parity check matrix $\mathrm{H}_g$ of the cyclic code $\mathcal{C}_g$ with generator polynomial $g(x)$ is $\mathrm{H}_g = [I_{n-k} \ \ -R_g^T]$; it follows that the columns of $\mathrm{H}_g$ correspond to the following polynomials:

$$h_i(x) = \begin{cases} x^i & \text{if } \ 0 \leq i \leq n-k-1 \\ r^{(i-n+k)}(x) & \text{if } \ n-k \leq i \leq n-1. \end{cases}$$

with $r^{(i-n+k)}(x)$ defined by: $x^{n-k+i} = g(x) \cdot q^{(i)}(x) + r^{(i)}(x), \ \ 0 \leq i \leq k-1$; the degree of the polynomial $r^{(i)}(x)$ is less than $n-k-1$. It follows that:

$$\begin{aligned} s(x) &= [v_0 + v_1 x + \ldots + v_{n-k-1}x^{n-k-1}] + [v_{n-k}r^0(x) + \ldots + v_{n-1}r^{k-1}(x)] \\ &= [v_0 + v_1 x + \ldots + v_{n-k-1}x^{n-k-1}] + \\ &\quad +[v_{n-k}(x^{n-k} - q^{(0)}(x)g(x)) + \ldots + v_{n-1}(x^{n-1} - q^{(k-1)}(x)g(x))] \\ &= [v_0 + v_1 x + \ldots + v_{n-k-1}x^{n-k-1} + v_{n-k}x^{n-k} + \ldots + v_{n-1}x^{n-1}] - \\ &\quad -[v_{n-k}q^{(0)}(x)g(x) + \ldots + v_{n-1}q^{(k-1)}(x)g(x)] \\ &= v(x) - [v_{n-k}q^{(0)}(x) + \ldots + v_{n-1}q^{(k-1)}(x)]g(x) \\ &= v(x) - h(x)g(x). \end{aligned}$$

Thus $v(x) = h(x)g(x) + s(x)$ with $h(x) = v_{n-k}q^{(0)}(x) + \ldots + v_{n-1}q^{(k-1)}(x)$. The uniqueness of the quotient and the remainder of polynomial division grantees that the syndrome polynomial $s(x)$ constructed following the definition of the syndrome for linear codes can be used to associate an error polynomial $e(x)$ with a polynomial $v(x)$ that may or may not be a codeword; if $s(x) \equiv 0$ then $e(x) \equiv 0$ and $v(x) \equiv c(x)$, as all codewords $c(x) \in \mathcal{C}_g$ are multiple of the generator polynomial $g(x)$. The degree of $s(x)$ is $\deg[s(x)] \leq (n-k-1)$.

□

An algorithm for decoding binary cyclic codes is discussed next. This algorithm corrects individual errors, as well as bursts of errors, a topic discussed in the next section. Before presenting the algorithm we introduce the concept of *cyclic run of length* $m \leq n$, a succession

of $m$ cyclically consecutive components of an $n$-tuple. For example $v = (011100101)$ has a cyclic run of two 0's and of three 1's.

**Algorithm.** *Error Trapping:* Given an $[n, k, d]$ cyclic code $\mathcal{C}_g$ with generator polynomial $g(x)$ and $d = 2t + 1$, let $s^{(i)}(x) = \sum_{j=0}^{n-k-1} s_j x^j$ be the syndrome polynomial for $x^i v(x)$ with $v(x) = c(x) + e(x)$. The error polynomial, $e(x)$, satisfies two conditions: (i) it corresponds to an error pattern of at most $t$ errors, $w[e(x)]) \leq t$; and (ii) it contains a cyclic run of at least $k$ zeros. To decode $v(x)$ we carry out the following steps:

1. Set $i = 0$; compute syndrome $s^{(0)}(x)$ as the remainder at the division of $v(x)$ by $g(x)$: $v(x) = q(x) \cdot g(x) + s^{(0)}(x)$.
2. If the Hamming weight of $s^{(i)}(x)$ satisfies the condition $w[s^{(i)}(x)] \leq t$ then the error polynomial is $e(x) = x^{n-i} s^{(i)}(x) \mod (x^n - 1)$ and terminate.
3. Increment $i$
   - If $i = n$ stop and report that the error is not trappable.
   - Else:
   $$s^{(i)}(x) = \begin{cases} xs^{(i-1)}(x) & \text{if } \deg[s^{(i-1)}(x)] < n-k-1 \\ xs^{(i-1)}(x) - g(x) & \text{if } \deg[s^{(i-1)}(x)] \geq n-k-1 \end{cases}$$
4. Go to step 2.

To justify this algorithm we show first that the syndrome of $xv(x)$ is

$$s^{(1)}(x) = \begin{cases} xs(x) & \text{if } \deg[s(x)] < n-k-1 \\ xs(x) - g(x) & \text{if } \deg[s(x)] = n-k-1 \end{cases}$$

According to the previous proposition, if $s(x)$ is the syndrome corresponding to a polynomial $v(x)$ then the following condition must be satisfied: $\deg[s(x)] \leq n-k-1$. It follows that $xs(x)$ is the syndrome of $xv(x)$ if $\deg[s(x)] < n-k-1$; then $\deg[xs(x)] \leq n-k-1$ thus, the condition is satisfied. If $\deg[s(x)] \geq n-k-1$ then we express $s(x)$, $g(x)$, and $xs(x)$ as:

$$s(x) = \sum_{i=0}^{n-k-1} s_i x^i = a(x) + s_{n-k-1} x^{n-k-1} \quad \text{with} \quad a(x) = \sum_{i=0}^{n-k-2} s_i x^i, \ \deg[a(x)] \leq n-k-2.$$

$$g(x) = \sum_{i=0}^{n-k} g_i x^i = b(x) + x^{n-k} \quad \text{with} \quad b(x) = \sum_{i=0}^{n-k-1} g_i x^i, \ \deg[b(x)] \leq n-k-1.$$

$$\begin{aligned} xs(x) &= xa(x) + xs_{n-k-1}x^{n-k-1} = xa(x) + s_{n-k-1}x^{n-k} \\ &= xa(x) + s_{n-k-1}[g(x) - b(x)] = s_{n-k-1}g(x) + [xa(x) - s_{n-k-1}b(x)] \end{aligned}$$

We see that $d(x) = [xa(x) - s_{n-k-1}b(x)]$ is the remainder when $xs(x)$ is divided by $g(x)$; $\deg[d(x)] < \deg[g(x)] \leq n-k-1$ thus, $d(x)$ satisfies the condition to be the syndrome associated with $xs(x)$.

Let us now return to the justification of the error trapping algorithm and recall that the error pattern has weight at most $t$ and has a run of at least $k$ 0's. Then there exists an integer $i$, with $0 \leq i \leq n-1$, such that a cyclic shift of $i$ positions would bring all non-zero components of $e(x)$ within the first $n-k$ components. This new polynomial $x^i e(x) \mod (x^n - 1)$ has the syndrome $s^{(i)}(x)$ and its Hamming weight is $w[s^{(i)}(x)] \leq t$; this syndrome is the remainder when we divide the polynomial $x^i e(x) \mod (x^n - 1)$ by $g(x)$. Indeed, if the parity-check matrix of a linear $[n, k, d]$ code is $\mathrm{H} = [I_{n-k} A_{n-k,k}]$ and the error pattern has the last $n-k$ bits equal to zero, $e = (e_1, e_2, \ldots, e_{n-k}, 0, 0 \ldots 0)$ then the syndrome is

$$s^T = \mathrm{H}e^T = [I_{n-k} A_{n-k,k}]e^T = (e_1, e_2, \ldots e_{n-k})^T.$$

In this case the syndrome is the same as the error polynomial.

**Example.** *Let $\mathcal{C}_g$ be the $[7, 4, 3]$ cyclic code with the generator polynomial $g(x) = 1 + x^2 + x^3$.* In this case $n = 7$, $k = 4$; the code can correct one error as $d = 2t + 1 = 3$ thus, $t = 1$. We see also that $x^7 - 1 = (x^3 + x^2 + 1)(x^4 + x^3 + x^2 + 1)$.

We apply the error trapping algorithm for the case when we transmit the code polynomial $c(x) = (1 + x) \cdot g(x)$ and $c(x)$ is affected by the error polynomial $e(x) = x^3$ and we receive $v(x) = c(x) + e(x)$. Then the codeword is $c(x) = 1 + x^2 + x^3 + x + x^3 + x^4 = 1 + x + x^2 + x^4$ and the received 7-tuple is $v(x) = c(x) + e(x) = 1 + x + x^2 + x^3 + x^4$.

Set $i = 0$ and compute the remainder when we divide $v(x)$ by $g(x)$

$$v(x) = q(x)g(x) + r^0(x) \quad \Longrightarrow \quad x^4 + x^3 + x^2 + x + 1 = x(x^3 + x^2 + 1) + (x^2 + 1).$$

Then $s^{(0)}(x) = 1 + x^2$ and the Hamming weight $w[s^{(0)}(x)] = 2 > t$ thus, we set $i = 1$ and evaluate $\deg[s^{(0)}(x)] = 2$ and $n - k - 1 = 7 - 4 - 1 = 2$. We see that $\deg[s^{(0)}(x)] \geq 2$ thus, we compute $s^{(1)}(x)$ as

$$s^{(1)}(x) = xs^{(0)}(x) - g(x) = x(1 + x^2) - (1 + x^2 + x^3) = 1 + x + x^2.$$

We see that $w[s^{(1)}(x)] = 3 > t$; we have to go to the next iteration, $i = 2$. Then $\deg[s^1(x)] = 2$ thus, we compute $s^{(2)}(x)$ as

$$s^{(2)}(x) = xs^{(1)}(x) - g(x) = x(1 + x + x^2) - (1 + x^2 + x^3) = 1 + x.$$

Once again $w[s^{(2)}(x)] = 2 > t$; we have to go to the next iteration $i = 3$. We evaluate $\deg[s^{(2)}(x)] = 1$ and the rule to compute $s^{(3)}(x)$ is

$$s^{(3)}(x) = xs^{(2)}(x) = x(1 + x) = x + x^2.$$

Again $w[s^3(x)] = 2 > t$ and we set $i = 4$. We evaluate $\deg[s^2(x)] = 2$ thus, the rule to compute $s^{(4)}(x)$ is:

$$s^{(4)}(x) = xs^{(3)}(x) - g(x) = x(x + x^2) - (1 + x^2 + x^3) = 1.$$

We have now reached the value of the Hamming weight which allows us to identify the error pattern: $w[s^{(4)}(x)] = 1$ thus $e(x) = x^{n-i} s^{(4)}(x) \mod (x^n - 1)$. Thus,

$$e(x) = x^{7-4} \mod (x^7 - 1) = x^3.$$

**Example.** *Cyclic redundancy check.* Cyclic codes are used for error detection by link layer protocols where the data transport unit is called a frame. It is easy to show that:

*Given $p(x)$, a primitive polynomial of degree $m$, the cyclic binary code with the generator polynomial $g(x) = (x+1)p(x)$ is an $[n, k, d]$ code with*

$$n = 2^m - 1, \qquad k = 2^m - m - 2, \qquad \text{and} \qquad d = 4.$$

The $n - k = m + 1$ parity check bits are called cyclic redundancy check (CRC) bits and are included in the trailer header of the frame. A cyclic code with $g(x)$ of this form can correct one error. The probability that two or more errors were undetected and the receiver accepted the frame, $P_{underr}$, is equal to the probability that the $(m+1)$ parity-check bits are all zero when $e \geq 2$, i.e., two or more errors occur:

$$P_{underr} = \text{Prob}(e \geq 2) \times 2^{-(m+1)}.$$

When $m$ is large, undetected errors have a vanishing probability.

Two standard CRC polynomials, CRC-ANSI and CRC-ITU-T are used [228]:

$$x^{16} + x^{15} + x^2 + 1 = (x+1)(x^{15} + x + 1)$$

and

$$x^{16} + x^{12} + x^5 + 1 = (x+1)(x^{15} + x^{14} + x^{13} + x^{12} + x^4 + x^3 + x^2 + x + 1).$$

In this case $2^{-17} < 10^{-5}$ and the probability of undetected errors is less than $10^{-3}$ if the probability of two or more errors is less than $10^{-2}$.

The question we address next is if there are bounds restricting our ability to design "good" cyclic codes.

## 4.13 The Minimum Distance of a Cyclic Code; BCH Bound

We wish to establish a bound on the distance of a cyclic code, a question answered by the so-called Bose-Chaudhuri-Hocquenghem (BCH) bound. We recall a theorem discussed in Section 4.6 which states that every set of $(d-1)$ columns of H, the parity check matrix of a linear code are linearly independent if and only if the code has a distance of at least $d$. Thus, we should concentrate on the properties of the parity-check matrix of a cyclic code. First, we introduce Vandermonde matrices and then review some properties of Vandermonde matrices with elements from a finite field, $GF(q)$ [228].

A Vandermonde matrix of order $n$ with elements $x_1, x_2, \ldots, x_n \in GF(q)$ is of the form

$$A_n = \begin{pmatrix} 1 & x_1 & (x_1)^2 & (x_1)^3 & \ldots & (x_1)^{n-1} \\ 1 & x_2 & (x_2)^2 & (x_2)^3 & \ldots & (x_2)^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & (x_n)^2 & (x_n)^3 & \ldots & (x_n)^{n-1} \end{pmatrix}.$$

**Proposition.** *Let $\beta$ be a primitive element of order $n$ of the finite field $GF(q)$; thus, $\beta^n = 1$ and $\beta^i \neq 1, 0 \leq i < n$. Let $x_j = \beta^{j-1}$. Then*

$$C = BA^T = 0 \quad \text{and} \quad D = AB^T = 0$$

with

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ x_1 & x_2 & x_3 & \dots & x_{n-1} & x_n \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & \dots & (x_{n-1})^2 & (x_n)^2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (x_1)^{\alpha-1} & (x_2)^{\alpha-1} & (x_3)^{\alpha-1} & \dots & (x_{n-1})^{\alpha-1} & (x_n)^{\alpha-1} \\ (x_1)^{\alpha} & (x_2)^{\alpha} & (x_3)^{\alpha} & \dots & (x_{n-1})^{\alpha} & (x_n)^{\alpha} \end{pmatrix}$$

$$B = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_{n-1} & x_n \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & \dots & (x_{n-1})^2 & (x_n)^2 \\ (x_1)^3 & (x_2)^3 & (x_3)^3 & \dots & (x_{n-1})^3 & (x_n)^3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (x_1)^{s-1} & (x_2)^{s-1} & (x_3)^{s-1} & & (x_{n-1})^{s-1} & (x_n)^{s-1} \\ (x_1)^{s} & (x_2)^{s} & (x_3)^{s} & \dots & (x_{n-1})^{s} & (x_n)^{s} \end{pmatrix}$$

and

$$s + \alpha + 1 \leq n.$$

*Proof:* If $C = [c_{ik}]$, $A = [a_{kj}]$, and $B = [b_{ij}]$ with $a_{kj} = (x_j)^{k-1}$ and $b_{ij} = (x_j)^i$, then:

$$c_{ik} = \sum_{j=1}^{n} b_{ij} a_{jk} = \sum_{j=1}^{n} (x_j)^i (x_j)^{k-1} = \sum_{j=1}^{n} (x_j)^{i+k-1} = \sum_{j=1}^{n} (\beta^{i+k-1})^{j-1}.$$

But $\beta^{i+k-1} \leq 1$ and

$$i + k - 1 \leq s + a \leq n - 1 \implies \beta^{i+k-1} \neq 1.$$

It follows immediately that the sum of the geometric series is

$$c_{ik} = \frac{(\beta^{i+k-1})^n - 1}{(\beta^{i+k-1}) - 1} = \frac{(\beta^n)^{i+k-1} - 1}{(\beta^{i+k-1}) - 1} = 0$$

because

$$(\beta^n)^{i+k-1} = 1.$$

The proof that $D = AB^T = 0$ follows a similar path. $\square$

**Proposition.** *Call $D_n = \det[A_n]$. Then:*

$$D_n = \prod_{i>j}^{n} (x_i - x_j)$$

$$= (x_2 - x_1)(x_3 - x_2)(x_3 - x_1)\dots(x_n - x_{n-1})(x_n - x_{n-2})\dots(x_n - x_1).$$

*Proof:* We prove this equality by induction:

(i) It is easy to see that $D_2 = x_2 - x_1$.

(ii) Assume that

$$D_{n-1} = \prod_{i>j}^{n-1}(x_i - x_j)$$

$$= (x_2 - x_1)(x_3 - x_2)(x_3 - x_1)\dots(x_{n-1} - x_{n-3})(x_{n-1} - x_{n-2})\dots(x_{n-1} - x_1).$$

We can consider $D_n$ as a polynomial of degree $(n-1)$ in $x_n$ with $x_1, x_2, \dots x_{n-1}$ as zeros. The coefficient of $(x_n)^{n-1}$ is $D_{n-1}$. Thus:

$$D_n = \prod_{i>j}^{n}(x_i - x_j) = \prod_{i=1}^{n-1}(x_n - x_i)D_{n-1}. \qquad \square$$

The BCH (Bose, Ray-Chaudhuri, Hocquenghem) bound. Let $g(x)$ be the generator polynomial of the $[n, k]$ cyclic code $\mathcal{C}$ over $GF(q)$. Let $\beta \in GF(q)$ be a primitive element of order $n$. If $g(x)$ has among its zeros $(d-1)$ elements of $GF(q)$, $\beta^\alpha, \beta^{\alpha+1}, \dots, \beta^{\alpha+d-2}$, then $d$ is the distance of the cyclic code.

*Proof:* Recall that the minimum distance between any pair of codewords called the distance of a linear code is equal to the number of independent column vectors of H, the parity check matrix of the code. The parity check matrix of code $\mathcal{C}$ is

$$\mathrm{H} = \begin{pmatrix} 1 & (\beta)^{\alpha} & (\beta)^{2\alpha} & (\beta)^{3\alpha} & \dots & (\beta)^{(n-1)\alpha} \\ 1 & (\beta)^{\alpha+1} & (\beta)^{2(\alpha+1)} & (\beta)^{3(\alpha+1)} & \dots & (\beta)^{(n-1)(\alpha+1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & (\beta)^{(\alpha+d-2)} & (\beta)^{2(\alpha+d-2)} & (\beta)^{3(\alpha+d-2)} & \dots & (\beta)^{(n-1)(\alpha+d-2)} \end{pmatrix}.$$

The determinant of $(d-1)$ columns of the parity check matrix H is

$$\det[\mathrm{H}] = \begin{vmatrix} (\beta)^{\alpha i_1} & (\beta)^{\alpha i_2} & (\beta)^{\alpha i_3} & \dots & (\beta)^{\alpha i_{d-1}} \\ (\beta)^{(\alpha+1)i_1} & (\beta)^{(\alpha+1)i_2} & (\beta)^{(\alpha+1)i_3} & \dots & (\beta)^{(\alpha+1)i_{d-1}} \\ \vdots & \vdots & \vdots & & \vdots \\ (\beta)^{(\alpha+d-2)i_1} & (\beta)^{(\alpha+d-2)i_2} & (\beta)^{(\alpha+d-2)i_3} & \dots & (\beta)^{(\alpha+d-2)i_{d-1}} \end{vmatrix}.$$

But $x_j = \beta^{i_j}$, thus

$$\det[\mathrm{H}] = \begin{vmatrix} (x_1)^{\alpha} & (x_2)^{\alpha} & (x_3)^{\alpha} & \dots & (x_{d-1})^{\alpha} \\ (x_1)^{(\alpha+1)} & (x_2)^{(\alpha+1)} & (x_3)^{(\alpha+1)} & \dots & (x_{d-1})^{(\alpha+1)} \\ \vdots & \vdots & \vdots & & \vdots \\ (x_1)^{(\alpha+d-2)} & (x_2)^{(\alpha+d-2)} & (x_3)^{(\alpha+d-2)} & \dots & (x_{d-1})^{(\alpha+d-2)} \end{vmatrix},$$

or

$$\det[\mathrm{H}] = x_1^\alpha x_2^\alpha x_3^\alpha \dots x_{d-1}^\alpha \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_{d-1} \\ \vdots & \vdots & \vdots & & \vdots \\ (x_1)^{(d-2)} & (x_2)^{(d-2)} & (x_3)^{(d-2)} & \dots & (x_{d-1})^{(d-2)} \end{vmatrix}.$$

Finally, $x_i \neq x_j$ thus

$$\det[\mathrm{H}] = x_1^\alpha x_2^\alpha x_3^\alpha \dots x_{d-1}^\alpha \prod_{i>j}^{d-1} (x_i - x_j) \neq 0.$$

This proves that $(d-1)$ columns of the parity check matrix are linearly independent, thus the minimum distance of the code is $d$.

## 4.14 Burst Errors. Interleaving

So far we have assumed that errors are uncorrelated and we have only been concerned with the construction of error correcting codes able to handle errors with a random distribution. Yet, physical systems may be affected by errors concentrated in a short time interval, or affecting a number of adjacent bits; communication channels, as well as memory and information processing systems may experience correlated errors. A burst of noise triggered by an electrostatic phenomena, a speck of dust on a recording media, a slight variation of the voltage of a power supply may affect a contiguous stream of bits transmitted, stored, or being processed. Codes designed to handle *bursts* of errors are called *Burst Error Correcting (BEC)* codes. We start this section with a few definitions, then present several properties and discuss several classes of BEC codes.

An error burst of length $l$ is an $n$-tuple whose non-zero symbols are confined to a span of no fewer than $l$ symbols. A wraparound burst of length $l$ is a cyclic shift of a burst of length $l$. A burst is specified by its *location*, the least significant digit of the burst, and the *burst pattern*. If $i$ is the location and $b(x)$ is the pattern ( $b_0 \neq 0$) then the error burst $e(x)$ and the wraparound burst are:

$$e(x) = x^i \cdot b(x), \qquad e(x) = x^i \cdot b(x) \mod (x^n - 1).$$

**Example.** *Assume* $n = 10$. Then 0001101000 is a burst of length 4, the location is 3, and the pattern is $1 + x + x^3$; 0100000011 is a wraparound burst of length 4, location 8, and the same pattern, $1 + x + x^3$.

Given an $[n, k]$ linear code capable of correcting error bursts of length $l$, its burst error correcting efficiency is defined as

$$\eta = \frac{2l}{(n-k)}.$$

The following two propositions show that cyclic codes and block codes can be used for burst error detection and correction.

**Proposition.** *A cyclic code with generator polynomial $g(x)$ of degree $r = n - k$ can detect all burst errors of length $l \leq r$.*

*Proof:* Call $c(x)$ the polynomial representing the codeword sent and $e(x)$ the error polynomial. We are able to detect an error when $v(x) = c(x) + e(x)$ is not a codeword; this condition requires that $e(x)$ is not a codeword, thus it is not a multiple of the generator polynomial $g(x) = g_0 + g_1 x + g_2 x^2 + ... + x^r$. Assume that the error burst starts at position $i$ and affects $l \leq r$ consecutive bits thus, $e(x) = x^i \cdot b(x)$. We see immediately that $g(x)$ and $x^i$ are relatively prime because $g_0 \neq 0$. The length of the burst $l \leq r$ and this implies that $\deg[b(x)] \leq r - 1 < \deg[g(x)] = r$ thus, $b(x)$ is not a multiple of $g(x)$. It follows immediately that $e(x) = x^i \cdot b(x)$ is not a multiple of $g(x)$, therefore it is not a codeword.

A cyclic code will detect most burst errors of length $l > r$; indeed, the probability that a binary cyclic code with generator polynomial of degree $r$ will detect error bursts of length $l > r$ is equal to $1 - 2^{-r}$ and the probability of detecting error bursts of length $l$ is $1 - 2^{-r+1}$.

**Proposition.** *A block code can correct all bursts of length $\leq l$ if and only if no two codewords differ by the sum of two bursts of length $\leq l$.*

An equivalent statement is that a linear block code can correct all bursts of length $\leq l$ if and only if no codeword is the sum of two bursts of length $\leq l$.

**Proposition.** *The Rieger Bound - the burst error correcting ability $l$ of an $[n, k]$ linear block code satisfies the condition*

$$l \leq \frac{n-k}{2}.$$

This inequality is an immediate consequence of the fact that all polynomials of degree smaller than $l$ must have distinct syndromes, otherwise the difference of two polynomials would be a codeword and, at the same time, the sum of two bursts of length $\leq l$.

We now examine several classes of burst error correcting codes and their burst error correcting efficiency. The Rieger bound imposes that linear codes have error correcting efficiency $\eta \leq 1$. On the other hand, Reed-Solomon codes discussed in Section 4.15 have $\eta = 1$. Reed-Solomon codes use $2t$ check symbols and can correct $t$ random errors and, thus, bursts of length $\leq t$. The Fire codes have an error correcting efficiency of $\approx 2/3$.

Cyclic codes have some desirable properties for burst error correction. Indeed, if symbol $j$ of cyclic shift $r^{(i)}(x)$ can be corrected, then the same error magnitude can be used to correct the symbol $(j - i) \mod n$ of $r(x)$. To decode a cyclic code we use a feedback shift register implementing the generator polynomial of the code, $g(x)$. If $r(x)$ is the received polynomial we use the following algorithm: (i) compute $s(x) = r(x) \mod g(x)$; (ii) Shift syndrome register with feedback until $s^{[i]}(x)$ contains a syndrome with a known error $e(x)$; (iii) decode $r(x)$ using $e^{(-i)}(x)$. For example, $g(x) = x^4 + x + 1$ is a primitive polynomial over $GF(2)$ which generates a $(15, 11)$ cyclic Hamming code. The code with $g(x) = (x + 1)(x^4 + x^3 + 1)$ can correct any bursts of length $\leq 2$.

Fire codes are cyclic codes over $GF(q)$ with a generator polynomial $g(x) = (x^{2t-1} - 1)p(x)$ such that: (i) $p(x)$ is a prime polynomial of degree $m$ and order $e$ over $GF(q)$, (ii) $m \geq t$, and (iii) $p(x)$ does not divide $(x^{2t-1} - 1)$. The block length of a Fire code, $n$, is the order of the generator polynomial, $g(x)$; $n = \text{lcm}(e, 2t - 1)$.

Interleaving is a simple technique to ensure burst error correction capabilities of a code originally designed for random error correction, or to enhance the capabilities of an existing burst error correcting code. The basic idea of interleaving is to alternate the transmission of the individual bits of a set of codewords, Figure 102. If we have an original $[n, k]$ code $\mathcal{C} = \{c_1, c_2 \dots c_m\}$ with $m$ codewords capable of correcting a burst of length $b \geq 1$ errors or less, then interleaving allows us to correct a burst of length $mb$ or less. We construct a table and every row of this table is a codeword of the original code. When transmitted or stored the symbols are interleaved; we send the symbols column wise, the first symbol of $c_1$, then the first symbol $c_2$, and so on. A burst of length $mb$ or less can have at most $b$ symbols in any row affected by the error. But each row is a codeword of the original code and it is able to correct a burst of length $b$ or less. If $b = 1$ we start with a code with no burst error correction capabilities.

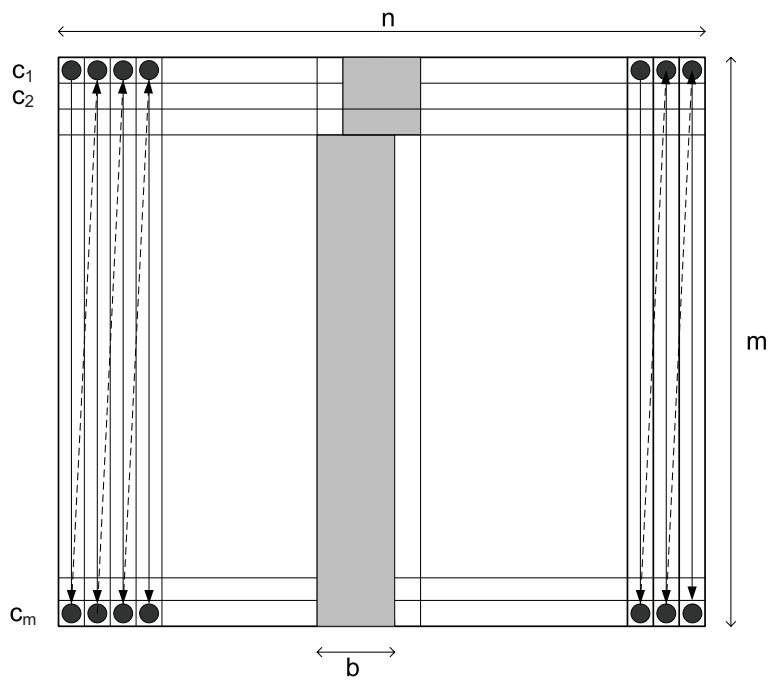



Figure 102: Schematic representation of interleaving. We wish to transmit $m$ codewords and construct a table with $m$ rows and $n$ columns; every row of this table is a codeword of length $n$ of the original code. We send the table column-wise thus, the symbols are interleaved. Now the code is capable of correcting a burst of length $mb$ or less because in any row at most $b$ symbols could be affected by the error and the original code can correct a burst of this length.

Let us now give an example of interleaving. We shall start with the Reed-Muller code with the generator matrix:

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

This code is able to correct a single error, its distance is $d = 3$, and the length of a codeword is $n = 8$. Consider the set of five messages $\mathcal{M} = \{0001, 0010, 0100, 0011, 0111\}$ encoded as the codewords $\mathcal{C} = \{00110110, 01011110, 10011010, 01101000, 11110000\}$, respectively. A first option is to send the five codewords in sequence:

$$00110110|01011110|10011010|01101000|11110000.$$

In this case a burst of errors of length $b \geq 2$ would go undetected. Note that the vertical bars are not part of the string transmitted; they mark the termination of each one of the five codewords. If we interleave the symbols and transmit the first bits of the five codewords followed by the second bits, and so on, then the string of bits transmitted is

$$00101|01011|10011|01001|01110|11000|11100|00000.$$

The interleaving allows us to correct any burst of errors of length $b \leq 4$. Now there are eight groups, one for each symbol of a codeword, and the vertical bars mark the end of a group.

## 4.15 Reed-Solomon Codes

Many of the codes we have examined so far are binary codes. We start our presentation of non-binary codes with a brief overview and then introduce a class of non-binary codes, the Reed-Solomon (R-S) codes. R-S codes are non-binary cyclic codes widely used for burst error correction; they were introduced in 1960 by Irving Reed and Gustav Solomon [354]. R-S codes are used for data storage on CDs, DVDs and Blu-ray Discs, in data transmission technologies such as DSL (Digital Subscriber Line) and WiMAX (Worldwide Interoperability for Microwave Access), in broadcast systems such as DVB (Digital Video Broadcasting), and in RAID 6 computer disk storage systems.

The density of codewords of the linear code $\mathcal{C}$ over the alphabet $\mathcal{A}$ is the ratio of the number of codewords to the number of $n$-tuples constructed with the symbols from the alphabet $\mathcal{A}$. Table 15 shows the density of the codewords $\mid \mathcal{C} \mid / \mid V_n \mid$ and the rate $R$ of binary and non-binary $[n, k]$ codes over $GF(2^m)$.

Table 15: The cardinality of the $n$-dimensional vector space $V_n$, the cardinality of an $[n, k]$ code $\mathcal{C}$, the density of the codewords, and the rates of a binary and of a non-binary code over $GF(2^m)$.

| Code | $\mid V_n \mid$ | $\mid \mathcal{C} \mid$ | $\mid \mathcal{C} \mid / \mid V_n \mid$ | R |
|---|---|---|---|---|
| Binary | $2^n$ | $2^k$ | $2^{-(n-k)}$ | $k/n$ |
| Non-binary | $q^n = 2^{mn}$ | $q^k = 2^{mk}$ | $2^{-(n-k)m}$ | $k2^m/n2^m = k/n$ |

We notice that the rates of a binary $[n, k]$ and of a $q = 2^m$-ary $[n, k]$ code are equal. We also see that non-binary codes have a lower density of codewords thus, a larger distance between codewords and this implies increased error correction capability.

**Example.** *A binary Hamming code versus a code using 3-bit symbols.* The binary $[7, 4, 3]$ code is a subspace of cardinality $\mid \mathcal{C} \mid = 2^4 = 16$ of a vector space of cardinality $\mid V_n \mid = 2^7 = 128$. The density of the codewords is

$$\frac{16}{128} = \frac{1}{8}.$$

If the code uses an alphabet with eight symbols, each symbol requiring $m = 3$ bits, then the total number of codewords becomes $2^{4\times 3} = 2^{12} = 4096$, while the number of 7-tuples is $2^{7\times 3} = 2^{21} = 2,097,152$. Now the density of the codewords is much lower:

$$\frac{4096}{2,097,152} = \frac{1}{2^9} = \frac{1}{512}.$$

Let $\mathbb{P}_{k-1}[x]$ be the set of polynomials $p_i(x)$ of degree at most $(k-1)$ over the finite field $GF(q)$. Given $n > k$ then an $[n, k]$ Reed-Solomon (R-S) code over $GF(q)$ consists of all codewords $c_i$ generated by polynomials $p_i \in \mathbb{P}_{k-1}[x]$:

$$c_i = (p_i(x_1), p_i(x_2), p_i(x_3), \ldots p_i(x_{n-1}), p_i(x_n))$$

where $x_1, x_2, \ldots \ldots x_n$ are $n$ different elements of $GF(q)$; $x_i = \beta^{i-1}, 1 \leq i \leq n$ with $\beta$ a primitive element of $GF(q)$. The polynomial $p_i(x)$ is called the generator of the codeword $c_i$.

Reed-Solomon codes have several desirable properties; for example, an $[n, k]$ R-S code can achieve the Singleton bound, the largest possible code distance (minimum distance between any pair of codewords) of any linear code: $d = n - k + 1$. This means that the code can correct at most $t$ errors with

$$t = \lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{n-k}{2} \rfloor.$$

Recall that a cyclic code $\mathcal{C}_g$ has a generator polynomial $g(x)$ and all codewords $c_i(x)$ are multiple of the generator polynomial, $c_i(x) = g(x) \cdot p_i(x),\ \forall c_i \in \mathcal{C}$. The Reed-Solomon codes are constructed differently, first we identify $p_i(x)$, the generator polynomial of each codeword $c_i(x)$ and then evaluate this polynomial for the elements $(x_1, x_2, \ldots, x_n)$ of the finite field $GF(q)$ to identify the $n$ coordinates $c_{i,j}$ of the codeword $c_i = \sum_j^n c_{i,j} x^j$. We shall later see that an R-S code is a cyclic code and we can define a generator polynomial. The difference between R-S codes and the BCH codes discussed at the end of this section is that for R-S codes $\beta$ is a primitive element of $GF(q)$ while in case of BCH codes $\beta$ is an element of an extension Galois Field.

As noted earlier, $\mathbb{P}_{k-1}[x] \subset GF(q)[x]$ is a vector space of dimension $(k-1)$. For example, when $k = 3$ the polynomials of degree at most 2 over $GF(2)$ are:

$$\begin{aligned} \mathbb{P}_2[x] \quad = \quad & \{p_0(x) = 1,\ p_1(x) = x,\ p_2(x) = x+1,\ p_3(x) = x^2,\ p_4(x) = x^2+1,\ p_5(x) = x^2+x, \\ & p_6(x) = x^2+x+1\} \end{aligned}$$

It is easy to see that an R-S code is a linear code. Let two codewords of an R-S code be:

$$c_i = (p_i(x_1), p_i(x_2), \ldots, p_i(x_n)) \quad \text{and} \quad c_j = (p_j(x_1), p_j(x_2), \ldots, p_j(x_n))$$

Then $\alpha c_i + \beta c_j = (g(x_1), g(x_2), \ldots, g(x_n))$ with $\alpha, \beta \in GF(q)$ and $g(x) = \alpha p_i(x) + \beta p_j(x)$.

*The minimum distance of an* $[n, k]$ *Reed-Solomon code is* $(n-k+1)$. Indeed, each codeword consists of $n$ symbols from $GF(q)$ and it is generated by a polynomial $p_i(x) \in \mathbb{P}_{k-1}$ of degree at most $(k-1)$. Such a polynomial may have at most $(k-1)$ zeros. It follows that the minimum weight of a codeword is $n-(k-1) = n-k+1$ thus, the minimum distance the code is $d_{min} = n-k+1$.

**Encoding R-S codes.** We can easily construct a generator and a parity-check matrix of an $[n, k]$ R-S code when the $n$ elements of the finite field $x_1, x_2, \ldots \ldots x_n \in GF(q)$ can be expressed as powers of $\beta$, the characteristic element of the finite field. Each row of the generator matrix G consists of one polynomial in $\mathbb{P}_k[x]$ evaluated for all elements $x_i \in GF(q)$. For example, the rows 1 to $k$ may consist of the polynomials $p_0(x) = x^0,\ p_1(x) = x^1, \ldots, p_{k-1}(x) = x^{k-1}$; thus, a generator matrix of an $[n, k]$ R-S code is [228]:

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ x_1 & x_2 & x_3 & \ldots & x_{n-1} & x_n \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & \ldots & (x_{n-1})^2 & (x_n)^2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (x_1)^{k-1} & (x_2)^{k-1} & (x_3)^{k-1} & \ldots & (x_{n-1})^{k-1} & (x_n)^{k-1} \end{pmatrix}.$$

To construct the parity-check matrix H of the R-S code with the generator matrix G we use a property discussed in Section 4.13: if G is a $k \times n$ matrix then there exist an $(n-k) \times n$ matrix H such that $\mathrm{HG}^T = \mathrm{GH}^T = 0$. It follows immediately that the parity check matrix H of the code is

$$\mathrm{H} = \begin{pmatrix} x_1 & x_2 & x_3 & \ldots & x_{n-1} & x_n \\ (x_1)^2 & (x_2)^2 & (x_3)^2 & \ldots & (x_{n-1})^2 & (x_n)^2 \\ (x_1)^3 & (x_2)^3 & (x_3)^3 & \ldots & (x_{n-1})^3 & (x_n)^3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (x_1)^{n-k} & (x_2)^{n-k} & (x_3)^{n-k} & \ldots & (x_{n-1})^{n-k} & (x_n)^{n-k} \end{pmatrix}.$$

The two matrices can be expressed in terms of the primitive element $\beta$ with the property $\beta^n = 1$ and $\beta^i \neq 1, 0 \leq i < n$. Then:

$$x_j = \beta^{j-1} \quad 1 \leq j \leq q-1 \quad \text{such that} \quad x_j^n = \beta^{j-1}, 1 \leq j \leq n, \quad \text{and} \quad n = q-1.$$

Now the generator and the parity-check matrices are:

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ 1 & (\beta)^1 & (\beta)^2 & \ldots & (\beta)^{n-2} & (\beta)^{n-1} \\ 1 & (\beta^2)^1 & (\beta^2)^2 & \ldots & (\beta^2)^{n-2} & (\beta^2)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & (\beta^{k-1})^1 & (\beta^{k-1})^2 & \ldots & (\beta^{k-1})^{n-2} & (\beta^{k-1})^{n-1} \end{pmatrix}$$

and:

$$\mathrm{H} = \begin{pmatrix} 1 & (\beta)^1 & (\beta)^2 & \dots & (\beta)^{n-2} & (\beta)^{n-1} \\ 1 & (\beta^2)^1 & (\beta^2)^2 & \dots & (\beta^2)^{n-2} & (\beta^2)^{n-1} \\ 1 & (\beta^3)^1 & (\beta^3)^2 & \dots & (\beta^3)^{n-2} & (\beta^3)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & (\beta^{n-k})^1 & (\beta^{n-k})^2 & \dots & (\beta^{n-k})^{n-2} & (\beta^{n-k})^{n-1} \end{pmatrix}$$

.

**Algorithm.** *To determine the codeword $c_i$ of an $[n, k]$ Reed-Solomon code generated by the information polynomial $p_i(x) = p_{i,0} + p_{i,1}x^1 + \ldots + \ldots + p_{i,n}x^n$ with $p_i(x) \in \mathbb{P}_{k-1}$ carry out the following three steps:*

*(i) choose $k$, $q$, and $n$ with $k < n < q$;*

*(ii) select $n$ elements $x_j \in GF(q), 1 \leq j \leq n$;*

*(iii) evaluate the polynomial $p_i(x)$ for $x = x_j, 1 \leq j \leq n$ and then construct:*

$$c_i = (p_i(x_1), p_i(x_2), \ldots, p_i(x_n))$$

**Example.** *Encoding for a $[10, 5]$ Reed-Solomon code over $GF(11)$.* In this case $n = 10$, $k = 5$, and $q = 11$. First, we have to find a primitive element $\beta^n = 1$; we notice that $2^{10} \mod 11 = 1$; then the $q - 1 = 10$ non-zero elements of $GF(11)$ can be expressed as $x_j = 2^{j-1}$ with $1 \leq j \leq 10$:

$$\begin{array}{l||l||l} x_1 = 2^0 \mod 11 = 1 & x_5 = 2^4 \mod 11 = 5 & x_9 = 2^8 \mod 11 = 3 \\ x_2 = 2^1 \mod 11 = 2 & x_6 = 2^5 \mod 11 = 10 & x_{10} = 2^9 \mod 11 = 6 \\ x_3 = 2^2 \mod 11 = 4 & x_7 = 2^6 \mod 11 = 9 & \\ x_4 = 2^3 \mod 11 = 8 & x_8 = 2^7 \mod 11 = 7 & \end{array}$$

The information vectors are denoted as $(i_0, i_1, i_2, i_3, i_4)$; the 31 polynomials of degree at most $k = 4$ with coefficients from $GF(2)$ corresponding to these vectors and the codewords they generate are shown in Table 16.

For example, when the information vector is $(0, 0, 0, 0, 1)$ the codeword is generated by the polynomial $p_5(x) = x^4$ then $c_{(0,0,0,0,1)} = (1, 5, 3, 4, 9, 1, 5, 3, 4, 9)$. Indeed,

$$\left| \begin{array}{l||l} p_5(x_1) = (x_1)^4 = 1^4 \mod 11 = 1 & p_5(x_6) = (x_6)^4 = 10^4 \mod 11 = 1 \\ p_5(x_2) = (x_2)^4 = 2^4 \mod 11 = 5 & p_5(x_7) = (x_7)^4 = 9^4 \mod 11 = 5 \\ p_5(x_3) = (x_3)^4 = 4^4 \mod 11 = 3 & p_5(x_8) = (x_8)^4 = 7^4 \mod 11 = 3 \\ p_5(x_4) = (x_4)^4 = 8^4 \mod 11 = 4 & p_5(x_9) = (x_9)^4 = 3^4 \mod 11 = 4 \\ p_5(x_5) = (x_5)^4 = 5^4 \mod 11 = 9 & p_5(x_{10}) = (x_{10})^4 = 6^4 \mod 11 = 9 \end{array} \right|$$

A generator matrix of the code is

$$\mathrm{G} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 5 & 10 & 9 & 7 & 3 & 6 \\ 1 & 4 & 5 & 9 & 3 & 1 & 4 & 5 & 9 & 3 \\ 1 & 8 & 9 & 6 & 4 & 10 & 3 & 2 & 5 & 7 \\ 1 & 5 & 3 & 4 & 9 & 1 & 5 & 3 & 4 & 9 \end{pmatrix}.$$

Table 16: The information vector, the generator polynomials and the codewords for the $[10, 4]$ R-S code over $GF(11)$.

| Vector | Generator polynomial | Codeword |
|---|---|---|
| $(1,0,0,0,0)$ | $p_1(x) = 1$ | $(1,1,1,1,1,1,1,1,1,1)$ |
| $(0,1,0,0,0)$ | $p_2(x) = x$ | $(1,2,4,8,5,10,9,7,3,6)$ |
| $(0,0,1,0,0)$ | $p_3(x) = x^2$ | $(1,4,5,9,3,1,4,5,9,3)$ |
| $(0,0,0,1,0)$ | $p_4(x) = x^3$ | $(1,8,9,6,4,10,3,2,5,7)$ |
| $(0,0,0,0,1)$ | $p_5(x) = x^4$ | $(1,5,3,4,9,1,5,3,4,9)$ |
| $(1,1,0,0,0)$ | $p_6(x) = 1 + x$ | $(2,3,5,9,6,0,10,8,4,7)$ |
| $(1,0,1,0,0)$ | $p_7(x) = 1 + x^2$ | $(2,5,6,10,4,2,5,6,10,4)$ |
| $(1,0,0,1,0)$ | $p_8(x) = 1 + x^3$ | $(2,9,10,7,5,0,4,3,6,8)$ |
| $(1,0,0,0,1)$ | $p_9(x) = 1 + x^4$ | $(2,6,4,5,10,2,6,4,5,10)$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $(1,1,1,1,1)$ | $p_{31}(x) = x^4 + x^3 + x^2 + x^1 + 1$ | $(5,9,0,6,0,1,5,0,0,4)$ |

**Decoding R-S codes.** Assume the codeword $c_i$ of an $[n, k]$ R-S code is generated by the polynomial $p_i(x)$; an error $e_i$ occurs when we send $c_i = (p_i(x_1), p_i(x_2), \ldots p_i(x_{n-1}), p_i(x_n))$ and we receive the $n$-tuple $v$ affected by at most $t$ errors:

$$v = c_i + e, \quad \text{with} \quad w(e) \leq t = \lfloor \frac{n-k}{2} \rfloor$$

Before presenting the decoding algorithm for an R-S codes we introduce two new concepts, the interpolating polynomial and the error locator polynomial; then we prove a theorem of existence for the interpolation polynomial and show that the polynomial used to generate the codeword $c_j$ is a ratio of the interpolating and error locating polynomials. The bivariate interpolating polynomial is defined as

$$Q(x, y) = Q_0(x) + yQ_1(x).$$

$Q(x, y)$ is a polynomial in $x$ and $y$ with $Q_0(x)$ and $Q_1(x)$ polynomials of degrees $m_0$ and $m_1$, respectively,

$$Q_0(x) = Q_{0,0} + Q_{0,1}x + \ldots + Q_{0,j}x^j + \ldots + Q_{0,m_0}x^{m_0}$$

and

$$Q_1(x) = Q_{1,0} + Q_{1,1}x + \ldots + Q_{1,j}x^j + \ldots + Q_{0,m_1}x^{m_1},$$

the so-called error locator polynomial.

**Proposition.** *When we receive an $n$-tuple $v = (v_0, v_2, \ldots, v_{n-1})$ an interpolation polynomial for the $[n, k]$ Reed-Solomon code with $t = \lfloor (n-k)/2 \rfloor$ exists if the following three conditions are satisfied:*

$(i)$ $\quad Q(x_i, v_i) = Q_0(x_i) + v_iQ_1(x_i) = 0, \quad i = 1, 2, \ldots (n-1), n, \quad x_i, v_i \in GF(q).$

(*ii*) $\deg[Q_0(x)] = m_0 \leq n-1-t$

(*iii*) $\deg[Q_1(x)] = m_1 \leq n-1-t-(k-1)$.

*Proof:* To prove the existence of $Q(x_i, v_i)$ we have to show that we can determine the $m_0+1$ coefficients $Q_{0,j}$ and the $m_1+1$ coefficients $Q_{1,j}$ given the $n$ equations $Q(x_i, v_i) = 0,\ \ 1 \leq i \leq n$, and that not all of them are equal to zero. The total number of coefficients of the bivariate polynomial is $(m_0+1)+(m_1+1) = (n-t)+(n-t-(k-1)) = 2n-k+1-2t$. If we substitute $2t = d = n-k+1$ we see that the total number of coefficients is $2n-k+1-n+k-1 = n$. The system of $n$ equations with $n$ unknowns has a non-zero solution if the determinant of the system is non-zero, and in this case a polynomial $Q(x_i, v_i)$ that satisfies the three conditions exists.

□

**Proposition.** *Let $\mathcal{C}$ be a Reed-Solomon code with $t = \lfloor \frac{n-k}{2} \rfloor$. If we transmit the codeword $c_i(x)$ generated by the polynomial $p_i(x) \in \mathbb{P}_{k-1}[x]$ and we receive $v(x) = \sum_j v_j x^i = c_i(x) + e(x)$ and if the number of errors is $w[e(x)] \leq t$ then the codeword $c_i(x)$ is generated by the polynomial $p_i(x) \in \mathbb{P}_{k-1}[x]$ given by*

$$p_i(x) = -\frac{Q_0(x)}{Q_1(x)}$$

*Proof:* If $v = (v_0, v_2, \ldots, v_{n-1})$ is the received $n$-tuple and $e(x) = (e_0, e_1, \ldots e_{n-1})$ is the error polynomial then at most $t$ components of $e$ are non-zero, $e_l = r_l$ thus, at least $n-t$ components $e_j$ of $e$ are zero. An interpolating polynomial $Q(x, v) = Q(x, (p_j(x) + e(x)))$ exists and has at least $n-t$ zeros, namely those values $x = x_j$ when $p_i(x_j) = v_j + e_j$ with $e_j = 0$. But the degree of $Q(x, p_i(x))$ is $m_0 \leq n-t-1$. It follows that $Q(x, p_i(x)) = 0$ for $x = x_j, j = 1, 2, \ldots n$. But $Q(x, p_i(x)) = 0$ implies that

$$Q_0(x) + p_i(x)Q_2(x) = 0 \quad \text{thus} \quad p_i(x) = -\frac{Q_0(x)}{Q_1(x)}$$

□

**Algorithm.** *Let $\mathcal{C}$ be an $[n, k]$ Reed-Solomon code able to correct at most $t = \lfloor (n-k)/2 \rfloor$ errors. We send $c_i = (p_i(x_1), p_i(x_2), \ldots, p_i(x_n))$ generated by the polynomial $p_i(x)$ and we receive $v = c_i + e$ with $v = (v_0, v_2, \ldots, v_{n-1})$. To determine $c_i$ we first compute the generator polynomial $p_i(x)$ and then evaluate it for $x = x_j, 1 \leq j \leq n$. The decoding algorithm consists of the following steps:*

*Step 1.* Compute $m_0 = n-1-t$ and $m_1 = n-1-t-(k-1)$. Solve the system of linear equations:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \ldots x_1^{m_0} & v_1 & v_1x_1 & \ldots v_1x_1^{m_1} \\ 1 & x_2 & x_2^2 & \ldots x_2^{m_0} & v_2 & v_2x_2 & \ldots v_2x_2^{m_1} \\ 1 & x_3 & x_3^2 & \ldots x_3^{m_0} & v_3 & v_3x_1 & \ldots v_3x_3^{m_1} \\ \vdots & & & & & & \\ 1 & x_n & x_n^2 & \ldots x_n^{m_0} & v_n & v_nx_n & \ldots v_nx_n^{m_1} \end{bmatrix} \begin{bmatrix} Q_{0,0} \\ Q_{0,1} \\ Q_{0,2} \\ \vdots \\ Q_{0,m_0} \\ Q_{1,0} \\ Q_{1,1} \\ Q_{1,2} \\ \vdots \\ Q_{1,m_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

*Step 2.* Construct the following polynomials:

$$Q_0(x) = \sum_{j=0}^{m_0} Q_{0,j}x^j, \qquad Q_1(x) = \sum_{j=1}^{m_1} Q_{1,j}x^j \qquad \text{and} \quad p(x) = -\frac{Q_0(x)}{Q_1(x)}.$$

*Step 3.* If $p(x) \in \mathbb{P}_{k-1}$ then the algorithm succeeds and produces $(p(x_1), p(x_2), \ldots p(x_n))$.

**An alternative definition of Reed-Solomon codes.** We discuss now several properties of extension finite fields $GF(2^m)$ and show that the polynomial $x^n - 1$ can be factored as a product of irreducible monic polynomial over $GF(2^m)$. Recall that the set of polynomials $p(x) = \sum_{i=0}^{k} \alpha_i x^i$ of degree at most $k$ and with $\alpha_i \in GF(2^m)$ is denoted as $\mathbb{P}_{2^m}[x]$.

**Proposition.** *If the polynomial $p(x) \in \mathbb{P}_{2^m}[x]$ then $p(x^2) = [p(x)]^2$.*

*Proof:* If $\alpha, \beta \in GF(2^m)$ then $(\alpha+\beta)^2 = \alpha^2+\beta^2$. Indeed, the term $2\alpha\cdot\beta = 0$ as $\alpha, \beta \in GF(2^m)$. If $p(x) = \alpha_k x^k + \alpha_{k-1}x^{k-1} + \ldots + \alpha_1 x + \alpha_0 x^0$, with $\alpha_0, \alpha_1, \ldots, \alpha_k \in GF(2^m)$ then

$$[p(x)]^2 = [\alpha_k x^k + \alpha_{k-1}x^{k-1} + \ldots + \alpha_1 x + \alpha_0 x^0]^2 = \alpha_k^2 x^{2k} + \alpha_{k-1}^2 x^{2(k-1)} + \ldots + \alpha_1^2 x^2 + \alpha_0^2.$$

$$p(x^2) = \alpha_k x^{2k} + \alpha_{k-1}x^{2(k-1)} + \ldots + \alpha_1 x^2 + \alpha_0.$$

Thus, $p(x^2) = [p(x)]^2$ if and only if $\alpha_i = (\alpha_i)^2,\ 1 \le i \le k$. But this is true if and only if $\alpha_i \in GF(2^m)$. It follows immediately that if $\beta$ is a zero of $p(x)$, then so is $\beta^2$:

$$p(\beta) = 0 \quad \Leftrightarrow \quad p(\beta^2) = 0.$$

□

**Proposition.** *The extension Galois field $GF(2^m)$ is a subfield of $GF(2^n)$ if and only if $m|n$ ($m$ divides $n$).*

*Proof:* It is easy to see that $a^q = a$ if and only if $a \in GF(q)$. If $a \in GF(q)$ then it can be expressed as a power of the primitive element $\beta$ of $GF(q)$ as $a = \beta^i$. Then

$$a^{q-1} = (\beta^i)^{(q-1)} = (\beta^{q-1})^i = 1^i = 1$$

as $\beta^{q-1} = 1$. If $a^{q-1} = 1$ then $a^q = a$ and $a$ is the primitive element of a finite field $GF(q)$.

We can thus express $GF(2^n) = \{x|x^{2^n} = x\}$ and $GF(2^m) = \{x|x^{2^m} = x\}$. In preparation for our proof we show that

$$(x^m - 1)|(x^n - 1) \quad \Leftrightarrow \quad m|n.$$

Given a non-negative integer $k$ we have the identity

$$(x^n - 1) = (x^m - 1)(x^{n-m} + x^{n-2m} + \ldots + x^{n-km}) + x^{n-km} - 1.$$

or,

$$\frac{x^n - 1}{x^m - 1} = (x^{n-m} + x^{n-2m} + \ldots + x^{n-km}) + \frac{x^{n-km} - 1}{x^m - 1}.$$

If $n - km = m$ then

$$\frac{x^n - 1}{x^m - 1} = (x^{n-m} + x^{n-2m} + \ldots + x^{n-km}) + 1.$$

In this case $n = m(k + 1)$ so that when $m \mid n$ then $(x^m - 1) \mid (x^n - 1)$.

Now we show that $m|n$ is a necessary and sufficient condition for $GF(2^m)$ to be a subfield of $GF(2^n)$. The condition is necessary: if $GF(2^m)$ is a subfield of $GF(2^n)$ them $m|n$. Indeed, $GF(2^m)$ contains an element of order $2^m - 1$. When $GF(2^m)$ is a subfield of $GF(2^n)$ then $GF(2^n)$ contains an element of order $2^m - 1$. Thus, $(x^m - 1)|(x^n - 1)$; therefore, according to the previous statement, $m|n$. The condition is sufficient: if $m|n$ then $GF(2^m)$ is a subfield of $GF(2^n)$. Indeed, $m|n \quad \rightarrow \quad (2^m - 1)|(2^n - 1)$ and this implies that $(x^{2^m} - 1)|(x^{2^n} - 1)$, in other words that $GF(2^m)$ is a subfield of $GF(2^n)$.

□

The minimal polynomial $m_\gamma(x)$ is the lowest degree polynomial with binary coefficients that has $\gamma \in GF(2^m)$ as a zero.

**Proposition.** *The minimal polynomial in $GF(2^m)$ has the following properties:*

1. $m_\gamma(x)$ exists and is unique.
2. $m_\gamma(x)$ is irreducible.
3. If $p(x)$ is a polynomial with binary coefficients and if $p(\gamma) = 0$ then minimal polynomial $m_\gamma(x)$ divides $p(x)$.
4. The polynomial $g(x) = x^{2^m} - x$ is the product of different minimal polynomials $m_{\gamma_i}(x)$ with $\gamma_i \in GF(2^m)$.
5. $\deg m_\gamma(x) \leq m$. If $\gamma$ is a primitive element of $GF(2^m)$ then $\deg m_\gamma(x) = m$.

*Proof:* If $\gamma$ is a primitive element of $GF(2^m)$ then $\gamma^{2^m-1} = 1$ and $\gamma^{2^m} - \gamma = 0$. There exists a polynomial such that $p(\gamma) = 0$; the polynomial $p(x) = x^{2^m} - x$ satisfies the condition $p(\gamma) = 0$. This polynomial must be unique; otherwise, if two polynomials, $f_1(x)$ and $f_2(x)$ of the same degree existed, then their difference $g(x) = f_1(x) - f_2(x)$ would have: (i) $\gamma$ as a zero; (ii) a lower degree, $\deg g(x) < \deg f_1(x)$. This contradicts the assumption that $f_1(x)$ and $f_2(x)$ are minimal polynomials.

To show that $m_\gamma(x)$ is irreducible we assume that $m_\gamma(x) = p(x)q(x)$. Then $\deg[p(x)], \deg[q(x)] < \deg[m_\gamma(x)]$ and at least one of the two polynomials $p(x)$ and $q(x)$ must have $\gamma$ as a root. This contradicts again the assumption that $m_\gamma(x)$ is a minimal polynomial.

It is easy to see that the minimal polynomial $m_\gamma(x)$ divides any polynomial $p(x)$ with binary coefficients if $p(\gamma) = 0$. Indeed we can express $p(x) = m_\gamma(x)q(x) + r(x)$. But $p(\gamma) = 0$, thus $r(\gamma) = 0$. Therefore $r(x) = 0$.

To prove that the polynomial $g(x) = x^{2^m} - x$ is the product of minimal polynomials $m_{\gamma_i}(x)$ with $\gamma_i \in GF(2^m)$ we observe that all elements $a \in GF(2^m)$ are zeros of $g(x)$ as they satisfy the equation $a^{2^m-1} - a = 0$. Then the minimal polynomial $m_a(x)$ exists and it is unique; according to the property (4) $m_a(x)$ divides the polynomial $g(x)$.

Finally, to show that $\deg m_\gamma(x) \leq m$, we recall that $GF(2^m)$ can be viewed as an $m$-dimensional vector space over $GF(2)$; thus, the elements $1, \gamma^1, \gamma^2, \ldots, \gamma^{m-1}, \gamma^m$ are linearly dependent. This means that there are $m+1$ coefficients $g_i \in GF(2),\ 0 \leq i \leq m$ such that:

$$g_m\gamma^m + g_{m-1}\gamma^{m-1} + \ldots + g_1\gamma^1 + g_0\gamma^0 = 0.$$

Thus $\deg m_\gamma(x)$ can be at most $m$. When $\gamma$ is a primitive element of $GF(2^m)$ then $GF(2^m)$ must consist of $2^m - 1$ different elements, thus $1, \gamma^1, \gamma^2, \ldots, \gamma^{m-1}$ must be linearly independent. This shows that when $\gamma$ is a primitive element of $GF(2^m)$ then $\deg m_\gamma(x) = m$. $\square$

The following proposition allows us to introduce the generator matrix of an R-S code.

**Proposition.** *If $\mathcal{C}_g$ is a cyclic code with generator polynomial $g(x)$ and $h(x) = (x^n - 1)/g(x)$ then the dual code $\mathcal{C}_g^\perp$ has a generator polynomial $g^\perp(x) = x^{k-1}h(1/x)$ with $h(x) = \sum_{i=0}^{k-1} h_i x^i$. Thus, $g^\perp(x) = h^0 x^{k-1} + h_1 x^{k-1} + \ldots + h_{k-1}$.*

Let $\alpha$ be a primitive element of $GF(q)$. If $n$ is a divisor of $(q-1)$ and $\beta = \alpha^{(q-1)/n}$ then $\beta^n = 1$. Let $\mathcal{C}$ be the $[n, k, d]$ <u>Reed-Solomon code</u> obtained by evaluating polynomials of degree at most $(k-1)$, $p(x) \in \mathbb{P}_{k-1}[x]$ over $GF(2^m)$ at $x_i = \beta^{i-1},\ 1 \leq i \leq n$. Then $\mathcal{C}$ is a cyclic code over $GF(q)$ with generator polynomial

$$g(x) = (x - \beta)(x - \beta^2)\ldots(x - \beta^{n-k}).$$

*Proof:* If $c \in \mathcal{C}$ then $c = (p(\beta^0), p(\beta^1), p(\beta^2), \ldots, p(\beta^{n-2}), p(\beta^{n-1}))$. A cyclic shift of the codeword $c$ leads to the $n$-tuple:

$$\hat{c} = \left(p(\beta^{n-1}), p(\beta^0), p(\beta^1), \ldots, p(\beta^{n-3}), p(\beta^{n-2})\right).$$

Recall that a cyclic code is a cyclic subspace of a vector space and from the fact that $\hat{c} \in \mathcal{C}$ we conclude that $\mathcal{C}$ is a cyclic code. We now construct the parity-check matrix of the code and then show that this parity check matrix corresponds to a cyclic code with a generator polynomial having $\beta, \beta^2, \ldots, \beta^{n-k}$ as roots. First, we express $\hat{c}$ as

$$\hat{c} = \left(p_1(\beta^0), p_1(\beta^1), p_1(\beta^2), \ldots, p_1(\beta^{n-2}), p_1(\beta^{n-1})\right)$$

with $p_1(x) = p(\beta^{-1})$. According to the previous proposition the generator matrix of the dual code is an $(n-k) \times n$ matrix $\mathrm{H} = [h_{ij}]$ with $h_{ij} = p(1/(x_j)^i),\ 1 \leq i \leq n-k, 1 \leq j \leq n$; when

we substitute $x_j = \beta^{j-1}$ the generator matrix of the dual code which is at the same time the parity check matrix of the R-S code can be expressed as

$$\mathrm{H} = \begin{pmatrix} 1 & (\beta)^1 & (\beta)^2 & \dots & (\beta)^{n-2} & (\beta)^{n-1} \\ 1 & (\beta^2)^1 & (\beta^2)^2 & \dots & (\beta^2)^{n-2} & (\beta^2)^{n-1} \\ 1 & (\beta^3)^1 & (\beta^3)^2 & \dots & (\beta^3)^{n-2} & (\beta^3)^{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & (\beta^{n-k})^1 & (\beta^{n-k})^2 & \dots & (\beta^{n-k})^{n-2} & (\beta^{n-k})^{n-1} \end{pmatrix}.$$

Recall from Section 4.13 that when $g(x)$ is the generator polynomial of the cyclic $[n, k]$ code $\mathcal{C}_g$ over $GF(q)$ and when $g(x)$ has among its zeros $(d-1)$ elements of $GF(q)$, $\beta^\alpha, \beta^{\alpha+1}, \dots, \beta^{\alpha+d-2}$ with $\beta \in GF(q)$ the parity check matrix of cyclic code $\mathcal{C}_g$ is

$$\mathrm{H}_g = \begin{pmatrix} 1 & (\beta)^\alpha & (\beta)^{2\alpha} & (\beta)^{3\alpha} & \dots & (\beta)^{(n-1)\alpha} \\ 1 & (\beta)^{\alpha+1} & (\beta)^{2(\alpha+1)} & (\beta)^{3(\alpha+1)} & \dots & (\beta)^{(n-1)(\alpha+1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & (\beta)^{(\alpha+d-2)} & (\beta)^{2(\alpha+d-2)} & (\beta)^{3(\alpha+d-2)} & \dots & (\beta)^{(n-1)(\alpha+d-2)} \end{pmatrix}.$$

The similarity of H and $\mathrm{H}_g$ allows us to conclude that $\beta, \beta^2, \dots, \beta^{n-k}$ are the zeros of the generator matrix of the R-S code thus,

$$g(x) = (x - \beta)(x - \beta^2) \dots (x - \beta^{n-k}). \qquad \square$$

**Example.** *Construct the generator polynomial of the* $[7, 3]$ *R-S code* $\mathcal{C}$ *over* $GF(2^3)$ *with the primitive element* $\beta \in GF(2^3)$ *satisfying the condition* $\beta^3 + \beta + 1 = 0$.

The distance of the code is $d = n - k + 1 = 5$. We can describe $GF(8)$ by binary 3-tuples, $000, 001, \dots 111$, by listing its elements expressed as powers of the primitive element $\beta$, or by polynomials $p(x)$ of degree at most $k - 1 = 2$. For example, the tuple 011 corresponds to $\beta^3$ and to the polynomial $p(x) = x + 1$.

Table 17: The correspondence between the 3-tuples, the powers of $\beta$, and the polynomials.

| Binary 3-tuple | Power of primitive element | Polynomial |
|---|---|---|
| 000 | | 0 |
| 001 | $\beta^0$ | 1 |
| 010 | $\beta^1$ | $x$ |
| 100 | $\beta^2$ | $x^2$ |
| 011 | $\beta^3$ | $x + 1$ |
| 101 | $\beta^4$ | $x^2 + x$ |
| 110 | $\beta^5$ | $x^2 + x + 1$ |
| 111 | $\beta^6$ | $x^2 + 1$ |

Table 17 shows the correspondence between the 3-tuples, the powers of the primitive element, $\beta$, and the polynomials. Indeed, we observe that the polynomial $x^2 + x + 1$ is irreducible and thus can be used to construct $GF(8)$. In $GF(8)$:

$p_0(x) = x^0 = 1, \quad p_1(x) = x^1 = x, \quad p_2(x) = x^2 = x^2$
$p_3(x) = x^3 = x + 1$
$p_4(x) = x^4 = x(x^3) = x(x+1) = x^2 + x$
$p_5(x) = x^5 = x(x^4) = x(x^2+x) = x^3 + x^2 = x^2 + x + 1$
$p_6(x) = x^6 = x(x^5) = x(x^2+x+1) = x^3 + x^2 + x = x + 1 + x^2 + x = x^2 + 1.$

The generator polynomial of $\mathcal{C}$ and of its dual $\mathcal{C}^\perp$ when $d = 5$ are respectively:

$$g(x) = (x - \beta^0)(x - \beta^1)(x - \beta^2)(x - \beta^3) = (x - 1)(x - \beta)(x - \beta^2)(x - \beta^3)$$

and

$$g^\perp(x) = (x - \beta^{-4})(x - \beta^{-5})(x - \beta^{-6}) = (x - \beta^3)(x - \beta^2)(x - \beta^1) = (x - \beta)(x - \beta^2)(x - \beta^3).$$

We notice that $g(x)/g^\perp(x) = x - 1$ thus, the code $\mathcal{C}$ is self-dual.

**Burst-error correction and Reed-Solomon codes.** R-S codes are well suited for burst-error correction. The following proposition quantifies the burst error correcting capability of an $R - S$ code with symbols from $GF(2^m)$.

**Proposition.** *If $\mathcal{C}$ is an $[n, k]$ R-S code cover $GF(2^m)$ with $t = \lfloor (n-k)/2 \rfloor$ then the corresponding $[n \times m, k \times m]$ binary code is able to correct burst errors of length $(m \times t)$ bits.*

An R-S code using an alphabet with $m$-bit symbols could correct any burst of $[(t-1)\times m+1]$ bits errors with $t$ the number of symbols in error. The code can correct a burst of $(m \times t)$ bit errors, provided that the errors occur in consecutive symbols and start at a symbol boundary; the code could not correct any arbitrary pattern of $(m \times t)$ bit errors because such a pattern could be spread over $(m \times t)$ symbols each with one bit in error.

**Example.** *The* $[255, 246]$ *R-S code.* In this case a symbol is an 8-bit byte, $m = 8$, $n = 255$, and $k = 246$, Figure 103. We transmit bytes of information in blocks of 255 bits with 246 information bits in each block. The code has a minimum distance of $d_{min} = n - k + 1 = 10$ and can correct $t$ symbols in error:

$$t = \lfloor \frac{n-k}{2} \rfloor = \lfloor \frac{255 - 246}{2} \rfloor = \lfloor \frac{9}{2} \rfloor = 4.$$

The R-S code could correct all bursts of at most $(t-1) \times m + 1 = 3 \times 8 + 1 = 25$ consecutive bits in error, see Figures 103 (a) and (b), and some bursts of $t \times m = 4 \times 8 = 32$ bits in error, Figure 103 (c). It cannot correct any arbitrary pattern of 32 errors, Figure 103 (d).

**Bose-Chaudhuri-Hocquenghem (BCH) codes.** Given the prime power $q$, and integers $m$ and $d$, the <u>BCH code</u> $BCH_{q;m;d}$ is obtained as follows: let $n = q^m$ and $GF(q^m)$ be an extension of $GF(q)$ and let $C'$ be the (extended) $[n; n-(d-1); d]_{q^m}$ Reed-Solomon code obtained by evaluating polynomials of degree at most $(n-d)$ over $GF(q^m)$ at all the points of $GF(q^m)$. Then the code $BCH_{q;m;d}$ is the $GF(q)$-subfield subcode of $C'$.

A simple argument shows a BCH code has dimension at least $n - m(d-1)$ [413]. Indeed, every function from $GF(q^m)$ to $GF(q)$ is a polynomial over $GF(q)$ of degree at most $(n-1)$.

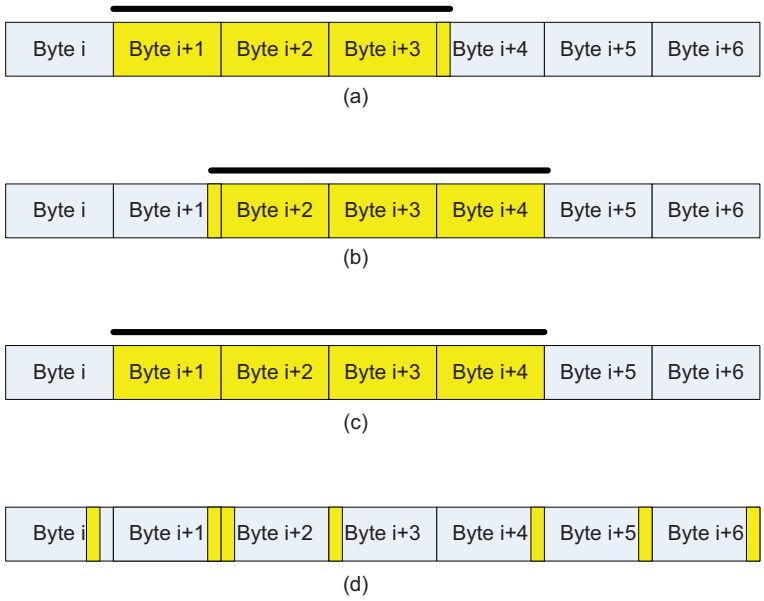



Figure 103: Correction of burst-errors with the $(255, 246)$ R-S code. The upper bar marks consecutive bits affected by error. The code can correct: (a), (b) any burst of 25 consecutive bit errors; (c) some bursts of 32 consecutive bit errors. The code cannot correct any arbitrary pattern of 32 bit errors, (d).

The space of polynomials from $GF(q^m)$ to $GF(q)$ is a linear space of dimension $n$. Now we ask what is the dimension of the subspace of polynomials of degree at most $(n-d)$. The restriction that a polynomial $p(x) = \sum_{i=0}^{n-1} c_i x^i$ has degree at most $(n-d)$ implies that the $c_i = 0, \quad i \in \{n-(d-1), \ldots, (n-1)\}$. Each such condition is a linear constraint over $GF(q^m)$ and this translates into a block of $m$ linear constraints over $GF(q)$. Since we have $(d-1)$ such blocks, the restriction that the functions have degree at most $(n-d)$ translates into at most $m(d-1)$ linear constraints. Thus, the resulting space has dimension at least $n - m(d-1)$.

BCH codes were independently discovered by Hocquenghem [209] and Bose and Ray-Chaudhuri [68]; the extension to the general q-ary case is due to Gorenstein and Zierler [173].

Next we examine codes used in radio and satellite communication to encode streams of data.

## 4.16 Convolutional Codes

Binary convolutional codes were introduced by Peter Elias in 1955 [141] as a generalization of block codes. Recall that in the case of a block code we encode $k$ information symbols into a block of $n$ symbols and the rate of the code is $R = k/n$. In some cases, for example in the case of R-S codes, we use one polynomial from a set of polynomials to generate one codeword. In our presentation we only discuss the encoding algorithm for convolutional codes; the Viterbi algorithm for decoding convolutional codes [440] is not analyzed here.

We consider an infinite sequence of blocks $\ldots b_{-2}, b_{-1}, b_0, b_1, b_2, \ldots$ with $k$ symbols per block. The input symbols are fed sequentially into an encoder which first creates a sequence of

information blocks $\ldots In_{-2}, In_{-1}, In_0, In_1, In_2, \ldots$ with $n$ symbols per block; then the encoder cyclically convolutes the information blocks with a generating vector $g = (g_0 g_1 \ldots g_m)$ with $g_0 = g_m = 1$ and produces the output $Out_j$; the rate of the cod is $R = k/n$.

To facilitate the presentation we consider an information vector consisting of $N$ bits or $NR$ information blocks obtained from $K$ input blocks; the output of the encoder is a codeword of $NR$ blocks or $N$ bits. The rate of the code is R=K/N=k/n. For example, in the binary case with $k = 1$ and $n = 2$ the rate is $R = 1/2$. If we choose $K = 8$ then the information vector and an output vector consist of $N = KR = 16$bits. A rule to construct the information vector could be

$$In = \{InBit_0, InBit_1, \ldots, InBit_{N-1}\} \quad \text{with} \quad InBit_j = \begin{cases} b_i & \text{if } j = \lfloor iR \rfloor \\ 0 & \text{otherwise.} \end{cases}$$

In this case each input block consists of one bit, each information block consists of two bits, the first equal to the corresponding bit of the information block and the second bit being a 0.

Given $N$, $R$, and $g = (g_0 g_1 \ldots g_m)$ with $m < N$ the convolutional code $\mathcal{C}$ encodes the information vector $In = \{In_0, In_1, \ldots, In_{N-1}\}$ as the $N$-vector $Out = \{Out_0, Out_1, \ldots Out_{N-1}\}$ with

$$Out_j = \sum_{i=0}^{m} g_i In_{j-i}, \quad 0 \leq j \leq N-1.$$

The sum is modulo 2 and all indices are modulo $N$. We recognize that an $N$-vector (codeword) results from the cyclic convolution of an information vector with the generating vector thus, the name convolutional code. The main difference between a convolutional encoder and a block code encoder is that the $j$-th code block $Out_j$ depends not only on the current $In_j$ but also on the earlier $m$ information blocks, $In_{j-1}, In_{j-2}, \ldots, In_{j-m}$.

**Example.** *Encode the frame $Fr = (11101010)$ from the input stream when $N = 16$, $R = 1/2$, and $g = (11.10.11)$.* To facilitate reading, in our examples the blocks of a codeword are separated by periods.

The process of codeword construction is illustrated in Figure 104. Each bit of the frame generates two bits of the information vector; Table 18 gives the expression of the individual bits of the codeword as the convolution of the generating vector and the information vector. Note that all indices are modulo $N = 16$ thus, $In_{-1} = In_{15}, In_{-2} = In_{14}, \ldots, In_{-5} = In_{11}$.

Given $N$, the length of a block, and $R$, the rate, the convolutional encoding algorithm discussed in this section generates a linear $[N, K]$ code iff $K = R \times N$. In this case there is a $K \times N$ generating matrix G of the code [228]. Then:

$$Out = Fr \cdot \mathrm{G}$$

The $K$ rows of the generator matrix are: the first row consists of the $m + 1$ bits of the generating vector followed by $N - (m + 1)$ zeros; rows 2 to $K$ are cyclic shifts of the previous row $1/R$ positions to the right.

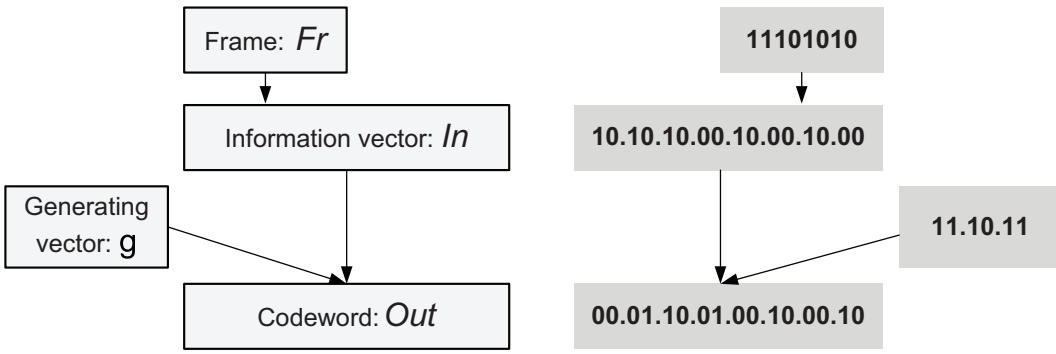



Figure 104: The construction of the codeword when $R = 1/2$, the frame size is $N/2 = 8$, the generating polynomial of degree $m = 5$ is $g = (11.10.11)$.

Table 18: The expression of the individual bits of the codeword when $Fr = (11101010)$, $N = 16$, $R = 1/2$, and $g = (11.10.11)$.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $Out_0 =$ | $g_0 \cdot In_0$ | $+g_1 \cdot In_{15}$ | $+g_2 \cdot In_{14}$ | $+g_3 \cdot In_{13}$ | $+g_4 \cdot In_{12}$ | $+g_5 \cdot In_{11}$ | $= 0$ |
| $Out_1 =$ | $g_0 \cdot In_1$ | $+g_1 \cdot In_0$ | $+g_2 \cdot In_{15}$ | $+g_3 \cdot In_{14}$ | $+g_4 \cdot In_{13}$ | $+g_5 \cdot In_{12}$ | $= 0$ |
| $Out_2 =$ | $g_0 \cdot In_2$ | $+g_1 \cdot In_1$ | $+g_2 \cdot In_0$ | $+g_3 \cdot In_{15}$ | $+g_4 \cdot In_{14}$ | $+g_5 \cdot In_{13}$ | $= 0$ |
| $Out_3 =$ | $g_0 \cdot In_3$ | $+g_1 \cdot In_2$ | $+g_2 \cdot In_1$ | $+g_3 \cdot In_0$ | $+g_4 \cdot In_{15}$ | $+g_5 \cdot In_{14}$ | $= 1$ |
| $Out_4 =$ | $g_0 \cdot In_4$ | $+g_1 \cdot In_3$ | $+g_2 \cdot In_2$ | $+g_3 \cdot In_1$ | $+g_4 \cdot In_0$ | $+g_5 \cdot In_{15}$ | $= 1$ |
| $Out_5 =$ | $g_0 \cdot In_5$ | $+g_1 \cdot In_4$ | $+g_2 \cdot In_3$ | $+g_3 \cdot In_2$ | $+g_4 \cdot In_1$ | $+g_5 \cdot In_0$ | $= 0$ |
| $Out_6 =$ | $g_0 \cdot In_6$ | $+g_1 \cdot In_5$ | $+g_2 \cdot In_4$ | $+g_3 \cdot In_3$ | $+g_4 \cdot In_2$ | $+g_5 \cdot In_1$ | $= 0$ |
| $Out_7 =$ | $g_0 \cdot In_7$ | $+g_1 \cdot In_6$ | $+g_2 \cdot In_5$ | $+g_3 \cdot In_4$ | $+g_4 \cdot In_3$ | $+g_5 \cdot In_2$ | $= 1$ |
| $Out_8 =$ | $g_0 \cdot In_8$ | $+g_1 \cdot In_7$ | $+g_2 \cdot In_6$ | $+g_3 \cdot In_5$ | $+g_4 \cdot In_4$ | $+g_5 \cdot In_3$ | $= 0$ |
| $Out_9 =$ | $g_0 \cdot In_9$ | $+g_1 \cdot In_8$ | $+g_2 \cdot In_7$ | $+g_3 \cdot In_6$ | $+g_4 \cdot In_5$ | $+g_5 \cdot In_4$ | $= 0$ |
| $Out_{10} =$ | $g_0 \cdot In_{10}$ | $+g_1 \cdot In_9$ | $+g_2 \cdot In_8$ | $+g_3 \cdot In_7$ | $+g_4 \cdot In_6$ | $+g_5 \cdot In_5$ | $= 1$ |
| $Out_{11} =$ | $g_0 \cdot In_{11}$ | $+g_1 \cdot In_{10}$ | $+g_2 \cdot In_9$ | $+g_3 \cdot In_8$ | $+g_4 \cdot In_7$ | $+g_5 \cdot In_6$ | $= 0$ |
| $Out_{12} =$ | $g_0 \cdot In_{12}$ | $+g_1 \cdot In_{11}$ | $+g_2 \cdot In_{10}$ | $+g_3 \cdot In_9$ | $+g_4 \cdot In_8$ | $+g_5 \cdot In_7$ | $= 0$ |
| $Out_{13} =$ | $g_0 \cdot In_{13}$ | $+g_1 \cdot In_{12}$ | $+g_2 \cdot In_{11}$ | $+g_3 \cdot In_{10}$ | $+g_4 \cdot In_9$ | $+g_5 \cdot In_8$ | $= 0$ |
| $Out_{14} =$ | $g_0 \cdot In_{14}$ | $+g_1 \cdot In_{13}$ | $+g_2 \cdot In_{12}$ | $+g_3 \cdot In_{11}$ | $+g_4 \cdot In_{10}$ | $+g_5 \cdot In_9$ | $= 1$ |
| $Out_{15} =$ | $g_0 \cdot In_{15}$ | $+g_1 \cdot In_{14}$ | $+g_2 \cdot In_{13}$ | $+g_3 \cdot In_{12}$ | $+g_4 \cdot In_{11}$ | $+g_5 \cdot In_{10}$ | $= 0$ |

**Example.** *The generator matrix of the convolutional code with* $N = 16$, $R = 1/2$, *and* $g = (11.10.11)$ is a $8 \times 16$ matrix. The first $m + 1 = 6$ elements of the first row are the elements of $g$, (111011), followed by $N - (m + 1) = 16 - 6 = 10$ zeros; the next $K - 1 = 7$ rows are cyclic shifts to the right of the previous row with two positions as $1/R = 2$.

$$
\mathrm{G} = \begin{pmatrix}
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{pmatrix}
$$

When $Fr = (11101010)$ it is easy to check that $Out = (00.01.10.01.00.10.00.10)$, a result consistent with the previous example. Indeed, $(11101010) \cdot \mathrm{G} = (00.01.10.01.00.10.00.10)$.

**Proposition.** *Every symbol encoded by a convolutional code depends on at most $M+1$ previous information symbols, where the integer $M = \lceil R \times (m+1) \rceil - 1$ is called the* <u>*memory*</u> *of the convolutional code.*

*Proof:* Recall that $Out_j = \sum_{i=0}^{m} g_i In_{j-i}$ thus, each symbol of the resulting $N$-vector is the result of the cyclic convolution of $(m+1)$ symbols of the information vector. Let us concentrate on the first of these $(m+1)$ symbols of the information vector, the symbol $In_j$ inherited from the $i$-th symbol of the frame, $Fr_i$, with $j = \lfloor i/R \rfloor$. Similarly, the information symbol $In_{j_{out}}$ is inherited from the symbol $Fr_{M+i+1}$ with $j_{out} = \lfloor (M+i+1)/R \rfloor$. It is easy to see that $Fr_{M+i+1}$ is outside of the range of the summation required by convolution. To show this we recall that when $a = \lceil b \times r \rceil$ and $r < 1$ then $\lfloor a/r \rfloor \geq b$; in our case $M+1 = \lceil R \times (m+1) \rceil$ with $R < 1$ thus, $\lfloor M+1/R \rfloor \geq M+1$. Recall that $j = \lfloor i/R \rfloor$ so we conclude that

$$
\lfloor \frac{M+i+1}{R} \rfloor \geq \lfloor \frac{i}{R} \rfloor + \lfloor \frac{M+1}{R} \rfloor \geq j + M + 1.
$$

It follows that $Out_j$ may only depend of at most $(M+1)$ symbols of the original frame.

□

The cyclic convolution rule formulated at the beginning of this section when used with relatively short frames justifies the name <u>tail-biting code</u> given to this class of codes. The encoding rule we discussed so far is too restrictive. Instead of a generating vector we can consider a periodic function $g$ alternating between $k$ generating vectors or polynomials; in this case the generating matrix consists of $k$ different types of rows.

An <u>$(n, k, M)$ convolutional code</u> is a sequence of block codes of length $N = q \times n$ encoded using the cyclic convolution rule; $q$ is an integer greater than 1. A <u>terminated $(n, k, M)$ convolutional code</u> is an $(n, k, M)$ convolutional code with the last $M$ input bits set to zero.

**Example.** *An encoder for a convolutional code with multiple generator polynomials.* The encoder consists of $m$ registers each holding one input bit, and $n$ modulo 2 adders. Each adder has a generator polynomial.

The encoded sequence of say $q$ input bits is of length $4q$ and represents a path in the state transition diagram. The diagram in Figure 106 shows that not all transitions are possible and gives us a hint on the decoding procedure for convolutional codes.

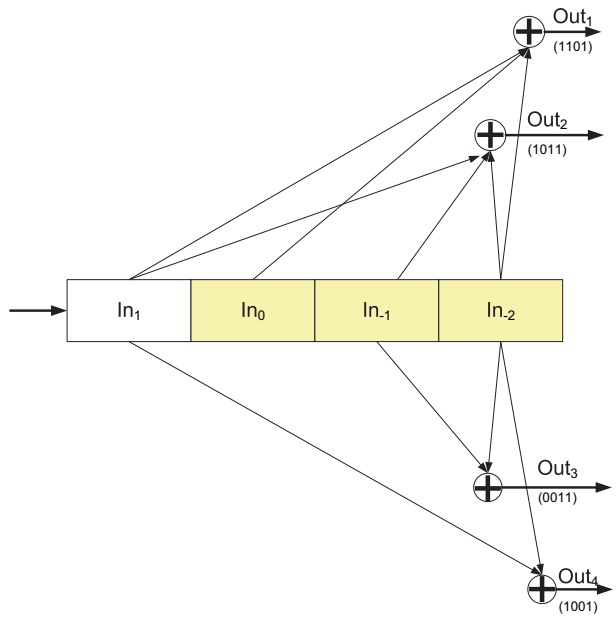



Figure 105: A convolutional non-recursive, non-systematic encoder with $k = 1$ and $n = 4$ and generator polynomials $g_1(x) = x^3 + x^2 + 1$, $g_2(x) = x^3 + x + 1$, $g_3(x) = x + 1$, and $g_4(x) = x^3 + 1$.

A sequence of symbols is valid only if it fits the state transition diagram of the encoder. If a sequence does not fit the state transition diagram of the encoder we have to identify the nearest sequence that fits the state transition diagram.

The encoder has $2^n$ states and operates as follows:

- initially the $m$ input registers are all set to 0;
- an input bit is fed to the $m$-bit register;
- the encoder outputs $n$ bits, each computed by one of the $n$ adders;
- the next input bit is fed and the register shifts one bit to the right;
- the next group of $n$ bits is calculated.

We present an encoder with $m = 1$, $n = 4$, and $k = 1$ in Figure 105; the encoder is a finite state machine and its transition diagram is shown in Figure 106.

The rate of the encoder is 1/4 and the encoder has 16 states. The four cells of the input register are labelled from left to right as: $In_1, In_0, In_{-1}$, and $In_{-2}$. The four generator polynomials are:

$$g_1(x) = x^3 + x^2 + 1, \quad g_2(x) = x^3 + x + 1, \quad g_3(x) = x + 1, \quad \text{and} \quad g_4(x) = x^3 + 1.$$

Thus, the four output symbols are calculated as follows:

$$\begin{array}{rcl|rcl} Out_1 & = & In_1 + In_0 + In_{-2} \mod 2 & Out_2 & = & In_1 + In_{-1} + In_{-2} \mod 2 \\ Out_3 & = & In_{-1} + In_{-2} \mod 2 & Out_4 & = & In_1 + In_{-2} \mod 2. \end{array}$$

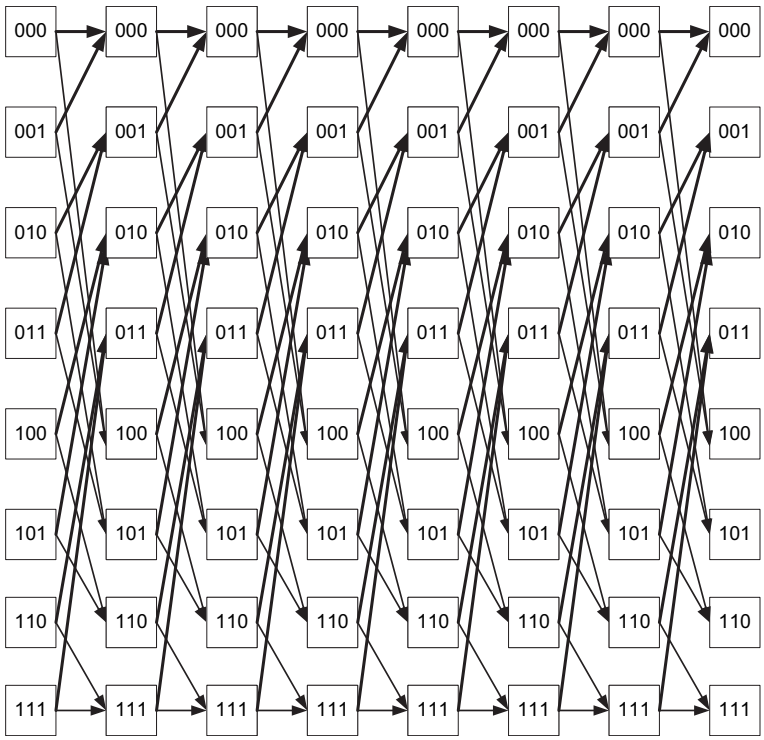



Figure 106: A trellis diagram describes the state transition of the finite state machine for the convolutional code encoder with k=1 and n=4 in Figure 105. For example, if the encoder is in state 010 ($In_0 = 0$, $In_{-1} = 1$, and $In_{-2} = 0$) and the input bit ($In_1$) is 0, then the next state is 001; if the input bit is 1, then the next state is 101. The diagram allows us to follow the path of eight input bits.

Table 19 summarizes the state transitions. The state is given by the contents of memory cells $In_0, In{-}1, In_{-2}$; the input is $In_1$. If at time $t$ the system is in state 010 ($In_0 = 0, In_{-1} = 1$ and $In_{-2} = 0$) and:

(a1) the input is $In_1 = 0$, then the next state of the encoder is 001 and the four output bits produced by the encoder are 0110 as

$$Out_1 = 0 + 0 + 0 = 0, \quad Out_2 = 0 + 1 + 0 = 1, \quad Out_3 = 1 + 0 = 1, \quad Out_4 = 0 + 0 = 0.$$

(b1) the input is $In_1 = 1$, then the next state of the encoder is 101 and the four output bits produced by the encoder are 1011 as

$$Out_1 = 1 + 0 + 0 = 1, \quad Out_2 = 1 + 1 + 0 = 0, \quad Out_3 = 1 + 0 = 1, \quad Out_4 = 1 + 0 = 1.$$

If at time $t + 1$ the system is in state 001 and

(a2) the input is $In_1 = 0$, then the next state of the encoder is 000 and the four output bits produced by the encoder are 1111 as

$$Out_1 = 0 + 0 + 1 = 1, \quad Out_2 = 0 + 0 + 1 = 1, \quad Out_3 = 0 + 1 = 1, \quad Out_4 = 0 + 1 = 1.$$

(b2) the input is $In_1 = 1$, then the next state of the encoder is 100 and the four output bits produced by the encoder are 0010 as

Table 19: The state transitions for the convolutional encoder.

| Time | Current state | Input | Input register | Output register | Next state |
|---|---|---|---|---|---|
| t | 010 | 0 | 0010 | 0110 | 001 |
| | | 1 | 1010 | 1011 | 101 |
| t+1 | 001 | 0 | 0001 | 1111 | 000 |
| | | 1 | 1001 | 0010 | 100 |
| | 101 | 0 | 0101 | 0111 | 010 |
| | | 1 | 1101 | 1010 | 110 |

$$Out_1 = 1 + 0 + 1 = 0, \quad Out_2 = 1 + 0 + 1 = 0, \quad Out_3 = 0 + 1 = 1, \quad Out_4 = 1 + 1 = 0.$$

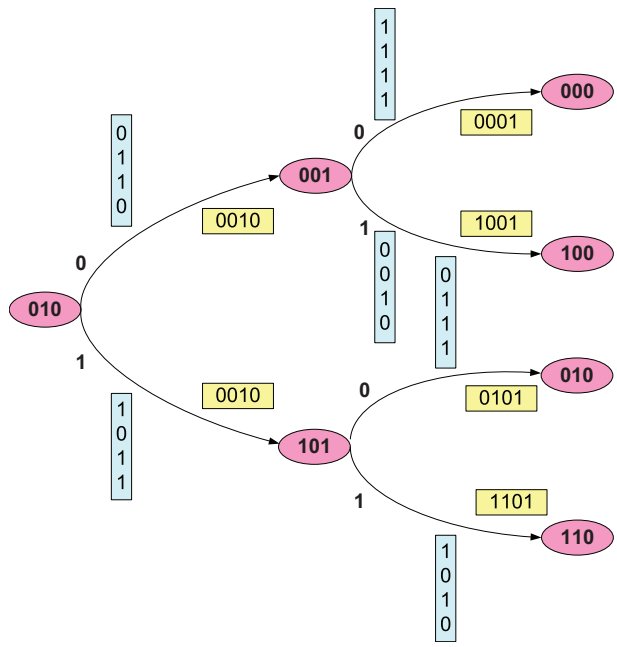



Figure 107: The state transition of the encoder in Figure 105 for a sequence of two input bits. The state is given by the contents of memory cells $In_0, In{-}1, In_{-2}$; the input is $In_1$. The vertical box next to a transition shows the output, $Out_1, Out2, Out_3, Out_4$, generated by the encoder and the horizontal box shows the contents of the four cells $In_1, In_0, In{-}1, In_{-2}$

Figure 107 shows the state transitions for a sequence of two input bits, the contents of the memory register, and the four output bits.

Convolutional codes are widely used in applications such as digital video, radio, mobile communication, and satellite communication. We next focus our attention to other families of codes able to encode large data frames and capable to approach the Shannon limit.

## 4.17 Product Codes

We first discuss encoding of large blocks of data and methods to construct composite codes from other codes. Then we survey some of the properties of product codes.

The question we address is how to encode large frames and decode them efficiently. Shannon channel encoding theorem tells us that we are able to transmit reliably at rates close to channel capacity, even though the number of errors per frame, $e = p \times N$, increases with the frame size, $N$, and with the probability of a symbol error, $p$. We also know that decoding complexity increase exponentially with $N$; indeed, after receiving a frame encoded as a codeword we check if errors have occurred and, if so, compute the likelihood of every possible codeword to correct the error(s).

A trivial solution is to split the frame into $N/n$ blocks of size $n$ and use short codes with small $n$. We distinguish *decoder errors,* the cases when an $n$-tuple in error is decoded as a valid codeword, from *decoding failures,* when the decoder fails to correct the errors; if we denote by $p_{err}$ probability of a decoding error and by $p_{fail}$ the probability of a decoding failure, then $P_{cF}$, the probability of a correct frame, and $P_{eF}$, the probability of a frame error, are:

$$P_{cF} = (1 - p_{err})^{\frac{N}{n}} \qquad \text{and} \qquad P_{eF} = 1 - (1 - p_{fail})^{\frac{N}{n}}.$$

The undetected errors have a small probability only if the error rate and the block size $n$ are small; but a small $n$ leads to a large $P_{eF}$. Thus, we have to search for other means to encode large frames. A possible solution is to construct codes from other codes using one of the following strategies:

- *Parity check.* Add a parity bit to an $[n, k, 2t-1]$ code to obtain an $[n+1, k, 2t]$ code.
- *Punctuating.* Delete a coordinate of an $[n, k, d]$ code to obtain an $[n-1, k, d-1]$ code.
- *Restricting.* Take the subcode corresponding to the first coordinate being "the most common element" and then delete the first coordinate: $[n, k, d] \quad \mapsto \quad [n-1, k-1, d]$.
- *Direct product.* Use two or more codes to encode a block of data.

The methods to construct codes from other codes weaken codes asymptotically, with one exception, the product codes discussed next.
A product code is a vector space of $n_1 \times n_2$ arrays. Each row of this array is a codeword of a linear $[n_1, k_1, d_1]$ code and each column is a codeword of the linear code $[n_2, k_2, d_2]$.

**Proposition.** *$\mathcal{C}$, the product code of $\mathcal{C}_1$ and $\mathcal{C}_2$, $[n_1, k_1, d_1]$ and $[n_2, k_2, d_2]$ linear codes, respectively, is an $[n_1 \times n_2, k_1 \times k_2, d_1 \times d_2]$ linear code.*

*Proof:* We arrange the information symbols into an $k_1 \times k_2$ array and compute the parity check symbols on rows and columns independently, using $\mathcal{C}_1$ for the rows and $\mathcal{C}_2$ for the columns; the parity checks on the parity checks computed either using $\mathcal{C}_1$ or $\mathcal{C}_2$ are the same due to linearity, Figure 108. It follows that $\mathcal{C}$ is an $[n_1 \times n_2, k_1 \times k_2]$ linear code.

The minimum distance codewords of $\mathcal{C}$ are those resulting from repeating a codeword of minimum weight equal to $d_1$ of $\mathcal{C}_1$ in at most $d_2$ rows; the weight of these codewords is $d_1 \times d_2$ and this is the distance of $\mathcal{C}$.

□

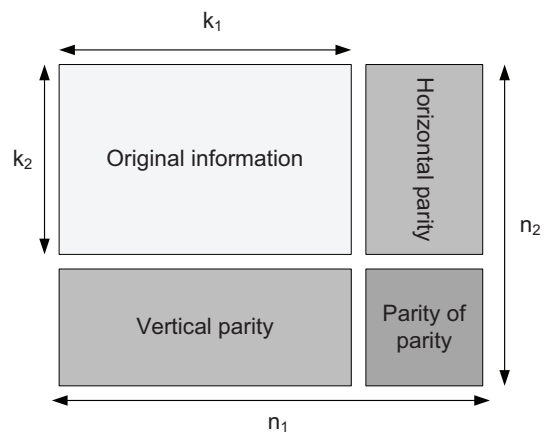



Figure 108: Schematic representation of product codes; horizontal, vertical and parity of parity symbols are added to the original information.

**Proposition.** *If $a_1$ and $a_2$ are the numbers of minimum weight codewords in $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively, then the product code $\mathcal{C}$ has $a_1 \times a_2$ codewords of minimum weight $d_1 \times d_2$.*

*Proof:* The minimum weight of any row is $d_1$ and there can be at most $a_1$ such rows; the minimum weight of any column is $d_2$ and there can be at most $a_2$ such columns. A minimum weight codeword of $\mathcal{C}$ is one corresponding to a choice of a row and of a column of minimum weight; there are only $a_1 \times a_2$ such choices of weight $d_1 \times d_2$.

Serial decoding of a product code $\mathcal{C}$: each column is decoded separately using the decoding algorithms for $\mathcal{C}_2$; then errors are corrected and the decoding algorithm for $\mathcal{C}_1$ is applied to each row.

**Proposition.** *When both $\mathcal{C}_1$ and $\mathcal{C}_2$ are $[n, k, d]$ codes then the resulting product code $\mathcal{C}$ is an $[n^2, k^2, d^2]$ linear code and can correct any error pattern of at most $e = (d^2 - 1)/4$ errors when serial decoding is used.*

*Proof:* $e = (d^2 - 1)/4 = [(d - 1)/2][(d + 1)/2]$; when $d = 2t + 1$ then $e = t(t + 1)$ and several columns affected by $t + 1$ errors will be decoded incorrectly; indeed, $\mathcal{C}_2$, the code used for encoding the columns, is an $[n, k, d]$ code and can correct at most $t$ errors. But, as the total number of errors is limited by $e = t(t + 1)$, we may have at most $t$ such columns affected by errors thus, the maximum number of errors left after the first decoding step in any row is $t$; these errors can be corrected by $\mathcal{C}_1$, the $[n, k, d]$ row code. $\square$

In the next section we introduce concatenated codes; we show that as the length of a codeword increases concatenated codes can be decoded more efficiently than other classes of composite codes.

## 4.18 Serially Concatenated Codes and Decoding Complexity

The concatenated codes were discovered by David Forney [154] at a time when the decoding problem was considered solved through (i) long, randomly chosen convolutional codes with

a dynamically, branching tree structure and (ii) sequential decoding, namely exclusive tree search techniques characterized by a probabilistic distribution of the decoding computations. The balance between code performance and decoding complexity was tilted in favor of code performance, measured by the number of errors corrected by the code.

Concatenated codes resulted from Forney's search for a class of codes having two simultaneous properties:

1. The probability of error decreases exponentially at all rates below channel capacity.
2. The decoding complexity increases only algebraically with the number of errors corrected by the code and, implicitly, with the length of a codeword.

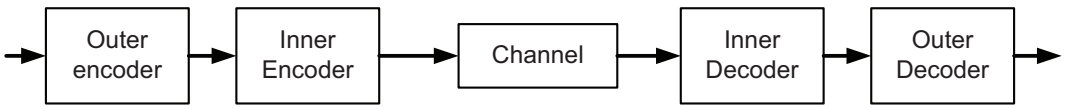



Figure 109: The original concatenated code was based on a relatively short random "inner" code and used maximum likelihood (ML) decoding to achieve a modest error probability, Prob($error$) $\approx 10^{-2}$, at a near capacity rate. A long, high-rate Reed-Solomon (R-S) code, and a generalized-minimum-distance (GMD) decoding scheme was used for the "outer" level.

Figure 109 illustrates the coding system proposed by Forney in his dissertation; a message is first encoded using an "outer" Reed-Solomon $[N, K, D]$ code over a large alphabet with $N$ symbols and then with an "inner" $[n, k, d]$ binary code with $k = \log_2 N$. The result is an $[N \times n, K \times k, D \times d]$ binary code, Figure 110. The decoding complexity of the code is dominated by the complexity of the algebraic decoder for the Reed-Solomon code while the probability of error decreases exponentially with $n$, the code length, at all rates below channel capacity.

Information transmission at Shannon rate is achievable using an efficient encoding and decoding scheme and Forney's concatenation method provides such an efficient decoding scheme. Consider a binary symmetric channel ($BSC_p$) with $p$ the probability of random errors. The following formulation of Shannon channel coding theorem [413] states that reliable information transmission through a binary symmetric channel is feasible at a rate

$$R = 1 - H(p) - \epsilon \qquad \text{with} \qquad \epsilon > 0.$$

**Theorem.** *There exist a code $\mathcal{C}$ with an encoding algorithm $\mathcal{E} : \{0,1\}^k \mapsto \{0,1\}^n$ and a decoding algorithm $\mathcal{D} : \{0,1\}^n \mapsto \{0,1\}^k$ with $k = (1 - H(p) - \epsilon) \times n$ such that:*

$$\text{Prob(decoding error for } BSC_p) < e^{-n}.$$

A corollary of this theorem discussed next shows that given a binary symmetric channel if we select an "outer" Reed-Solomon $(N, K, \epsilon N)$ code with $K = (1 - \epsilon)N$ and an "inner" binary $(n, k, d)$ code with $k = (1 - H(p) - \epsilon) \times n$, then we should be able to correct random errors occurring with probability $p$ and in this process limit the decoding failures by $e^{-n}$ when $k \geq log_2 N$.

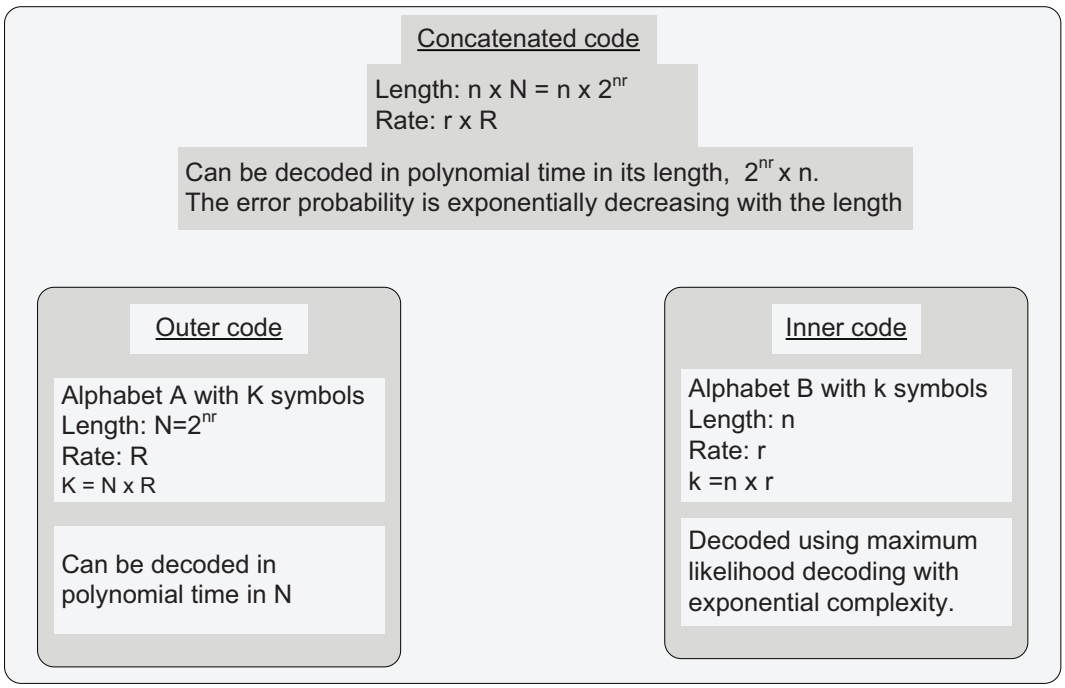



Figure 110: The "inner" and the "outer" codes of a concatenated code.

**Corollary.** *There exist a code $\mathcal{C}$ with an encoding algorithm $\mathcal{E}' : \{0,1\}^K \mapsto \{0,1\}^N$ and a decoding algorithm $\mathcal{D}' : \{0,1\}^N \mapsto \{0,1\}^K$ such that:*

$$\text{Prob(decoding error)} < \frac{1}{N}$$

*where $\mathcal{E}'$ and $\mathcal{D}'$ are polynomial time algorithms.*

*Proof:* We divide a message $m \in \{0,1\}^K$ into $K/k$ blocks of size $k = \log K$. Then we use Shannon's encoding function $\mathcal{E}$ to encode each of these blocks into words of size $n$ and concatenate the results, such that $N = (K/k)n$. The rate is preserved:

$$\frac{K}{N} = \frac{k}{n}.$$

It follows that

$$\text{Prob(decoding failure of } [\mathcal{E}', \mathcal{D}']_K) \leq \frac{K}{k}\text{Prob(decoding failure of } [\mathcal{E}, \mathcal{D}]_k).$$

This probability can be made smaller or equal to $\frac{1}{KN}$ by selecting $k = \textit{constant} \times \log K$ with a large enough constant. It is assumed that the encoding and decoding algorithms are exponential in time, by picking up a linear encoding function.

□

The resulting code has the following properties [413]:

(i) A slightly lower rate $R = (1 - H(p) - \epsilon)(1 - \epsilon) \geq 1 - H(p) - 2\epsilon$.

(ii) The running time is exponential in $k$ and polynomial in N, $\exp(k)\text{poly}(N)$.

(iii) The probability of decoding failure is bounded by $e^{-nN}$

$$\begin{aligned}
\text{Prob(decoding failure)} &\leq \text{Prob}(\tfrac{\epsilon N}{2} \text{ inner blocks leading to decoding failure}) \\
&\leq \binom{N}{\epsilon N/2} e^{-n\epsilon N/2} \\
&\leq e^{-nN}.
\end{aligned}$$

The decoding effort can be broken into the time to decode the R-S code, $\mathcal{O}(n^2)$ and the time to decode the "inner code" $n$ times (once for each "inner" codeword for $n \log n$ bits). Thus the decoding complexity for the concatenated code is $\mathcal{O}(n^2 \log n)$.

Concatenated codes with an interleaver placed between the outer and the inner encoder to spread out bursts of errors and with a deinterleaver between the inner and outer decoder have been used for space exploration since late 1970s. The serially concatenated codes have a very large minimum distance; they could be improved if the inner decoder could provide information to the outer decoder, but such a scheme would introduce an asymmetry between the two. The turbo codes discussed next address these issues.

## 4.19 Parallel Concatenated Codes - Turbo Codes

Turbo codes were introduced by Claude Berrou and his coworkers in 1993 as a refinement of concatenated codes [52]. The encoding structure of concatenated codes is complemented by an iterative algorithm for decoding; by analogy with a turbo engine, the decoder uses the processed output values as a posteriori input for the next iteration, thus the name "turbo," as shown in Figure 111. Turbo codes achieve small bit error rates at information rates much closer to the capacity of the communication channel than previously possible. If the rates of the inner and outer codes are $r$ and $R$, respectively, then the rate of a serially concatenated code is $R_s = r \times R$ while the rate of a parallel concatenated code is

$$R_p = \frac{r \times R}{1 - (1-r)(1-R)}.$$

We start the discussion of turbo codes with a review of <u>likelihood functions</u>. Consider a random variable $X$ with the probability density function $p_X(x)$; the result of an observation of the random variable $X$ should be classified in one of $M$ classes. Bayes's theorem expresses the "a posteriori probability (APP)" of a decision $d$ to classify the observation of $X$ in a class $i$, $d = i$, $1 \leq i \leq M$ conditioned by an observation $x$ as

$$\text{Prob}(d = i|x) = \frac{p_X[x|(d=i)]\text{Prob}(d=i)}{p_X(x)} \quad 1 \leq i \leq M,$$

where

$$p_X(x) = \sum_{i=1}^{M} p_X[x|(d=i)]\text{Prob}(d=i).$$

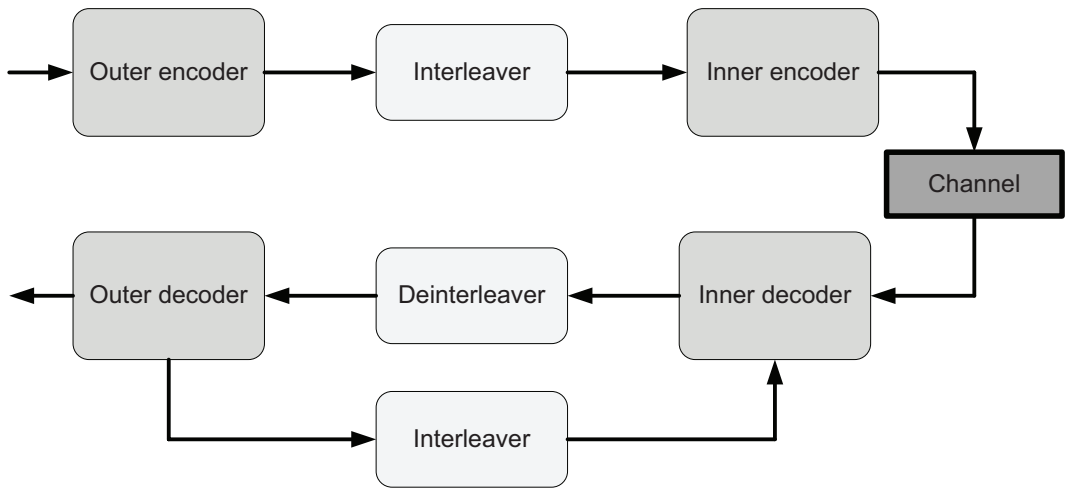



Figure 111: The feedback from the outer to the inner decoder of a turbo code.

Two decisions rules are widely used, *ML, maximum likelihood*, Figure 112, and *MAP, maximum a posteriori*.

When $d$ is a binary decision, $d = \{0, 1\}$,

$$p_X(x) = p_X[x \mid (d = 0)\text{Prob}(d = 0) + p_X[x \mid (d = 1)\text{Prob}(d = 1)$$

and the <u>maximum a posteriori decision</u> is a minimum probability error rule formulated as:

$$\text{Prob}(d = 0|x) > \text{Prob}(d = 1|x) \Longrightarrow d = 0 \quad \text{and} \quad \text{Prob}(d = 1|x) > \text{Prob}(d = 0|x) \Longrightarrow d = 1.$$

These conditions can be expressed as:

$$\frac{p_X[x|(d = 1)]}{p_X[x|(d = 0)]} < \frac{\text{Prob}(d = 0|x)}{\text{Prob}(d = 1|x)} \quad \Longrightarrow d = 0 \quad \text{and} \quad \frac{p_X[x|(d = 1)]}{p_X[x|(d = 0)]} > \frac{\text{Prob}(d = 0|x)}{\text{Prob}(d = 1|x)} \quad \Longrightarrow d = 1.$$

Next we introduce the *log-likelihood ratios, LLR*, real numbers used by a decoder; the absolute value of LLR quantifies the reliability of the decision and represents the "soft" component of the decision, while the sign selects either the value 0 or the value 1 and represents the "hard" component of the decision. The LLR, the logarithm of ratio of probabilities required by the MAP rule is

$$L(d|x) = \log\left[\frac{\text{Prob}(d = 1|x)}{\text{Prob}(d = 0|x)}\right] = \log\left[\frac{p_X[x|(d = 1)]\text{Prob}(d = 1)}{p_X[x|(d = 0)]\text{Prob}(d = 0)}\right],$$

or

$$L(d|x) = \log\left[\frac{p_X[x|(d = 1)]}{p_X[x|(d = 1)]}\right] + \log\left[\frac{\text{Prob}(d = 1)}{\text{Prob}(d = 0)}\right].$$

We rewrite the previous equality as

$$L'_X(\hat{d}) = L_{channel}(x) + L_X(d).$$

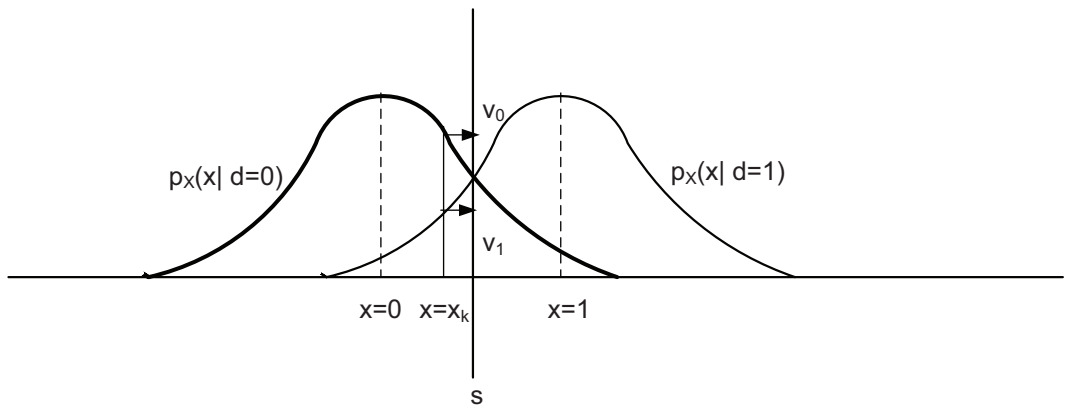



Figure 112: The probability distribution functions $p_X[x|(d = 0)]$ and $p_X[x|(d = 1)]$ for a continuous random variable $X$ and a binary decision. The decision line is $s$. In case of maximum likelihood we decide that $x_k$, the value of the random variable $X$ observed at time $k$, is 0 ($d = 0$) because $v_0 > v_1$, $x_k$ falls on the left of the decision line.

In this expression $L_{channel}(x)$ is the result of a measurement performed on the output of the channel and $L_X(d)$ is based on the a priori knowledge of the input data.

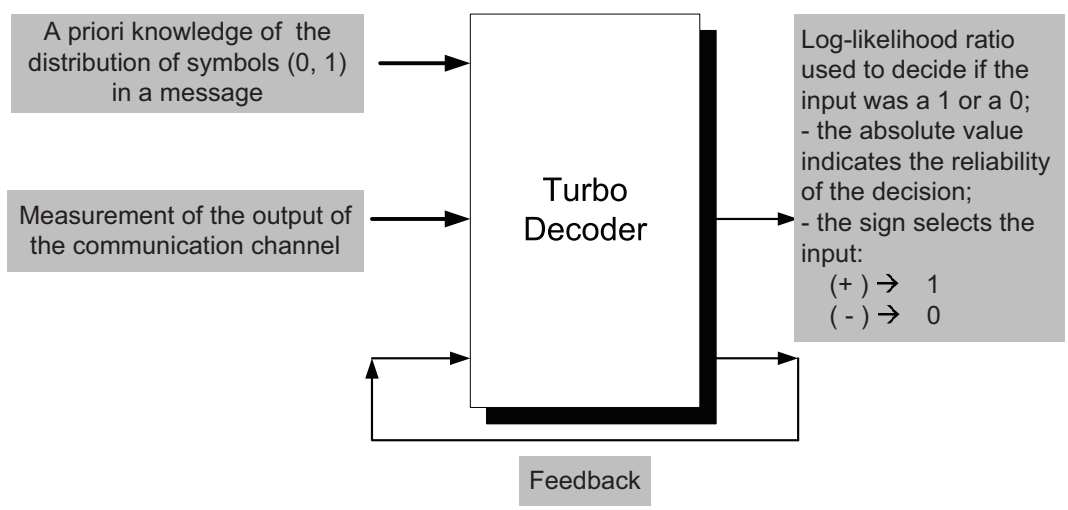



Figure 113: The decoder of a turbo code uses feedback to complement a priori information and the result of output channel measurement.

A turbo-decoder is at the heart of the scheme discussed in this section. As we can see in Figure 113 the decoder needs as input:

- A channel measurement,
- The a priory knowledge about the data transmitted, and
- Feedback from the previous iteration, called "extrinsic" information.

The soft output of the decoder, denoted as $L(\hat{d})$ will consist of two components, $L'_X(\hat{d})$, the LLR representing the soft decision of the decoder, and $L_{extrn}(\hat{d})$, the feedback reflecting the knowledge from the previous iteration of the decoding process:

$$L(\hat{d}) = L'_X(\hat{d}) + L_{extrn}(\hat{d}) = L_{channel}(x) + L_X(d) + L_{extrn}(\hat{d}).$$

We illustrate these ideas with an example for a two-dimensional code from [394]. An $[n_1, k_1]$ code is used for horizontal encoding and an $[n_2, k_2]$ code is used for vertical encoding see Figure 114.

At each step we compute also the horizontal and vertical extrinsic LLRs $L_{extrnHoriz}$ and $L_{extrnVert}$, respectively. The decoding algorithm consists of the following steps:

1. If a priori probability is available set $L(d)$ to that value; if not set $L(d) = 0$.
2. Decode horizontally and set $L_{extrnHoriz}(\hat{d}) = L(\hat{d}) - L_{chan}(x) - L(d)$.
3. Set $L(d) = L_{extrnHoriz}(\hat{d})$.
4. Decode vertically and set $L_{extrnVert}(\hat{d}) = L(\hat{d}) - L_{chan}(x) + L(d))$
5. Set $L(d) = L_{extrnVert}(\hat{d})$
6. Repeat steps (2) to (5) until a reliable decision can be made then provide the soft output $L(\hat{d}) = L_{chan}(x) + L_{extrnHoriz}(\hat{d}) + L_{extrnVert}(\hat{d})$.

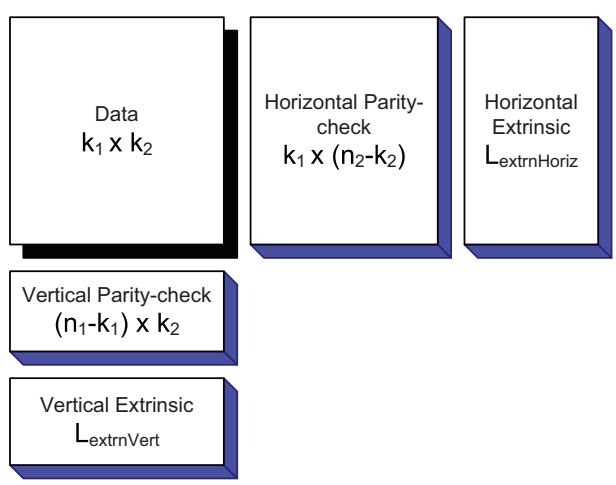



Figure 114: Schematic representation of a two-dimensional turbo code constructed from a horizontal $[n_1, k_1]$ code and a vertical $[n_2, k_2]$ code; the data is represented by a $k_1 \times k_2$ rectangle, the horizontal parity check bits by a $k_1 \times (n_2 - k_2)$ rectangle, and the vertical parity check bits by a $k_2 \times (n_1 - k_1)$ rectangle. The horizontal extrinsic and the vertical extrinsic LLR bits computed at each step are shown.

To verify that indeed the reliability of a decision made using two independent random variables is the minimum of the absolute value of individual LLRs and the "hard" component

of the decision is given by the signs of the two LLR we consider symbol-by-symbol MAP estimation of a set of symbols $c_k$ of a transmitted vector of symbols $c$ encoded, transmitted through a communication channel subject to a Gauss-Markov process[37] according to the distribution $p_{X|Y}(r|c)$ , and finally received as a vector $v = (v_1, v_2, \ldots, v_k, \ldots)$.

To distinguish between two distinct random variables, instead of $L(d)$ we use the notation $L_X(d)$ and assume that:

$$P_X(x=0) = \frac{e^{-L_X(d)}}{1+e^{-L_X(d)}} \quad \text{and} \quad P_X(x=1) = \frac{e^{L_X(d)}}{1+e^{L_X(d)}}.$$

If $X_1$ and $X_2$ are independent binary random variables then

$$L_{X_1,X_2}(d_1 \oplus d_2) = \log \frac{1+e^{L_{X_1}(d_1)}e^{L_{X_2}(d_2)}}{e^{L_{X_1}(d_1)}+e^{L_{X_2}(d_2)}}$$

with $\oplus$ the addition modulo two. This expression becomes

$$L_{X_1,X_2}(d_1 \oplus d_2) \approx \text{sign}[L_{X_1}(d_1)] \cdot \text{sign}[L_{X_2}(d_2)] \cdot \min[\mid L_{X_1}(d_1) \mid, \mid L_{X_2}(d_2) \mid \ ].$$

This brief introduction of turbo codes concludes our presentation of classical error correcting codes.

## 4.20 History Notes

While Claude Shannon was developing the information theory, Richard Hamming, a colleague of Shannon's at Bell Labs, understood that a more sophisticated method than the parity checking used in relay-based computers was necessary. Hamming realized the need for error correction and in the early 1950s he discovered the single error correcting binary Hamming codes and the single error correcting, double error detecting extended binary Hamming codes. Hamming's work marked the beginning of coding theory; he introduced fundamental concepts of coding theory such as Hamming distance, Hamming weight, and Hamming bound.

In 1955, Peter Elias introduced probabilistic convolutional codes with a dynamic structure pictured at that time as a branching tree [141] and in 1961, John Wozencraft and Barney Reiffen proposed *sequential decoding* based on exhaustive tree search techniques for long convolutional codes [463]. The probability that the number of operations required for the sequential decoding exceeds a threshold $D$ for an optimal choice of algorithm parameters is a Pareto (algebraic) distribution:

$$\text{Prob(Number of Operations} > D) \approx D^{-\alpha(R)}$$

with $\alpha(R) > 1$ and $R$ the transmission rate; this implies that the number of operations is bounded when $R < R_0$ and $R_0$ was regarded as the "practical capacity" of a memoryless

[37] A Gauss-Markov process is a stochastic process $X(t)$ with three properties: (i) If $h(t)$ is a non-zero scalar function of t, then $Z(t) = h(t)X(t)$ is also a Gauss-Markov process. (ii)If $f(t)$ is a non-decreasing scalar function of t, then $Z(t) = X(f(t))$ is also a Gauss-Markov process. (iii) There exists a non-zero scalar function $h(t)$ and a non-decreasing scalar function $f(t)$ such that $X(t) = h(t)W(f(t))$, where $W(t)$ is the standard Wiener process.

channel [155]. In 1963, Robert Fano developed an efficient algorithm for sequential decoding [145].

Work on algebraic codes started in late 1950s with codes over Galois fields [68] and the Reed-Solomon codes [354]. This line of research continued throughout the decade of 1960s with the development of cyclic codes including the BCH codes, and the publication of the seminal book of Elwyn Berlekamp [50].

In the early 1960s the center of gravity in the field of error correcting codes moved to MIT. The information theory group at MIT was focused on two problems: (a) how to reduce the probability of error of block codes, and (b) how to approach channel capacity. In 1966, Robert Gallagher [167] showed that when a randomly chosen code and maximum-likelihood decoding are used, then the probability of error decreases exponentially with the block length:

$$\text{Prob}(E) \approx e^{-NE(R)}$$

with $R$ the transmission rate, $C$ the cannel capacity, and $E(R) > 0$ the function relating the error and the transmission rate. The probability of error decreases algebraically with $\mathcal{D}$, the decoding complexity:

$$\text{Prob}(E) \approx \mathcal{D}^{-E(R)/R}$$

In mid 1960s David Forney introduced concatenated codes [154]. At the time when concatenated coded were introduced the balance between code performance and decoding complexity was tilted in favor of code performance, measured by the number of errors corrected by the code while decoding complexity was largely ignored. The original concatenated code was based on a relatively short random "inner code." A long, high-rate Reed-Solomon code, and a generalized-minimum-distance (GMD) decoding scheme was used for the "outer" level. The decoding complexity of this code is dominated by the complexity of the algebraic decoder for the Reed-Solomon code while the probability of error decreases exponentially with $n$, the code length, at all rates below channel capacity.

In his 1995 Shannon Lecture Forney writes: "I arrived at MIT in early sixties, during what can be seen in retrospect as the first golden age of information theory research....It was well understood by this time that the key obstacle to practically approaching channel capacity was not the construction of specific good long codes, although the difficulties in finding asymptotically good codes were already apparent (as expressed by the contemporary folk theorem *All codes are good except those we know of* [463]). Rather it was the problem of decoding complexity."

A major contribution to the field of error correcting codes is the introduction in 1993 of turbo codes. Turbo codes discovered by Claude Berrou and his coworkers [52] are a refinement of concatenated codes; they use the encoding structure of concatenated codes and an iterative algorithm for decoding.

## 4.21 Summary and Further Readings

*Coding Theory* is concerned with the process of increasing the redundancy of a message by packing a number $k$ of information symbols from an alphabet $\mathcal{A}$ into longer sequences of $n > k$ symbols. A *block code* of length $n$ over the alphabet $\mathcal{A}$ is a set of $M$, $n$-tuples where

each $n$-tuple takes its components from $\mathcal{A}$ and is called a codeword. We call the block code an $[n, M]$ code over $\mathcal{A}$.

Linear algebra allows an alternative description of a large class of $[n, k]$ block codes. A linear code $\mathcal{C}$ is a $k$-dimensional subspace, $V_k(F)$ of an $n$-dimensional vector space, $V_n$, over the field $F$. An $[n, k]$ linear code, $\mathcal{C}$, is characterized by a $k \times n$ matrix whose rows forms a vector space basis for $\mathcal{C}$. This matrix is called the *generator matrix* of $\mathcal{C}$. Once we select a set of $k$ linearly independent vectors in $V_n$ we can construct a generator matrix G for $\mathcal{C}$. Then the code $\mathcal{C}$ can be constructed by multiplying the generator matrix G with message vectors (vectors with $k$ components, or k-tuples). Given two $[n, k]$ linear codes over the filed $F$, $\mathcal{C}_1$ and $\mathcal{C}_2$ with generator matrices $\mathrm{G}_1$ and $\mathrm{G}_2$ respectively, we say that they are equivalent if $\mathrm{G}_2 = \mathrm{G}_1 P$ with $P$ a permutation matrix.

Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$ the orthogonal complement, or the *dual of code* $\mathcal{C}$, denoted as $\mathcal{C}^\perp$ consists of vectors (or $n$-tuples) orthogonal to every vector in $\mathcal{C}$. Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$, its dual, $\mathcal{C}^\perp$, is an $[n, n-k]$ linear code over the same field, $F$. If $\mathrm{G} = [I_k A]$ is a generator matrix for $\mathcal{C}$, or for a code equivalent to $\mathcal{C}$, then $\mathrm{H} = [-A^T I_{n-k}]$ is a generator matrix for its dual code, $\mathcal{C}^\perp$.

Given an $[n, k]$-linear code $\mathcal{C}$ over the field $F$ let H be the generator matrix of the dual code $\mathcal{C}^\perp$. Then $\forall c \in V_n \;\; \mathrm{H}c^T = 0 \;\; \Leftrightarrow \;\; c \in \mathcal{C}$. The matrix H is called the *parity-check matrix* of $\mathcal{C}$. The error syndrome $s$ of an $n$-tuple $v$ is defined as $s^T = \mathrm{H}v^T$. There is a one-to-one correspondence between the error syndrome and the bits in error, thus, the syndrome is used to determine if an error has occurred and to identify the bit in error.

Ideally, we want codes with a large number of codewords (thus, large value of $k$), which can correct as many errors, $e$, as possible, and with the shortest possible length of a codeword, $n$. These contradictory requirements cannot be satisfied concurrently. Several bounds for linear codes exist. Let $\mathcal{C}$ be an $[n, k]$ code over $GF(q)$ capable of correcting $e$ errors. Then $\mathcal{C}$ satisfies the Hamming bound

$$q^n \geq q^k \sum_{i=0}^{e} \binom{n}{i} (q-1)^i.$$

Every $[n, k, d]$ linear code $\mathcal{C}$ over the field $GF(q)$ satisfies the Singleton bound $k + d \leq n + 1$.

If $\mathcal{C}$ is an $[n, k, d]$ linear code over $GF(q)$ and if $M(n, d) =\mid \mathcal{C} \mid$ is the largest possible number of codewords for a given $n$ and $d$ values, then the code satisfies the Gilbert-Varshamov bound

$$M(n, d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}.$$

If $\mathcal{C}$ is an $[n, k, d]$-linear code and if $2d > n$ then the code satisfies the Plotkin bound

$$M \leq 2 \lfloor \frac{d}{2d-n} \rfloor$$

with $M =\mid \mathcal{C} \mid$.

A $q$-ary Hamming code of order $r$ over the field $GF(q)$ is an $[n, k, 3]$-linear code with $n = q^r - 1/q - 1$ and $k = n - r$. The parity-check matrix of the code $\mathrm{H}_r$ is an $r \times n$ matrix such that it contains no all zeros column and no two columns are scalar multiples of each other. The Hamming code of order $r$ over $GF(q)$ is a perfect code $\mathcal{C}_r$.

The first-order Reed-Muller code, $R(1,r)$ is a binary code with the generator matrix

$$\mathrm{G} = \begin{pmatrix} \mathbf{1} & 1 \\ \mathrm{H}_r & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ B_r \end{pmatrix}.$$

A cyclic subspace $\mathcal{C}$ of an $n$-dimensional vector space $V_n$ over the field $F$ is a set of vectors $\{c\} \in \mathcal{C}$ with the property:

$$c = (c_0 c_1 c_2 \ldots c_{n-2} c_{n-1}) \in \mathcal{C} \quad \Longrightarrow \quad c' = (c_{n-1} c_0 c_1 \ldots c_{n-2}) \in \mathcal{C}.$$

A linear code $\mathcal{C}$ is a cyclic code if $\mathcal{C}$ is a cyclic subspace. There is a one-to one correspondence between the cyclic subspaces of an $n$-dimensional vector space $V_n$ over the field of $F$ and the monic polynomials $g(x) \in F[x]$ which divide $f(x) = x^n - 1$, with $F[x]$ a ring of polynomials. If

$$f(x) = \prod_{i=1}^{q} g_i^{a_i}(x)$$

with $a_i$ positive integers and $g_i^{a_i}(x), 1 \leq i \leq q$ are distinct irreducible monic polynomials then $V_n$ contains

$$Q = \prod_{i=1}^{q} (a_i + 1)$$

cyclic subspaces. If $g(x)$ is a monic polynomial of degree $n-k$ and it divides $f(x) = x^n - 1$, then $g(x)$ is the generator polynomial of a cyclic subspace of $V_n$. The dimension of this cyclic subspace is equal to $k$.

Let $g(x)$ be the generator polynomial of the cyclic $[n,k]$ code $\mathcal{C}$ over $GF(q)$. Let $\beta \in GF(q)$ be a primitive element of order $n$. If $g(x)$ has among its zeros $\beta^{\alpha}, \beta^{\alpha+1}, \ldots, \beta^{\alpha+d-2}$ then the minimum distance of the cyclic code is $d$.

Codes designed to handle *bursts* of errors are called *Burst Error Correcting* codes (BEC). A burst of length $l$ is an $n$-tuple whose non-zero symbols are confined to a span of no fewer than $l$ symbols. A block code can correct all bursts of length $\leq l$ if and only if no two codewords differ by the sum of two bursts of length less or equal to $l$. The burst error correcting ability $l$ of an $[n,k]$ linear block code satisfies the Rieger Bound:

$$l \leq \frac{n-k}{2}$$

*Interleaving* is a simple technique to ensure burst error correction capabilities of a code originally designed for random error correction, or to enhance the capabilities of an existing burst error correcting code. The basic idea of interleaving is to alternate the transmission of the individual bits of a set of codewords.

Reed-Solomon (R-S) codes, are non-binary cyclic codes widely used for burst error correction. Non-binary codes allow a lower density of codewords (the ratio of the number of codewords to the number of $n$-tuples), thus, a larger distance between codewords, that, in turn, means increased error correction capability.

Call $\mathbb{P}_k$ the set of polynomials of degree at most $k$ over the finite filed $GF(q)$; $\mathbb{P}_k \subset GF(q)[x]$ is a vector space of dimension $k$. Consider $n > k$ elements of the finite field $x_1, x_2, \ldots \ldots x_n \in GF(q)$. Call $f_i(x)$ the generator of the codeword $c_i$ with

$$c_i = (f_i(x_1), f_i(x_2), f_i(x_3), \ldots f_i(x_{n-1}), f_i(x_n))$$

The $[n, k]$ Reed-Solomon (R-S) code consists of all codewords generated by polynomials $f_i \in \mathbb{P}_k$. An $[n, k]$ R-S code can achieve the Singleton bound, the largest possible code distance (minimum distance between any pair of codewords) of any linear code, $d = n - k + 1$; this means that the code can correct at most $t$ errors with:

$$t = \lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{n-k}{2} \rfloor.$$

Concatenated codes have two simultaneous properties: (i) the probability of error decreases exponentially at all rates below channel, capacity; and (ii) the decoding complexity increases only algebraically with the number of errors corrected by the code and implicitly the length of a codeword.

Information transmission at Shannon rate is achievable using an efficient encoding and decoding scheme. Forney's concatenation method provides such an efficient decoding scheme. In this method a message is first encoded using an "outer" R-S code $[N, K, N - K]$ over a large alphabet with $N$ symbols and an "inner" $[n, k, d]_2$ binary code with $k = \log_2 N$. The result is an $[Nn, Kk, Dd]_2$ binary code. The decoding is block by block. It is shown that there exist an encoding algorithm $\mathcal{E}' : \{0, 1\}^K \mapsto \{0, 1\}^N$ and a decoding algorithm $\mathcal{D}' : \{0, 1\}^N \mapsto \{0, 1\}^K$ with similar parameters such that Prob(decoding error) $< 1/N$ where $\mathcal{E}'$ and $\mathcal{D}'$ are polynomial time algorithms.

Turbo codes are a refinement of concatenated codes and achieve small bit error rates at information rates much closer to the capacity of the communication channel than previously possible.

There is an extensive literature on error correcting codes. The papers by Elias [141], Reed and Solomon [354], Bose [68], Wozencraft and Reiffen [463], Fano [145], Gallagher [167], Berrou, Glavieux, and Thitimajshima [52] reflect best the evolution of ideas in the field of error correcting codes, they are the original sources of information for more recently discovered codes. The second edition of the original text by Peterson [322] provides a rigorous discussion of error correcting codes. The text by MacWilliams and Sloane [281] is an encyclopedic guide to the field of error correcting codes. We also recommend the book by Berlekamp [50] for its treatment of cyclic codes, BCH codes, as well as Reed-Solomon codes. The book by Forney [154] covers the fundamentals of concatenated codes. More recent books, [228] and [436], provide a rather accessible introduction to error correction. The comprehensive 2003 book by MacKay, "Information Theory, Inference, and Learning Algorithms" is also a very good reference for error correcting codes.

## 4.22 Exercises and Problems

**Problem 1.** Show that there exists a characteristic element in $GF(q)$. Prove first that if $a \in GF(q), a \neq 0$ then the order of $a$ divides $q - 1$.

**Problem 2.** Show that $a^q = a$ if and only if $a \in GF(q)$

**Problem 3.** Show that if $\beta$ is a characteristic element of $GF(q)$ then $\forall a, b \in GF(q)$ we have $(a+b)^\beta = a^\beta + b^\beta$.

**Problem 4.** Construct $GF(3^2)$ using the irreducible polynomial $f(x) = x^2 + x + 2$ and display its addition and multiplication tables.

**Problem 5.** Construct $GF(2^4)$ using the irreducible polynomial $f(x) = x^4 + x^3 + x^2 + x + 1$ and display its addition and multiplication tables.

**Problem 6.** Prove that given a binary $[n, k]$ linear code $\mathcal{C}$, the number of $n$-tuples at distance $i$ from a codeword with weight $w$ and at distance $j$ from the decoded codeword is given by

$$T(i, j, w) = \binom{w}{i-k}\binom{n-w}{k}$$

with $k = 1/2(i + j - w)$ when $(i + k - w)$ is even and when $w - j \leq i \leq w + j$; otherwise $T = 0$.

**Problem 7.** Consider an $[n, k, d]$ linear code $\mathcal{C}$ and a binary symmetric channel; let $p$ be the probability of a symbol being in error. Prove that in case of bounded distance decoding the probability of a decoding failure, $P_{fail}$ and of decoding error, $P_{err}$ are:

$$p_{fail} = 1 - \sum_{i=0}^{e} \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=e+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

and

$$p_{err} = \sum_{w>0} \sum_{i=w-e}^{w+e} \sum_{j=0}^{e} A_w T(i, j, w) p^i (1-p)^{n-i}.$$

with $A_w$ the number of codewords of code $\mathcal{C}$ of weight equal to $w$.

**Problem 8.** Let $\mathcal{C}$ be an $[n, k]$ linear code and $\mathcal{C}^\perp$ its dual. Show that:

$$\sum_{w \in \mathcal{C}} (-1)^{v \cdot w} = \begin{cases} \mid \mathcal{C} \mid & \text{if } v \in \mathcal{C}^\perp \\ 0 & \text{if } v \notin \mathcal{C}^\perp \end{cases}$$

**Problem 9.** Let $\mathcal{C}$ be an $[n, k]$ linear code with generator matrix G and $\mathcal{C}^\perp$ its dual. Show that $\mathcal{C}$ is weakly self-dual, $\mathcal{C}^\perp \subset \mathcal{C}$ if and only if:

$$\mathrm{G}^T \mathrm{G} = 0.$$

**Problem 10.** Let G be the generator matrix of a binary $[32, 16]$ code:

$$\mathrm{G} = \begin{pmatrix} A & I & I & I \\ I & A & I & I \\ I & I & A & I \\ I & I & I & A \end{pmatrix}$$

with $I$ and $A$ given by:

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

10.1 Construct H, the parity check matrix of the code.

10.2 Show that the code is self dual.

10.3 Construct the cosets of the code.

10.4 What is the minimum distance of the code and how many errors can it correct?

**Problem 11.** Consider a finite field with $q$ elements, $F(q)$ and an element $\alpha \in F$. Prove the following properties:

1. $\alpha^j = 1 \iff \text{ord}(\alpha) = j$.
2. $\text{ord}(\alpha) = s \implies \alpha_1 \neq \alpha_2 \neq \ldots \neq \alpha^s$.
3. $\text{ord}(\alpha) = s,\ \text{ord}(\beta) = r,\ \gcd(s, r) = 1 \implies \text{ord}(\alpha \cdot \beta) = s \cdot r$.
4. $\text{ord}(\alpha^j) = \frac{\text{ord}(\alpha)}{\gcd(\text{ord}(\alpha), j)}$.
5. $F(q)$ has an element of order $(q-1)$.
6. The order of any element in $F(q)$ divides $(q-1)$.
7. $\gamma^q - \gamma = 0, \quad \forall \gamma \in F(q)$.

**Problem 12.** Show that among the set of generator matrices of an $[n, k]$ linear code $\mathcal{C}$ over the field $F$ there is one of the form $\text{G} = [I_k A]$.

**Problem 13.** Given an $[n, k]$ linear code $\mathcal{C}$ over the field $F$ the orthogonal complement, or the dual of $\mathcal{C}$, denoted as $\mathcal{C}^\perp$ consists of vectors (or $n$ tuples) orthogonal to every vector in $\mathcal{C}$:

$$\mathcal{C}^\perp \ = \ \{v \in V_n : \quad v \cdot c = 0, \quad \forall\, c \in \mathcal{C}\}$$

Prove that $\mathcal{C}^\perp$ is a subspace of $V_n$.

**Problem 14.** Show that if you pick up a different set of base vectors you obtain a different set of codewords for the linear code discussed in the example from Section 4.5. Determine a base such that the first three symbols in a codeword are precisely the information symbols in each of the eight messages.

**Problem 15.** An $[n, k, d]$ code over $GF(q)$ $\mathcal{C}$ is an MDS (Minimum Distance Separable) code if reaches the Singleton bound thus, $d = n - k + 1$. Show that the number of codewords of weight $(n - k + 1)$ of $\mathcal{C}$ is equal to $(q-1)\binom{n}{d}$. Show that its dual, $\mathcal{C}^\perp$ has distance $k + 1$.

**Problem 16.** Prove that the congruence modulo a subgroup is an equivalence relationship that partitions an Abelian group into disjoint equivalence classes.

**Problem 17.** Show that the $r+1$ rows of the generator matrix G of a first order Reed-Muller code are linearly independent.

**Problem 18.** If $F[x]$ denotes the set of all polynomials in $x$ with coefficients from the field $F$, $f(x)$ is a non-zero polynomial, $f(x) \in F[x]$, and $F[x]/(f(x))$ denotes the set of equivalence classes of polynomials modulo $f(x)$ then the following propositions are true:

(i) $F[x]$ is a principal ideal ring.

(ii) $F[x]/(f(x))$ is a principal ideal ring.

**Problem 19.** Let $S$ be a subspace of an $n$-dimensional vector space $V_n$ over the field F, $S \subset V_n$. Let $R$ be the ring of polynomials associated with $V_n$ and let $I$ be the set of polynomials in $R$ corresponding to $S$. Show that $S$ is a cyclic subspace of $V_n$ if and only if $I$ is an ideal in $R$.

**Problem 20.** Let $f(x) = x^n - 1$ and let $R$ be the ring of equivalence classes of polynomials modulo $f(x)$ with coefficients in the field $F$, $R = F[x]/(f(x))$. If $g(x)$ is a monic polynomial which divides $f(x)$ then $g(x)$ is the generator of the ideal of polynomials $I = m(x) \cdot g(x), \quad m(x) \in R$.

**Problem 21.** Show that there is a one-to-one correspondence between the cyclic subspaces of $V_n$ and the monic polynomials $g(x)$ which divide $f(x) = x^n - 1$.

**Problem 22.** Show that the two codes defined in Section 4.11 $\mathcal{C}_{\bar{h}}$ and $\mathcal{C}_h$ are equivalent codes.

**Problem 23.** In Section 4.11 we give an alternative method to construct the generator matrix of an $[n, k]$ cyclic code with generator polynomial $g(x)$. We divide $x^{n-k+i}$, $0 \leq i \leq k-1$ by the generator polynomial and we obtain a quotient $q^{(i)}(x)$ and a remainder $r^{(i)}(x)$:

$$x^{n-k+i} = g(x) \cdot q^{(i)}(x) + r^{(i)}(x), \quad 0 \leq i \leq k-1.$$

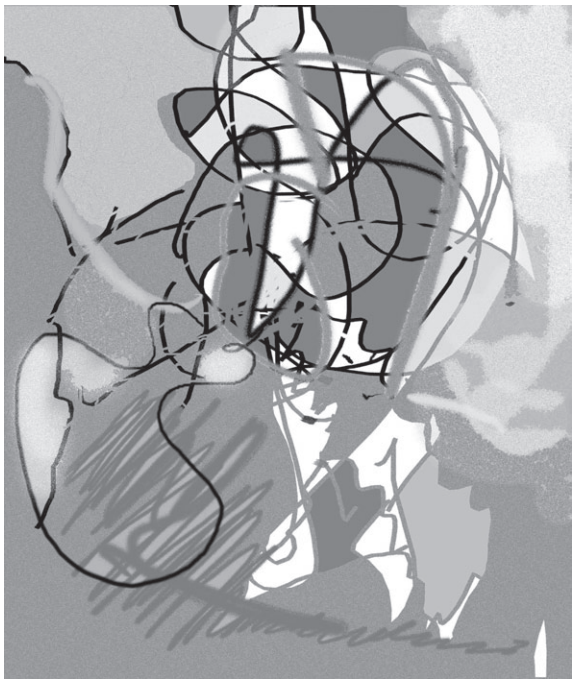
Consider an $(n-k) \times k$ matrix $R$ with rows corresponding to $-r^{(i)}(x)$ and construct the generator matrix as $\mathrm{G} = [RI_k]$. Show that when we encode a message $m = (m_0 m_1 \ldots m_{k-1})$ as a codeword $w = m\mathrm{G}$, then the last $k$ bits of the codeword contain the original message.

**Problem 24.** Show that binary Hamming codes are equivalent to cyclic codes.

**Problem 25.** Consider the following finite fields: (1) $GF(2^8)$; (2) $GF(2^9)$; (3) $GF(3^3)$; (4) $GF(5^2)$; (5) $GF(7^3)$; (6) $GF(39)$; (8) $GF(101)$. Determine the values $n$, $k$, and $d$ such that an $[n, k, d]$ Reed-Solomon code exists over each one of these fields.

**Problem 26.** Show that the dual of an $[n, k]$ Reed-Solomon code over $GF(q)$ is an $[n, (n-k)]$ Reed-Solomon code over $GF(q)$.

"To know that we know what we know, and to know that we do not know what we do not know, that is true knowledge." Copernicus.



# 5 Quantum Error Correcting Codes

Quantum systems are affected by *decoherence,* the alteration of the quantum state as a result of the interaction with the environment; decoherence occurs as the state of the quantum system embodying the information is affected by physical processes which cannot be reversed. Decoherence could be seen as *the effect of measurements of the state of the quantum system carried out by the environment* [471].

Correction of the errors caused by decoherence is one of the most challenging problems faced by quantum computing and quantum communication [342]. The first question that comes to mind is if quantum error correction is possible at all as quantum ensembles are analog systems and error correction can only be applied to digital systems. The answer is that we can *digitize* the errors and project small errors to a state where no error has occurred, or to a state with a large error [342] thus, quantum error correction is possible.

Quantum error correction allows us to reverse the effects of decoherence and reliably store, process, and transmit quantum information; thus, the topics covered in this chapter have multiple applications including:

1. Transmission of quantum information over a quantum channel expected to preserve some degree of coherence.

2. The effect of passage of time over a set of qubits which interact with the environment.

3. The result of the transformation performed by noisy quantum gates on a set of qubits.

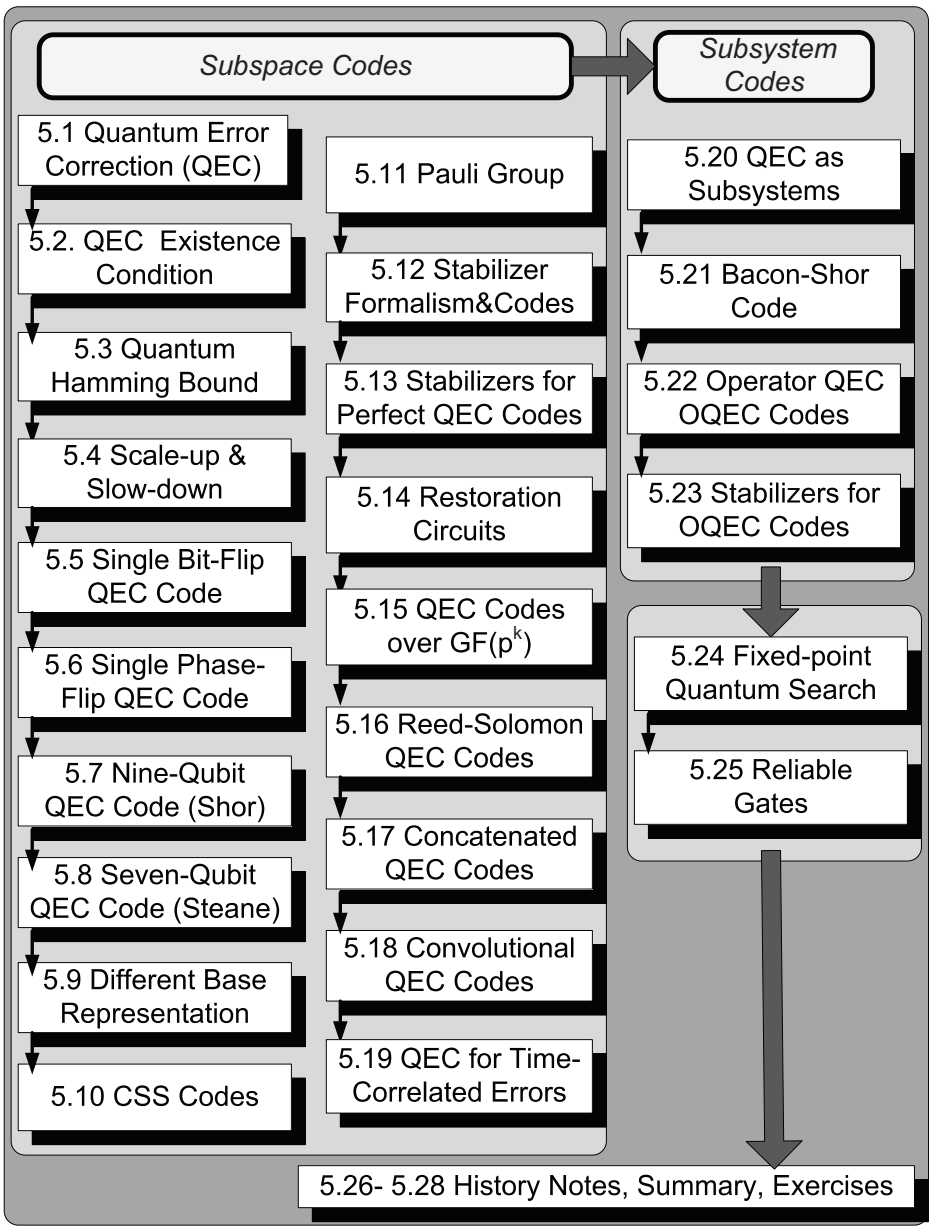



Figure 115: Chapter 5 at a glance.

According to Steane [403]: "a quantum error-correcting code is a method of storing or transmitting $k$ bits of information in such a way that if an arbitrary subset of the $n$ qubits

undergoes arbitrary errors, the transmitted quantum information can nevertheless be recovered exactly... In this context an *error* is any physical process that cannot be undone simply by applying, upon reception, a time-reversed version of the *error* process."

Efficient quantum error-correcting codes were discovered in mid 1990's [81, 369, 381, 383]; these were subspace codes closely related to classical error correcting codes. Then, a decade later, a new class of codes, the subsystem codes were introduced [17, 18, 260].

Figure 115 presents the organization of this chapter; Sections 5.1 to 5.19 cover the subspace codes. These codes are subspaces of the state space of the qubits with the property that an initial state of the subspace can be recovered when few qubits are affected by errors [40, 246, 247, 248, 381, 403]. The subspace codes are based on the so called standard model of error correction; this model is described by the triplet $(\mathcal{R}, \mathcal{E}, \mathcal{C})$ where the quantum code $\mathcal{C}$ is a subspace of the Hilbert space $\mathcal{H}$ associated with a given quantum system. The quantum operators $\mathcal{E}, \mathcal{R}$ are the error and the recovery operator; the recovery $\mathcal{R}$ undoes the effect of the error $\mathcal{E}$. In Section 5.20 and subsequent sections we discuss the subsystem codes when the error correcting codes can be viewed as subsystems where there is a clear separation between the subsystem where the protected information resides and there is no need for error correction other than the procedure to make the information available in a standard form [250]. Now we partition the Hilbert space as $\mathcal{H} = (\mathcal{C} \otimes \mathcal{D}) \oplus \mathcal{K}$ into two subspaces, $\mathcal{K}$ and a subspace orthogonal to $\mathcal{K}$ and on this orthogonal subspace we introduce a tensor product structure $\mathcal{C} \otimes \mathcal{D}$ and encode information in the "noiseless" subsystem $\mathcal{C}$.

We conclude the chapter with a brief discussion of fixed-point quantum search and its applications to quantum error correction; we also present an application of quantum error correction for the design of reliable quantum gates. The description of quantum error correction suggests a close analogy with classical error correction and, indeed, several subspace quantum codes are based on classical error correcting codes; thus, a review of the material discussed in Chapter 4 is helpful. In the next section we discuss basic principles of quantum error correction.

## 5.1 Quantum Error Correction

The aim of quantum error correction is to identify an encoding of information such that the set of correctable errors includes the *most likely errors* [405]. Many concepts in quantum error correction have a classical correspondent; as expected, concepts and techniques from classical error correcting codes cannot be applied directly to quantum error correction due to the special properties of quantum information. Quantum error corrections adds new challenges to the ones posed by classical error correction:

- To detect and to correct errors affecting classical information we increase the redundancy. Quantum states cannot be cloned; for error correction we have to take advantage of special properties of quantum information such as superposition and entanglement.

- The states of a system used to embody classical information are perfectly distinguishable. We can distinguish with absolute certainty *only orthogonal quantum states;* if a quantum state is transformed by two distinct quantum errors to two non-orthogonal states, then neither error can be corrected because the two states are not distinguishable.

- A measurement of a classical system does not affect the state of the system; we can determine if an error has occurred during transmission of classical information and then identify the bit(s) in error without altering the state of the classical system. A quantum measurement projects the state thus, in the general case, it alters the state of a quantum system, as we have seen in Chapter 2; we have to devise ingenious strategies to carry out the measurements necessary to identify a qubit in error.

- A classical bit may only be affected by a bit-flip error when "0" is flipped to "1" or "1" is flipped to "0." A qubit may be affected by bit-flip, phase-flip, or bit- and phase-flip errors; a quantum error can be represented as a linear combination of the bit-flip, phase-flip, and bit- and phase-flip errors

- Classical error correction codes assume uncorrelated errors and there is no classical form of entanglement. An error may propagate to some or all of the qubits entangled with one another and we can expect non-Markovian quantum errors, errors correlated in time and/or space; entanglement adds a new dimension to quantum error correction.

**Quantum errors.** The most general type of error in the state of a qubit can be described as a linear combination of the two basic types of quantum errors:

1. *Amplitude error*, or *bit-flip*. Changes of the form $\mid 0\rangle \quad \leftrightarrow \quad \mid 1\rangle$. Pauli operator $X$, applied to a qubit leads to a bit-flip. The classical errors correspond to quantum bit-flip errors.

2. *Phase error*, or *phase-flip*. Changes of the form $\mid 0\rangle + \mid 1\rangle \quad \leftrightarrow \quad \mid 0\rangle - \mid 1\rangle$. Pauli operator $Z$ applied to a qubit state leads to a phase-flip. There is no correspondent of quantum phase-flip in classical error correction.

When we say that *a qubit is in error* we assume that the qubit has suffered a transformation described by one of the Pauli operators: a bit-flip error is the result of applying the transformation $X$ implemented by an `X`-gate, a phase-flip error is the result of applying the transformation $Z$ implemented by an `Z`-gate, and a bit- and phase-flip error is the result of applying the transformation $Y$ implemented by a `Y`-gate.

A qubit originally in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ affected by bit-flip, phase-flip, or both will end up in a different state:

$$\text{bit-flip} \Longrightarrow \quad X \mid \psi\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_0 \end{pmatrix} \qquad \mapsto \quad \alpha_1 \mid 0\rangle + \alpha_0 \mid 1\rangle$$

$$\text{phase-flip} \Longrightarrow \quad Z \mid \psi\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ -\alpha_1 \end{pmatrix} \qquad \mapsto \quad \alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle$$

$$\text{bit \& phase-flip} \Longrightarrow \quad Y \mid \psi\rangle = i \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = i \begin{pmatrix} -\alpha_1 \\ \alpha_0 \end{pmatrix} \quad \mapsto \quad -(\alpha_1 \mid 0\rangle - \alpha_0 \mid 1\rangle)$$

When an error affects one qubit it is equally likely to be a $X, Y$, or $Z$ error. We assume that errors occurring in the block of qubits used to encode a single qubit are uncorrelated. If $\epsilon$ is the probability of one qubit in the bock to be in error, then the probability of two qubits to be in error is $\epsilon^2$. If $\epsilon$ is small enough, e.g. $\epsilon \leq 10^{-4}$, then $\epsilon^2$ is very small indeed and we can focus on the treatment of single errors. Even though two or more errors have a very low probability of occurrence, when such events occur, the code is not capable of correcting the errors.

**Encoding quantum information.** Figure 116 illustrates the basic quantum error correction process, whereby a single logical qubit described by the density matrix $\rho_i^{(1)}$, is encoded to an $n$-qubit register with the density matrix $\rho^{(n)}(0)$. The $n$ qubits are affected by errors and their density matrix at time $t$ is $\rho^{(n)}(t)$. The $n$-qubit register is then decoded to yield a single logical qubit, with the density matrix $\rho_f^{(1)}$.

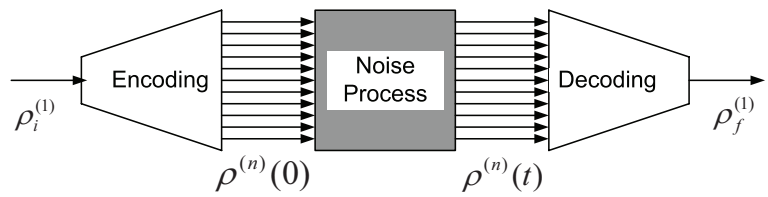



Figure 116: The quantum error correction process: a single logical qubit is encoded to an $n$-qubit register; the $n$ physical qubits are affected by errors; finally, the error correction process restores the state of the logical qubit.

In classical information processing systems error correction is ensured by redundant encoding of information. For example, in the case of linear codes we map a $k$-dimensional subspace of a vector space to an $n$-dimensional space, $n > k$, such that the $2^k$ codewords selected among the $2^n$ $n$-tuples satisfy some distance properties.

A more intuitive view of classical error correction is that we partition the $2^n$ elements in the space of binary $n$-tuples to $2^k$ disjoint Hamming spheres of radius $t$ and then we select the centers of these Hamming spheres as the codewords. Thus, we are able to encode $2^k$ messages. If at most $t$ errors occur when we send a codeword $w$, then the resulting $n$-tuple will still be within, or on the surface of the Hamming sphere $s_{(w,t)}$ of radius $t$ about $w$. Thus, we can correctly decode the $n$-tuple in error as $w$, the center of the sphere. *All the words inside a Hamming sphere of radius $t$ come from at most $t$ errors acting on the same codeword.*

Quantum error correction is based on similar ideas. For example, we can encode the state of a qubit

$$| \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$$

as a linear combination of $\mid 000\rangle$ and $\mid 111\rangle$:

$$| \varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle.$$

A random error can cause departures from the subspace spanned by $\mid 000\rangle$ and $\mid 111\rangle$. We shall be able to correct small errors because the component which was $\mid 000\rangle$ is likely to remain in a sub-space spanned by the four vectors, $\mid 000\rangle, \mid 001\rangle, \mid 010\rangle$, and $\mid 100\rangle$, while the

component which was $|\,111\rangle$ is likely to remain in a sub-space spanned by the four vectors, $|\,111\rangle, |\,110\rangle, |\,101\rangle$, and $|\,011\rangle$ [402].

In the general case a quantum code encodes $k$ *logical qubits* to $n$ *physical qubits* and has $2^k$ basis codewords. Any linear combination of the basis codewords is also a valid codeword. The *coding space*, the space of valid codewords of a code, is a subspace of a the corresponding Hilbert space $\mathcal{H}_{2^n}$.

In the process of correcting quantum errors it is possible to make a partial measurement which extracts only error information and leaves the encoded state untouched. *We can avoid the collapse of quantum information by choosing an encoding such that the standard errors translate the coding subspace to orthogonal subspaces.*

**The role of entanglement in quantum error correction.** The decoherence is due to the entanglement of the quantum system carrying information with the environment, which attempts to alter the state of the system. Initially, it was suspected that quantum error correction is not possible because the no-cloning theorem does not allow redundant quantum systems.

A major discovery is due to Peter Shor [381] who in 1995 showed that it is possible to restore the state of a quantum system using only partial information about the state. Until then it was not known how to perform an error correction step without a measurement which would alter the state. Shor summarizes the fundamental idea of quantum error correction as to "fight entanglement with entanglement." A quantum error correcting scheme takes advantage of entanglement in two ways:

(i) We entangle one qubit carrying information with $(n-1)$ qubits and create an $n$-qubit quantum codeword which is more resilient to errors, and

(ii) We entangle the $n$ qubits of this quantum codeword with ancilla qubits and construct the error syndrome without measuring the qubits of the codeword and thus, without altering their state.

**Non-demolition measurements.** The syndrome $\mathrm{H}e^T$ of a linear code with the parity check matrix H uniquely identifies the bit(s) in error and we have polynomial algorithms to determine the bit(s) in error once the syndrome is known. We apply the same basic strategy to identify a qubit in error, we calculate the syndrome and then measure it.

The computation of the syndrome for a quantum code requires a number of ancilla qubits; each ancilla qubit is entangled with multiple qubits of the encoded quantum word. The degree of entanglement of an ancilla qubit with each one of the $n$ physical qubits is relatively weak; a low degree of entanglement of the ancilla qubits with the physical qubits of the codeword allows us to carry out a *non-demolition measurement* of the ancilla qubits and determine the syndrome without altering the state of the $n$ physical qubits. We can also use a classical computer to compute the error from the syndrome once we have measured it.

For subspace codes the number of ancilla qubits is at least equal to the number of rows of the parity check matrix of the linear code used for quantum error correction. For now we assume that each ancilla qubit is the target of several `CNOT`s; the control qubits of the `CNOT` gates are a subset of the $n$ physical qubits used to encode the logical qubit, see for example Shor and Steane codes in Sections 5.7 and 5.8, respectively.

The scheme outlined above has several weaknesses e.g., an ancilla qubit may itself be affected by errors. As we know (see Section 1.19) phase-shifts allow the roles of the control and target qubits of a `CNOT` gate to be reversed; a qubit acting as the target of multiple `CNOT` gates can thus affect by back-propagation the state of all the control qubits. To avoid back-propagation of ancilla qubit errors we increase the number of ancilla qubits; each ancilla qubit is coupled with only one of the $n$ physical qubits, as we shall see in Section 5.8. The QEC process should remove more noise than it generates when itself is imperfect.

**Formal description of quantum error correction (QEC).** Given a system $A$ in state $\mid \varphi_A\rangle$ in the presence of noise the quantum error correction process can be succinctly described as follows:

- We introduce an additional system $C$ similar to $A$ in an initial state $\mid \varphi_C\rangle = \mid 0\rangle$.
- The "encoding" $E$ maps the joint state of $A$ and $C$ to a new state,
$$\mid \varphi_A\rangle \otimes \mid \varphi_C\rangle \quad \mapsto \quad \mid \varphi_{AC}^E\rangle = E\left(\mid \varphi_A\rangle \otimes \mid \varphi_C\rangle\right).$$
- The encoded state $\mid \varphi_{AC}^E\rangle$ is transmitted or stored and during this process it is subject to noise
$$\mid \varphi_{AC}^E\rangle \quad \mapsto \quad e_s \mid \varphi_{AC}^E\rangle.$$
- The "decoding" operation, $E^\dagger$ (the dual of $E$) is able to recover $\mid \varphi_A\rangle$, the correct state of $A$, if and only if
$$E^\dagger\left(\mid \varphi_{AC}^E\rangle\right) = \mid \varphi_A\rangle \otimes \mid \varphi_C\rangle.$$
The noise process is in general of the form
$$\mid \varphi_{AC}^E\rangle \quad \mapsto \quad \sum_s e_s \mid \varphi_{AC}^E\rangle.$$

In this description of QEC we assume that the errors affect only qubits either in storage or transmitted through a communication channel, and do not occur during the encoding $E$ or decoding $E^\dagger$. In other words, we assume that the quantum circuits involved in $E$ and $E^\dagger$ are not *faulty.* To deal with faulty quantum circuits we introduce an ancilla system $A$ initially in state $\mid \varphi_A\rangle = \mid 0_A\rangle$ and a *recovery operator*, $\mathcal{R}$, which corrects the encoded system $AC$ without decoding it and *carries the noise to the ancilla* $A$. The recovery operator performs the transformation
$$\mathcal{R}\, e_s\left(\mid \varphi_{AC}^E\rangle \otimes \mid 0_A\rangle\right) \quad \mapsto \quad \mid \varphi_{AC}^E\rangle \otimes \mid s_A\rangle, \quad \forall e_s \in S.$$

**The steps for quantum error correction.** The subspace codes generally use the following procedure to correct errors:

1. Construct two basis vectors in $\mathcal{H}_{2^n}$
$$\mid 0\rangle_L \equiv \mid 000\ldots0\rangle \quad \text{and} \quad \mid 1\rangle_L \equiv \mid 111\ldots1\rangle.$$

2. Encode one logical qubit as $n$ physical qubits

$$\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle \quad \Longrightarrow \quad \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$$

3. Process - transmit, store, or transform the resulting codeword.

4. Periodically compute the syndrome without altering the codeword.

5. Use the information provided by the error syndrome to locate the single error of any one of the $n$ qubits of the codeword.

6. Correct the error.

Figure 117 in Section 5.5 illustrates these steps for the correction of a single bit-flip error. Since multiple types of errors may affect the state of a qubit we have to perform error correction in multiple bases. This is the reason why a quantum error correcting code is based on multiple classical error correcting codes. This requirement is spelled out by the Quantum Error Correction Theorem discussed in Section 5.10.

A question we may ask is if it is easier to correct pure or mixed states of a quantum system. Mixed states are ensembles of pure states, therefore if we can correct all pure states of an ensemble then we are able to correct the mixed state. If we wish to think in terms of a noisy quantum communication channel we can say that the channel applies a superoperator to the input density matrix [174]. We can write this superoperator as a direct sum of a number of distinct operators, each acting on different input pure states. A code capable of correcting any of these operators can correct the mixed state. The decomposition of the superoperator is done by diagonalization of the matrix representing the superoperator to possibly non-unitary matrices.

The next question we address if it is always possible to construct a quantum code which allows us to distinguish the effects of two distinct quantum errors. A quantum code is a subspace of a Hilbert space so we have to analyze the effects of the two error operators on the basis sates of the code.

## 5.2 A Necessary Condition for the Existence of a Quantum Code

We now discuss necessary conditions for the existence of a quantum code. First, the quantum code should distinguish the effects of two distinct errors applied to the same state of a quantum system. A quantum state can be represented as a superposition of the basis states thus, a necessary condition for distinguishability is that orthogonal basis states be transformed to orthogonal states by the two distinct error operators. Second, we impose the condition that the measurement required by the error correction procedure should not modify the state of the system, therefore should not reveal any information about the state of the system.

A quantum error is an undesirable change of state due to the interaction of the system with the environment. Assume a quantum error is represented by an error operator $\mathbf{E}$. When applied to a system in state $\mid \psi\rangle$ the system ends up in state $\mid \varphi\rangle$:

$$| \varphi\rangle = \mathbf{E} \mid \psi\rangle.$$

If we can correct errors $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$ then we can correct any linear combination of them $\mathbf{E} = \alpha\mathbf{E}_\alpha + \beta\mathbf{E}_\beta$.

Recall that we can reliably distinguish two quantum states $\mid \psi_i\rangle, \mid \psi_j\rangle \in \mathcal{H}_{2^n}$ if and only if the two states are orthogonal, $\langle\psi_i \mid \psi_j\rangle = 0, \quad i \neq j$.

The coding space of a $[n, k]$ quantum code has $2^k$ basis vectors and it is a subspace of the Hilbert space $\mathcal{H}_{2^n}$. Consider two distinct basis sates $\mid \psi_i\rangle$ and $\mid \psi_j\rangle$, $i \neq j$. To correct the errors $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$, we should be able to distinguish the states resulting from applying the two error operators to these basis states, $\mathbf{E}_\alpha \mid \psi_i\rangle$ and $\mathbf{E}_\beta \mid \psi_j\rangle$. These states must be orthogonal thus, the formal description of the distinguishability condition is

$$\langle\psi_i \mid \mathbf{E}_\alpha^\dagger \mathbf{E}_\beta \mid \psi_j\rangle = 0.$$

A major problem we are faced with in quantum error correction is that if we gather information about the quantum states $\mid \psi_i\rangle$ and/or $\mid \psi_j\rangle$ we disturb the state, see Knill and Laflamme [245]. We gather information about the state $\mid \psi_i\rangle$ by measuring $\langle\psi_i \mid \mathbf{E}_\alpha^\dagger \mathbf{E}_\beta \mid \psi_i\rangle$ for all possible errors $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$. We can claim that we have not obtained any new information about the state if indeed $\langle\psi_i \mid \mathbf{E}_\alpha^\dagger \mathbf{E}_\beta \mid \psi_i\rangle$ is the same for all basis states

$$\langle\psi_i \mid \mathbf{E}_\alpha^\dagger \mathbf{E}_\beta \mid \psi_i\rangle = constant, \quad \forall\ i \in \{1, 2^k\}.$$

This condition must also be satisfied for all error operators $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$. We combine the last two equations as

$$\langle\psi_i \mid \mathbf{E}_\alpha^\dagger \mathbf{E}_\beta \mid \psi_j\rangle = C_{\alpha\beta}\delta_{ij}$$

with $C_{\alpha\beta}$ a constant which does not depend on $i$ and $j$; $\delta_{ij}$ is the Kronecker delta function. This equation gives a necessary condition for the existence of a quantum code capable of correcting errors described by error operators $\mathbf{E}_\alpha$ and $\mathbf{E}_\beta$.

## 5.3 Quantum Hamming Bound

Our ability to construct "good" classical codes is limited; given the probability of error per bit, a "good" or efficient classical error correcting code, is one with a *rate* $R$, the ratio $R = k/n$, that does not become increasingly smaller as $k$ increases. The Hamming bound gives a lower bound for the number of parity check bits, $r = n - k$, of an $[n, k]$ classical code

$$2^r - r \ \geq\ k + 1 \qquad \rightarrow \qquad 2^r \ \geq\ n + 1.$$

While classical codes increase the information redundancy by adding parity check bits, quantum codes encode a logical qubit to multiple physical qubits; the quantum Hamming bound establishes a *lower bound on the number of physical qubits required for every logical qubit.*

Recall that a perfect classical code able to correct $t$ errors is one where Hamming spheres of radius $t$ about each codeword are *disjoint* and exhaust the entire space of $n$-tuples. Similarly, the Hilbert space of dimension $2^n$ should be large enough to accommodate *orthogonal*

*subspaces of the Hilbert space* for possible bit-flip, phase-flip, and bit- and phase-flip errors affecting each one of the $n$ physical qubits.

To introduce the concept of a perfect quantum error correcting code we discuss an inequality established by Laflamme, Miquel, Paz, and Zurek in [264]. We consider the case $k = 1$ and a quantum error correcting code which encodes one logical qubit to $n$ physical qubits in the Hilbert space $\mathcal{H}_{2^n}$ with basis vectors $\{| 0\rangle, | 1\rangle, \ldots | i\rangle, \ldots | 2^n - 1\rangle\}$:

$$| 0\rangle_L = \sum_i \alpha_i \mid i\rangle \qquad \text{and} \qquad | 1\rangle_L = \sum_i \beta_i \mid i\rangle$$

The two logical qubits are entangled states in $\mathcal{H}_{2^n}$ and the quantum error correcting code must map coherently the two-dimensional Hilbert space spanned by $| 0\rangle_L$ and $| 1\rangle_L$ to two-dimensional Hilbert spaces to ensure that the code is capable of correcting the three types of errors (bit-flip, phase-flip, as well as bit- and phase-flip) for each of the $n$ qubits.

The Hilbert space $\mathcal{H}_{2^n}$ must be large enough to accommodate $n$ two-dimensional subspaces corresponding to bit flip errors of each of the $n$ qubits, and in addition, $n$ subspaces for the phase-flip errors, $n$ subspaces for bit-flip and phase-flip errors, and, finally, one subspace for the entangled state $| 0\rangle_L$ without errors. A similar requirement for $| 1\rangle_L$ leads us to conclude that $n$ must satisfy the quantum Hamming bound

$$2(3n + 1) \leq 2^n.$$

We see immediately that $n = 5$ is the smallest number of physical qubits required to encode the two superposition states $| 0\rangle_L$ and $| 1\rangle_L$, and then be able to recover them regardless of the qubit in error and the type of error.

## 5.4 Scale-up and Slow-down

While the lower bound for the number of physical qubits for error correction is relatively modest, in reality the quantum circuitry for error correction requires a dramatic increase of the number of physical qubits for a fault-tolerant quantum circuit, as well as, the number of steps required by a fault-tolerant computation.

Fault-tolerance is a complex subject beyond the scope of this book; nevertheless, we have to stress at this early stage of our discussion that quantum error correcting codes in themselves are not sufficient to ensure fault-tolerance. Fault-tolerant quantum information processing requires fault-tolerant quantum gates and circuits; to illustrate the difficulty of problems posed by fault-tolerance we dissect the construction of a reliable `CNOT` gate based on the seven qubit code of Steane in Section 5.25.

We now introduce two quantitative measures for these two undesirable but unavoidable consequences of quantum error correction and of special measures required by fault-tolerance. Let $k$ be the number of logical qubits and $n > k$ the number of physical qubits used for the implementation of a quantum circuit. Let $Q$ the number of elementary steps (operations on logical qubits), and $T > Q$ the number of elementary gates required for a specific computation. Then the *overhead of the error correction* is measured by the scale-up, the ratio of the number of physical to logical qubits, $\sigma = n/k$ and the slow-down, the ratio of the number of gates to

the number of the elementary steps of the computation, $\tau = T/Q$. The scale-up is the inverse of the classical code rate $\sigma = 1/R$.

The maximum computation size is given by the product $\mu = kQ$ and it is a function of several factors that quantify:

- The noise and the imprecision that can be tolerated.
- The degree of parallelism of the quantum hardware and the increase in the number of qubits required [410]. Parallelism means that a number of gates can be acted on in parallel. Parallelism is a critical ingredient of quantum error corrections, as it allows us to keep storage errors under control.

At this point we should caution the reader that the quantum error correction process discussed in this chapter does not distinguish between different physical realizations of quantum devices. As we shall see in Chapter 6 the characteristics of the circuits based on ion traps are very different from those based on liquid NMR, quantum dots, or optical systems. Table 31 in Section 5.25 shows that the decoherence time ranges from $10^4$ seconds for nuclear spin based systems to $10^{-9}$ for quantum dot systems based on charge; it is considerably more difficult to build reliable quantum circuits for systems with a very short decoherence time.

While the results regarding quantum error correction are general, the complexity of the circuits involved in quantum error correction, reflected by the scale-up and the slow-down measures, is far beyond today's technological possibilities; for example, a fault-tolerant implementation of Shor's algorithm would most likely require thousands of physical qubits, at least two orders of magnitude more qubits than the systems reported in the literature have been able to harness. It may be possible though to resort to techniques which exploit the specific properties of individual physical realization of quantum devices to manage the complexity of the quantum circuits for fault-tolerant systems.

Once we are convinced that in spite of its overhead, error correction is feasible, we focus our attention on quantum error correcting codes. Recall that in Chapter 1 we started our discussion of quantum algorithms with several "toy" problems; we follow the same strategy; first, we examine a few simple codes to illustrate the basic concepts and then analyze on more complex quantum error correcting codes.

## 5.5 A Repetitive Quantum Code for a Single Bit-flip Error

In this section and in the next one we analyze the trivial case of quantum repetitive codes; to facilitate the understanding of the quantum error correction schemes we present the details of the transformations and the circuits used to encode, to decode, and to correct errors.

We consider only *bit-flip* errors affecting a qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$; a bit-flip occurs with probability $p$ if we observe the state of the qubit being $X \mid \psi\rangle$ with probability $p$. Figure 117 illustrate the encoding and correction of one bit-flip error. We prepare several ancilla qubits, entangle them with the original qubits, and compute the error syndrome and store it in ancilla qubits without altering the state of the $n$-qubit word; then, we correct the error using the result of a non-demolition measurement of the syndrome.

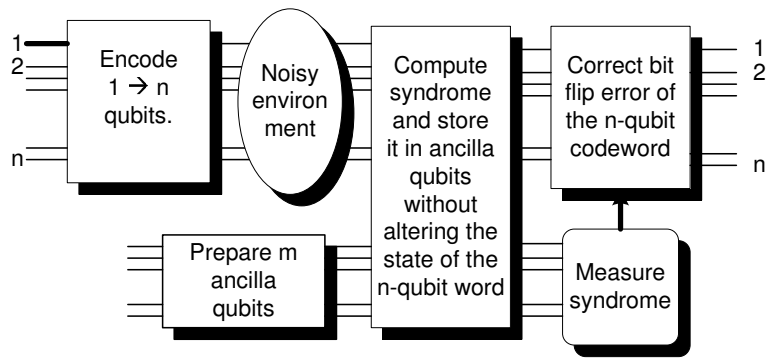



Figure 117: Encoding and correction of one bit-flip error.

**Error detection.** We encode the original qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as $\mid \varphi\rangle = \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$. Call $\mid \varphi_i^{bf}\rangle, \quad 0 \leq i \leq n$ the state of $\mid \psi\rangle$ after a bit-flip of qubit $i$ and $\mid \bar{\varphi}_i^{bf}\rangle, \quad 0 \leq i \leq n$ the complement of the error codeword - the binary string with bit $b'_j = \bar{b}_j, 1 \leq j \leq n$ with $b_j$ the j-th bit of $\mid \varphi_i^{bf}\rangle$. Let $0_{L,i}$ be the bit string with the $i$-th bit flipped to 1 and $1_{L,i}$ the bit string with the $i$-th bit flipped to 0; then

$$\mid \varphi_i^{bf}\rangle = \alpha_0 \mid 0_{L,i}\rangle + \alpha_1 \mid 1_{L,i}\rangle$$

The $(n+1)$ projectors used to identify the qubit in the error are

$$\mathbf{P}_i^{bf} = \mid \varphi_i^{bf}\rangle\langle\varphi_i^{bf} \mid + \mid \bar{\varphi}_i^{bf}\rangle\langle\bar{\varphi}_i^{bf} \mid \quad 0 \leq i \leq n.$$

The $(n+1)$ syndromes indicating that no error has occurred or that qubit $i$ suffered a bit-flip error are

$$\Sigma_i^{bf} = \langle\varphi_i^{bf} \mid \mathbf{P}_i^{bf} \mid \varphi_i^{bf}\rangle, \quad 0 \leq i \leq n.$$

If the $k$-th qubit is in error then

$$\Sigma_i^{bf} = \begin{cases} 1 & \text{if} \quad i = k \\ 0 & \text{if} \quad i \neq k. \end{cases}$$

To compute the syndrome we use $m$ ancilla qubits entangled with the $n$ qubits of the codeword; then a non-demolition measurement allows us to measure the error syndrome without altering the original qubits. Each ancilla qubit is entangled with multiple qubits of $\mid \varphi_i^{bf}\rangle$, and the low "degree of entanglement" allows us to carry out such a non-demolition measurement.

**Example.** *Correction of a single bit flip error of a repetitive code.* The minimum number of qubits that satisfy the quantum Hamming bound is $n = 5$; nevertheless, to illustrate the basic ideas of quantum error correction we consider a repetitive code with which maps one logical qubit to three physical qubits; the mapping of the basis vectors is $\mid 0\rangle \mapsto \mid 000\rangle$ and $\mid 1\rangle \mapsto \mid 111\rangle$. The circuit in Figure 118 encodes the qubit in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as $\mid \varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$. The input state $\mid \xi\rangle$ is

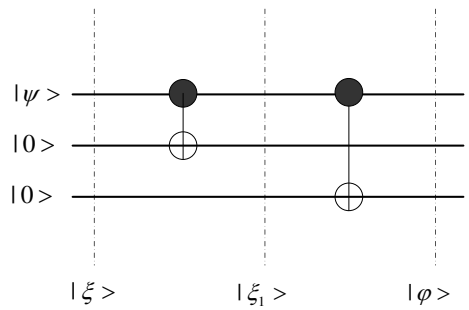



Figure 118: A circuit to encode the qubit $|\,\psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as $|\,\varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$. The input is the tensor product $|\,\xi\rangle = |\,\psi\rangle \otimes \mid 0\rangle \otimes \mid 0\rangle$.

$$|\,\xi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ \alpha_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The transfer matrix of the first stage of the circuit in Figure 118 is

$$G_1 = G_{CNOT} \otimes \mathbb{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

The state after the first stage is

$$|\,\xi_1\rangle = G_1 \mid \xi\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \\ 0 \end{pmatrix}.$$

The transfer matrix of the second stage was derived in [284] as

$$G_Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

The output of the encoding circuit is

$$\mid \varphi\rangle = G_Q \mid \xi_1\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \end{pmatrix}.$$

Thus, we have

$$\mid \varphi\rangle = \alpha_0 \mid 000\rangle \ + \ \alpha_1 \mid 111\rangle.$$

We turn our attention from encoding to error detection and error correction of the quantum codeword $\mid \varphi\rangle$ for a bit-flipping channel; there are four projection operators, $\mathbf{P}_i^{bf}$, corresponding to the four possible cases: no error; the first qubit, the second qubit, or the third qubit are bit-flipped

$$\begin{array}{llll} \text{No error} & (i=0) & \mapsto & \mathbf{P}_0^{bf} = \mid 000\rangle\langle 000 \mid + \mid 111\rangle\langle 111 \mid \\ \text{Qubit 1 in error} & (i=1) & \mapsto & \mathbf{P}_1^{bf} = \mid 100\rangle\langle 100 \mid + \mid 011\rangle\langle 011 \mid \\ \text{Qubit 2 in error} & (i=2) & \mapsto & \mathbf{P}_2^{bf} = \mid 010\rangle\langle 010 \mid + \mid 101\rangle\langle 101 \mid \\ \text{Qubit 3 in error} & (i=3) & \mapsto & \mathbf{P}_3^{bf} = \mid 001\rangle\langle 001 \mid + \mid 110\rangle\langle 110 \mid \end{array}$$

with:

$$\mathbf{P}_0^{bf} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{P}_1^{bf} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{P}_2^{bf} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{P}_3^{bf} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The error correction procedure is based on a measurement leading to the calculation of the *error syndrome*. Call $\mid \varphi_i^{bf}\rangle, \quad 0 \leq i \leq 3$ the error codeword with $i = 0$ corresponding to no error, $i = 1, 2, 3$ corresponding to bit-flip of the first, second or the third qubit of the quantum codeword $\mid \varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$, respectively. Then the error syndromes corresponding to the four possible cases are respectively:

| | |
|---|---|
| $\langle\varphi_i^{bf} \mid \mathbf{P}_0^{bf} \mid \varphi_i^{bf}\rangle,\ 0 \leq i \leq 3$ | $\langle\varphi_i^{bf} \mid \mathbf{P}_1^{bf} \mid \varphi_i^{bf}\rangle,\ 0 \leq i \leq 3$ |
| $\langle\varphi_i^{bf} \mid \mathbf{P}_2^{bf} \mid \varphi_i^{bf}\rangle,\ 0 \leq i \leq 3$ | $\langle\varphi_i^{bf} \mid \mathbf{P}_3^{bf} \mid \varphi_i^{bf}\rangle,\ 0 \leq i \leq 3$ |

When one bit-flip error occurs there is a one-to-one correspondence between the error syndrome and the qubit in error; consider first the case of an error in the first qubit

$$\mid \varphi_1^{bf}\rangle = \alpha_0 \mid 100\rangle + \alpha_1 \mid 011\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \\ \alpha_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{or} \qquad \langle\varphi_1^{bf} \mid \ \mapsto \ (0\ 0\ 0\ \alpha_1^*\ \alpha_0^*\ 0\ 0\ 0)$$

It is easy to see that $\langle\varphi_1^{bf} \mid \mathbf{P}_1^{bf} \mid \varphi_1^{bf}\rangle = 1$ while

$$\langle\varphi_1^{bf} \mid \mathbf{P}_0^{bf} \mid \varphi_1^{bf}\rangle = 0, \quad \langle\varphi_1^{bf} \mid \mathbf{P}_2^{bf} \mid \varphi_1^{bf}\rangle = 0, \quad \langle\varphi_1^{bf} \mid \mathbf{P}_3^{bf} \mid \varphi_1^{bf}\rangle = 0.$$

Indeed,

$$\langle\varphi_1^{bf} \mid \mathbf{P}_1^{bf} \ = \ (0\ 0\ 0\ \alpha_1^*\ \alpha_0^*\ 0\ 0\ 0) \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (0\ 0\ 0\ \alpha_1^*\ \alpha_0^*\ 0\ 0\ 0)$$

$$\langle \varphi_1^{bf} \mid \mathbf{P}_1^{bf} \mid \varphi_1^{bf} \rangle \;=\; (0\ 0\ 0\ \alpha_1^*\ \alpha_0^*\ 0\ 0\ 0) \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \\ \alpha_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \;=\; \mid \alpha_0 \mid^2 + \mid \alpha_1 \mid^2 = 1.$$

We only show that in this case $\langle \varphi_1^{bf} \mid \mathbf{P}_0^{bf} \mid \varphi_1^{bf} \rangle \;=\; 0$:

$$\langle \varphi_1^{bf} \mid \mathbf{P}_0^{bf} \;=\; (0\ 0\ 0\ \alpha_1^*\ \alpha_0^*\ 0\ 0\ 0) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \;=\; (0\ 0\ 0\ 0\ 0\ 0\ 0\ 0).$$

$$\langle \varphi_1^{bf} \mid \mathbf{P}_0^{bf} \mid \varphi_1^{bf} \rangle \;=\; (0\ 0\ 0\ 0\ 0\ 0\ 0\ 0) \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \\ \alpha_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \;=\; 0.$$

The case of a bit-flip of the second and third qubit are treated similarly.

Once we have calculated the syndrome the error correction is straightforward, we flip back the corresponding qubit of the quantum codeword if an error is detected. We apply a transformation given by a unitary operator $\mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_2$ or $\mathbf{C}_3$ to the codeword:

$$\begin{array}{rcl} \text{No Error} & \mapsto & \mathbf{C}_0 = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \\ \text{Bit-flip error of first qubit} & \mapsto & \mathbf{C}_1 = X \otimes \mathbb{I} \otimes \mathbb{I} \\ \text{Bit-flip error of second qubit} & \mapsto & \mathbf{C}_2 = \mathbb{I} \otimes X \otimes \mathbb{I} \\ \text{Bit-flip error of third qubit} & \mapsto & \mathbf{C}_3 = \mathbb{I} \otimes \mathbb{I} \otimes X \end{array}.$$

The syndrome is computed as $\mathrm{H}e_i^T$; the parity check matrix of the code discussed in this section is

$$\mathrm{H} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

The circuit to calculate the syndrome requires a number of ancilla larger or equal to the number of rows of the parity check matrix; in case of equality *each ancilla is entangled with the physical qubits selected by the corresponding row of the parity check matrix.*

A circuit for encoding and syndrome calculation for this code is presented in Figure 119; the three qubits of the encoded state $\mid \varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$ act as control qubits of the `CNOT` gates used to set the ancilla qubits. Each ancilla qubit is entangled with multiple qubits of the original codeword and the resulting degree of entanglement allows for non-demolition measurements. The first ancilla qubit is set to 1 if and only if only one of the second or third qubit is flipped; the second ancilla qubit is set to 1 if and only if only one of the first or third qubit is flipped. So the pair of ancilla qubits is set to: 00 if there is no error, 01 if the topmost qubit is bit-flipped, 10 is the middle qubit is bit-flipped, and to 11 if the third qubit is bit-flipped. If two qubits are bit-flipped the error cannot be corrected. The syndrome does not contain any information regarding the values of $\alpha_0$ or $\alpha_1$.

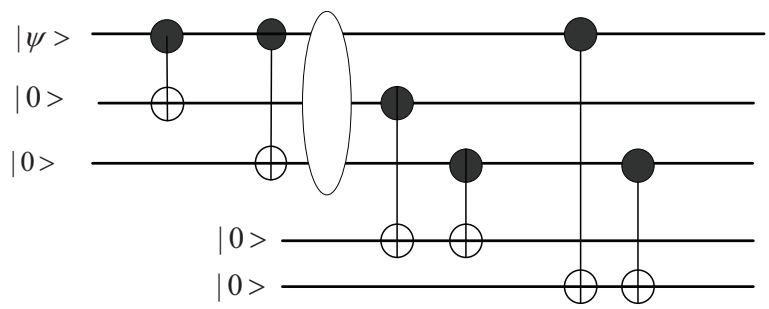



Figure 119: A circuit for encoding the qubit $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \mid \alpha_1 1\rangle$ as $\alpha_0 \mid 000\rangle + \mid \alpha_1 111\rangle$ and for syndrome calculation; the qubits are numbered top to bottom. To calculate the syndrome we need two ancilla qubits; the first ancilla is the target qubit of `CNOT` gates selected by the first row of the parity check matrix, 011, and the second is the target selected by the second row of the parity check matrix, 101.

An alternative analysis of the error correction process considers an observable $q_{12}$ which provides information whether the first or the second qubit have been flipped and returns the value $+1$ if none of them has been flipped and $-1$ if one has been flipped. The projective measurement uses two projectors $(\mid 00\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid) \otimes \mathbb{I}$ and $(\mid 01\rangle\langle 01 \mid + \mid 10\rangle\langle 10 \mid) \otimes \mathbb{I}$ and has the spectral decomposition

$$q_{12} = (\mid 00\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid) \otimes \mathbb{I} - (\mid 01\rangle\langle 01 \mid + \mid 10\rangle\langle 10 \mid) \otimes \mathbb{I}.$$

A second observable $q_{23}$ provides information whether the second or the third qubit have been flipped and returns the value $+1$ if none of them has been flipped and $-1$ if one of them has been flipped. The projectors are

$$\mathbb{I} \otimes (\mid 00\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid) \quad \text{and} \quad \mathbb{I} \otimes (\mid 01\rangle\langle 01 \mid + \mid 10\rangle\langle 10 \mid).$$

The projective measurement has the spectral decomposition

$$q_{23} = \mathbb{I} \otimes (\mid 00\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid) - \mathbb{I} \otimes (\mid 01\rangle\langle 01 \mid + \mid 10\rangle\langle 10 \mid).$$

By combining the two measurements we can determine if any of the three qubits of the quantum word has been affected by an error and if so, which one. Our next challenge is to correct phase-flip errors, the topic of the following section.

## 5.6 A Repetitive Quantum Code for a Single Phase-flip Error

A phase-flip error is described as the effect of applying the $Z$ operator to a qubit in state $\mid \psi\rangle$

$$Z \mid \psi\rangle = Z\left(\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle\right) = \alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ -\alpha_1 \end{pmatrix}.$$

The phase-flip error has no classical equivalent and the solution is to convert the phase-flip to a bit-flip error by a change of basis carried out by the Hadamard Transform $H$:

$$\{\mid 0\rangle, \mid 1\rangle\} \overset{H}{\mapsto} \{\mid +\rangle, \mid -\rangle\} \overset{H}{\mapsto} \{\mid 0\rangle, \mid 1\rangle\}$$

with $HH = \mathbb{I}$; the new basis state vectors $\mid +\rangle, \mid -\rangle$ are

$$\mid +\rangle = H \mid 0\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \quad \Leftrightarrow \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$\mid -\rangle = H \mid 1\rangle = \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}} \quad \Leftrightarrow \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

In the rotated basis $Z$ becomes a bit-flip operator

$$Z \mid +\rangle = \mid -\rangle \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and

$$Z \mid -\rangle = \mid +\rangle \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The error correction process follows closely the one presented in Section 5.5. Call $\mid \varphi_i^{phf}\rangle$ the state of $\mid \psi\rangle$ after a phase-flip of qubit $i$ with $0 \leq i \leq n$; the $(n+1)$ projectors used to identify the qubit in the error are

$$\mathbf{P}_i^{phf} = \mid \varphi_i^{phf}\rangle\langle\varphi_i^{phf} \mid + \mid \bar{\varphi}_i^{phf}\rangle\langle\bar{\varphi}_i^{phf} \mid, \qquad 0 \leq i \leq n.$$

The $(n+1)$ syndromes indicating that no error has occurred or that qubit $i$ suffered a phase-flip error are

$$\Sigma_i^{phf} = \langle\varphi_i^{bf} \mid \mathbf{P}_i^{bf} \mid \varphi_i^{bf}\rangle, \quad 0 \leq i \leq n.$$

If the $k$-th qubit is in error then

$$\Sigma_i^{phf} = \begin{cases} 1 & \text{if} \quad i = k \\ 0 & \text{if} \quad i \neq k. \end{cases}$$

**Example.** *Consider the three-qubit repetitive code discussed in Section 5.5 and apply a rotation of the basis vectors to correct phase-flip errors.*

The scheme for error correction follows closely the one presented in Figure 117; the states $\mid 0\rangle$ and $\mid 1\rangle$ are encoded as

$$\mid 0\rangle \mapsto \mid +++\rangle \quad \text{and} \quad \mid 1\rangle \mapsto \mid ---\rangle$$

The encoding, the error detection and the recovery are performed with respect to logical states $\mid +++\rangle$ and $\mid ---\rangle$; the circuit in Figure 120 encodes a qubit state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ to $\mid \zeta\rangle = \alpha_0 \mid +++\rangle + \alpha_1 \mid ---\rangle$.

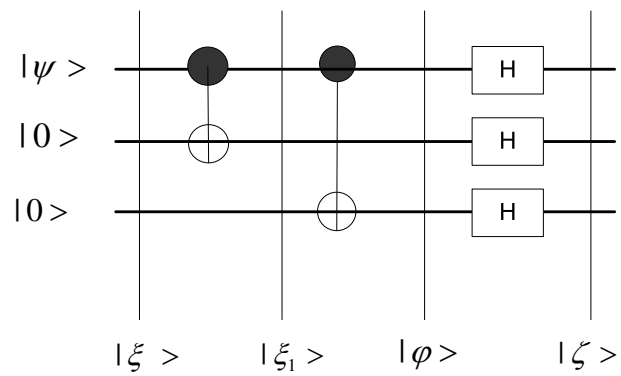



Figure 120: An encoding circuit to correct a phase-flip error. The state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ of the qubit is encoded first to $\mid \varphi\rangle = \alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$ as in the case of bit-flip error correction; then $\mid \varphi\rangle$ is encoded as $\mid \zeta\rangle = \alpha_0 \mid +++\rangle + \alpha_1 \mid ---\rangle$.

The first two stages of this circuit are similar to the circuit in Figure 118 used for encoding in basis $\{\mid 000\rangle, \mid 111\rangle\}$ and the third stage consists of Hadamard gates, one for each qubit. If the qubit to be encoded is in state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$, the input state for the third stage is $\mid \varphi\rangle$ and the transfer matrix of the third stage is

$$H^{\otimes 3} = H \otimes H \otimes H = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The state of the three-qubit system at the output of the encoding circuit is

$$\mid \zeta\rangle = H^{\otimes 3} \mid \varphi\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha_1 \end{pmatrix} = \frac{1}{2\sqrt{2}} \begin{pmatrix} \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \end{pmatrix}$$

The detailed calculations related to the first two stages of the circuit in Figure 120 are presented in Section 5.5; in Section 5.9) we show that

$$|+++\rangle = \frac{1}{2\sqrt{2}}(|\,000\rangle + |\,001\rangle + |\,010\rangle + |\,011\rangle + |\,100\rangle + |\,101\rangle + |\,110\rangle + |\,111\rangle)$$

and

$$|---\rangle = \frac{1}{2\sqrt{2}}(|\,000\rangle - |\,001\rangle - |\,010\rangle + |\,011\rangle - |\,100\rangle + |\,101\rangle + |\,110\rangle - |\,111\rangle).$$

Thus, our circuit performs the desired transformation:

$$|\,\psi\rangle = \alpha_0\,|\,0\rangle + \alpha_1\,|\,1\rangle \quad \mapsto \quad |\,\zeta\rangle = \alpha_0\,|+++\rangle + \alpha_1\,|---\rangle.$$

We now discuss the projector operators for phase-flip errors; there are four projectors, $\mathbf{P}_i^{phf}$ corresponding to the four possible cases: no error, the first qubit, the second qubit, or the third qubit are phase-flipped

$$\begin{array}{llll}
\text{No error} & (i=0) & \mapsto & \mathbf{P}_0^{phf} = |+++\rangle\langle+++| + |---\rangle\langle---| \\
\text{Qubit 1 in error} & (i=1) & \mapsto & \mathbf{P}_1^{phf} = |-++\rangle\langle-++| + |+--\rangle\langle+--| \\
\text{Qubit 2 in error} & (i=2) & \mapsto & \mathbf{P}_2^{phf} = |+-+\rangle\langle+-+| + |-+-\rangle\langle-+-| \\
\text{Qubit 3 in error} & (i=3) & \mapsto & \mathbf{P}_3^{phf} = |++-\rangle\langle++-| + |--+\rangle\langle--+|
\end{array}$$

The phase-flip projectors are related to the ones for bit-flip errors by the following relation

$$\mathbf{P}_i^{phf} = H^{\otimes 3}\mathbf{P}_i^{bf}H^{\otimes 3}, \quad 0 \le i \le 3$$

Thus

$$\mathbf{P}_0^{phf} = H^{\otimes 3}\mathbf{P}_0^{bf}H^{\otimes 3} = \frac{1}{4}\begin{pmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1
\end{pmatrix},$$

$$\mathbf{P}_1^{phf} = H^{\otimes 3}\mathbf{P}_1^{bf}H^{\otimes 3} = \frac{1}{4}\begin{pmatrix}
1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 \\
0 & 1 & 1 & 0 & -1 & 0 & 0 & -1 \\
0 & 1 & 1 & 0 & -1 & 0 & 0 & -1 \\
1 & 0 & 0 & 1 & 0 & -1 & -1 & 0 \\
0 & -1 & -1 & 0 & 1 & 0 & 0 & 1 \\
-1 & 0 & 0 & -1 & 0 & 1 & 1 & 0 \\
-1 & 0 & 0 & -1 & 0 & 1 & 1 & 0 \\
0 & -1 & -1 & 0 & 1 & 0 & 0 & 1
\end{pmatrix},$$

$$\mathbf{P}_2^{phf} = H^{\otimes 3}\mathbf{P}_2^{bf}H^{\otimes 3} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 & -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 & 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix},$$

and

$$\mathbf{P}_3^{phf} = H^{\otimes 3}\mathbf{P}_3^{bf}H^{\otimes 3} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}.$$

Call $\mid \zeta_i^{phf}\rangle$ the codeword with $i = 1, 2, 3$ corresponding to the phase-flip of the first, second or the third qubit of the quantum codeword $\mid \zeta\rangle = \alpha_0 \mid +++\rangle + \alpha_1 \mid ---\rangle$, respectively and $i = 0$ corresponding to no error. The error syndromes for the four possible cases are respectively:

| | |
|---|---|
| $\langle \zeta_i^{phf} \mid \mathbf{P}_0^{phf} \mid \zeta_i^{phf}\rangle,\ 0 \le i \le 3$ | $\langle \zeta_i^{phf} \mid \mathbf{P}_1^{phf} \mid \zeta_i^{phf}\rangle,\ 0 \le i \le 3$ |
| $\langle \zeta_i^{phf} \mid \mathbf{P}_2^{phf} \mid \zeta_i^{phf}\rangle,\ 0 \le i \le 3$ | $\langle \zeta_i^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_i^{phf}\rangle,\ 0 \le i \le 3$ |

There is a one-to-one correspondence between the error syndrome and the qubit in error, if one error is present; we discuss only the phase-flip of the third qubit and show that

$$\langle \zeta_0^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_0^{phf}\rangle = \langle \zeta_1^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_1^{phf}\rangle = \langle \zeta_2^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_2^{phf}\rangle = 0,$$

and

$$\langle \zeta_3^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_3^{phf}\rangle = 1.$$

$$\mid \zeta_3^{phf}\rangle = \alpha_0 \mid ++-\rangle + \alpha_1 \mid --+\rangle = \frac{1}{2\sqrt{2}}\begin{pmatrix} \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \end{pmatrix}$$

Now

$$\mathbf{P}_3^{phf} \mid \zeta_3^{phf}\rangle = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix} \frac{1}{2\sqrt{2}} \begin{pmatrix} \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \end{pmatrix}$$

$$= \frac{1}{8\sqrt{2}} \begin{pmatrix} 4\alpha_0 + 4\alpha_1 \\ -4\alpha_0 + 4\alpha_1 \\ 4\alpha_0 - 4\alpha_1 \\ -4\alpha_0 - 4\alpha_1 \\ 4\alpha_0 - 4\alpha_1 \\ -4\alpha_0 - 4\alpha_1 \\ 4\alpha_0 + 4\alpha_1 \\ -4\alpha_0 + 4\alpha_1 \end{pmatrix} = \frac{1}{2\sqrt{2}} \begin{pmatrix} \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ -\alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ -\alpha_0 + \alpha_1 \end{pmatrix} = \mid \zeta_3^{phf}\rangle$$

and

$$\langle \zeta_3^{phf} \mid \zeta_3^{phf}\rangle = 1$$

In case of no error

$$\mid \zeta_0^{phf}\rangle = \alpha_0 \mid +++\rangle + \alpha_1 \mid ---\rangle = \frac{1}{2\sqrt{2}} \begin{pmatrix} \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \end{pmatrix}$$

and

$$\mathbf{P}_3^{phf} \mid \zeta_0^{phf}\rangle = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix} \frac{1}{2\sqrt{2}} \begin{pmatrix} \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 + \alpha_1 \\ \alpha_0 - \alpha_1 \end{pmatrix}$$

$$\mathbf{P}_3^{phf} \mid \zeta_0^{phf}\rangle = \frac{1}{8\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Thus, $\langle \zeta_0^{phf} \mid \mathbf{P}_3^{phf} \mid \zeta_0^{phf}\rangle = 0$. In summary, when we apply a Hadamard Transform to each qubit we transform phase-flip errors to bit-flip-errors; then we correct the bit-flip errors, and finally apply another Hadamard Transform to each qubit of the register to return to the original basis.

## 5.7 The Nine Qubit Error Correcting Code of Shor

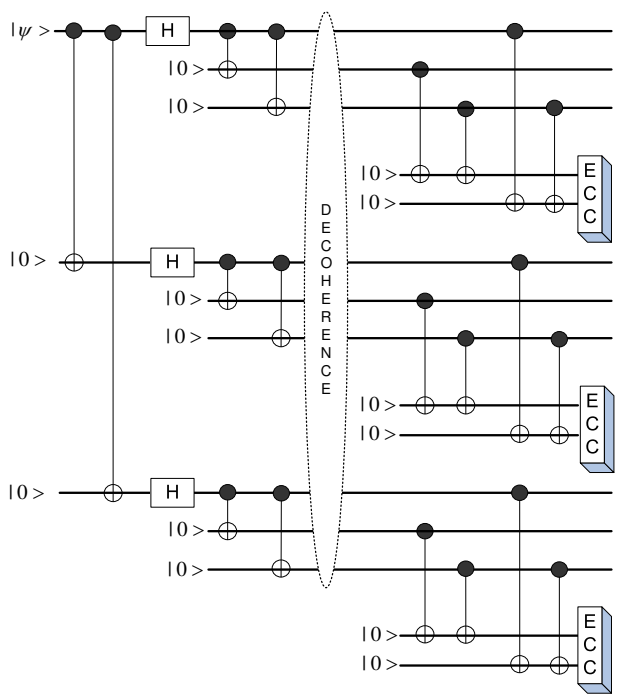



Figure 121: A circuit for Shor's algorithm.

The repetitive quantum code presented in Section 5.5 can only correct a single bit-flip error and the one in Section 5.6 can only correct a single phase-flip error. Shor created a nine qubit quantum error correcting code able to correct a single bit-flip error, a single phase-flip error, or a single bit- and phase-flip error; the distance of the code is 3.

A qubit is first encoded using the phase-flip error correction code in Section 5.6 then each of the three qubits is encoded using the circuit for bit-flip error correction discussed in Section 5.5. The quantum circuits for syndrome calculation for each group of three qubits are shown

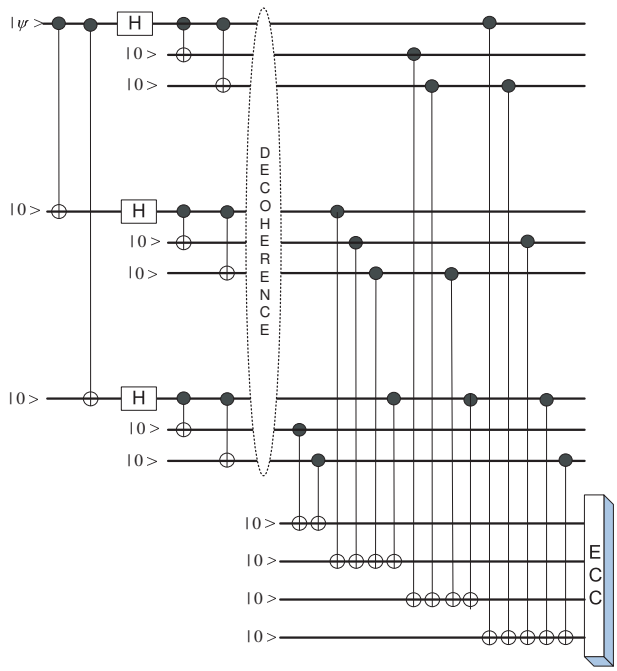



Figure 122: An alternative circuit for syndrome decoding for Shor's algorithm which uses only four ancilla qubits while the one in Figure 121 needs six. Each one of the four ancilla qubits reflects the transformations expressed by one row of the parity check matrix of the code.

in Figure 121; the qubits are counted left to right starting with qubit 0. The circuit uses six ancilla qubits, two for each group of three qubits though we only need four ancilla qubits for he computation of the syndrome as seen in Figure 122.

First, a qubit is encoded for phase-flip error correction, then each of the resulting three qubits is encoded for bit-flip error correction. The two basis vectors in $\mathcal{H}_2$, $\mid 0\rangle, \mid 1\rangle\}$ are transformed in $\{\{\mid 0\rangle_L, \mid 1\rangle_L\}$ in $\mathcal{H}_{2^9}$; the encoding is

$$\begin{array}{ccccccc} & & Phase-flip & & & & Bit-flip \\ \mid 0\rangle & \mapsto & \mid +++\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle+|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle+|1\rangle}{\sqrt{2}} & \mapsto & \mid 0\rangle_L = \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}} \\ \mid 1\rangle & \mapsto & \mid ---\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}} \otimes \frac{|0\rangle-|1\rangle}{\sqrt{2}} & \mapsto & \mid 1\rangle_L = \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}} \end{array}$$

The parity check matrix for this code is

$$\mathrm{H} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

The error syndrome identifies the qubit in error; for example, when the fifth qubit is in error then the error vector is $e_5 = (000010000)$ then the syndrome is $\mathrm{H}e_5^T = (0101)$ and $(0101)_2 = 5$.

## 5.8 The Seven Qubit Error Correcting Code of Steane

Now we turn our attention to a quantum system where we encode a single qubit as a block of seven qubits and correct one bit-flip or one phase-flip error. Steane's approach is to encode an arbitrary state in a two-dimensional Hilbert space, $\mathcal{H}_2$ by embedding $\mathcal{H}_2$ into $\mathcal{H}_{2^7}$. Steane's code is based on a $(7, 4, 3)$ Hamming code $\mathcal{C}$ with the generator matrix G given by

$$\mathrm{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = [I_4 A] \quad \text{with} \quad A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

$\mathcal{C}$ is a code which maps $k = 4$ information symbols to codewords of length $n = 7$; the $2^4 = 16$ codewords of $\mathcal{C}$ are generated by multiplying all 4-tuples $v_i,\ 1 \le i \le 16$ with the generator matrix G of the code, $\mathcal{C} = \{v_i\mathrm{G},\ 1 \le i \le 16\}$, as shown in Table 20.

The four information symbols are the prefix of each codeword; the subset of codewords of $\mathcal{C}$ marked $\perp$ are also the codewords of the dual code, $\mathcal{C}^\perp$. The distance of code $\mathcal{C}$ is equal to the weight of the minimum weight codeword $d = 3$; the number of linearly independent columns of G is equal to 4.

To construct the dual code, $\mathcal{C}^\perp$, we first derive H, the parity check matrix of $\mathcal{C}$, which is the generator matrix of the dual code, $\mathcal{C}^\perp$; from Section 4.5 we know that

$$\mathrm{H} = [A^T I_{n-k}] = [A^T I_3] = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The codewords of $\mathcal{C}^\perp$ are constructed using H as the generator matrix as shown in Table 21. We observe that $\mathcal{C}^\perp \subset \mathcal{C}$ as $\mathcal{C}^\perp$ consists of all the even codewords of $\mathcal{C}$.

Code $\mathcal{C}$ can be used to encode one bit of information to 7 bits. If the $i$-th bit of a codeword $c \in \mathcal{C}$ is affected by an error, instead of $c$ we receive a different 7-tuple, $(c + e_i)$.

The error 7-tuple, $e_i$, has only one 1 in the position $i$ in error. There is a one-to-one correspondence between an error $e_i$ and the syndrome $\Sigma_i$

$$\Sigma_i = \mathrm{H}(c + e_i)^T = \mathrm{H}{e_i}^T \quad \text{as} \quad \mathrm{H}c^T = 0 \ \text{ when } \ c \in \mathcal{C}.$$

**Encoding for Steane's code.** The logical zero qubit is the superposition of all codewords with an even number of 1's (even codewords) of the Hamming code

$$\begin{aligned} |\,0\rangle_L = \tfrac{1}{\sqrt{8}}(&|\,0000000\rangle + |\,0011101\rangle + |\,0100111\rangle + |\,0111010\rangle + |\,1001110\rangle + |\,1010011\rangle \\ &+ |\,1101001\rangle + |\,1110100\rangle) \end{aligned}$$

The logical one qubit is the superposition of all codewords with an odd number of 1's (odd codewords) of the Hamming code

$$\begin{aligned} |\,1\rangle_L = \tfrac{1}{\sqrt{8}}(&|\,1111111\rangle + |\,1100010\rangle + |\,1011000\rangle + |\,1000101\rangle + |\,0110001\rangle + |\,0101100\rangle \\ &+ |\,0010110\rangle + |\,0001011\rangle). \end{aligned}$$

Table 20: Mapping of information symbols to the codewords for the code $\mathcal{C}$ used by the seven qubit quantum code of Steane.

| $i$ | $v_i$ | | | | $\mathcal{C}$ |
|---|---|---|---|---|---|
| 1 | 0000 | | | | **0000**000 $\perp$ |
| 2 | 0001 | | | | **0001**111 $\perp$ |
| 3 | 0010 | | | | **0010**110 |
| 4 | 0011 | | | | **0011**001 |
| 5 | 0100 | | | | **0100**101 |
| 6 | 0101 | | | | **0101**010 |
| 7 | 0110 | | | | **0110**011 $\perp$ |
| 8 | 0111 | $\times$ | $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$ | $\mapsto$ | **0111**100 $\perp$ |
| 9 | 1000 | | | | **1000**011 |
| 10 | 1001 | | | | **1001**100 |
| 11 | 1010 | | | | **1010**101 $\perp$ |
| 12 | 1011 | | | | **1011**010 $\perp$ |
| 13 | 1100 | | | | **1100**110 $\perp$ |
| 14 | 1101 | | | | **1101**001 $\perp$ |
| 15 | 1110 | | | | **1110**000 |
| 16 | 1111 | | | | **1111**111 |

Table 21: Mapping information symbols to the codewords for the dual code $\mathcal{C}^{\perp}$ used by the seven qubit quantum code of Steane.

| $v'$ | | | | $\mathcal{C}^{\perp}$ |
|---|---|---|---|---|
| 000 | | | | 0000**000** |
| 001 | | | | 1101**001** |
| 010 | | | | 1011**010** |
| 011 | $\times$ | $\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$ | $\mapsto$ | 0110**011** |
| 100 | | | | 0111**100** |
| 101 | | | | 1010**101** |
| 110 | | | | 1100**110** |
| 111 | | | | 0001**111** |

The quantum circuit in Figure 123 encodes a qubit $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as a block of 7 qubits, $\mid \varphi\rangle = \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$; two `CNOT` gates encode the state $\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as the state $\alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$, a superposition of even and odd Hamming codewords. The three Hadamard gates prepare an equal superposition of the all eight possible values of the last three qubits. The state of the system at this stage is

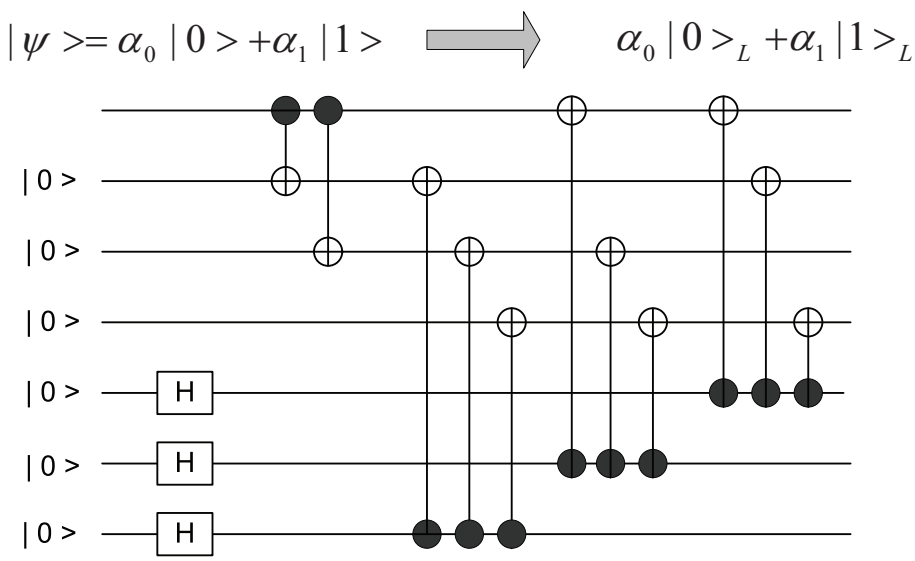



Figure 123: A circuit for encoding the qubit $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as a block of seven qubits $\mid \varphi\rangle = \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$ for Steane's code. The two `CNOT` gates to encode the state $\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as the state $\alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle$. The three Hadamard gates prepare an equal superposition of the all eight possible values of the last three qubits. The three groups of `CNOT`s gates switch on the parity bits dictated by $H$; the first group of `CNOT`s corresponds to the prefix (0 1 1 1) of the first row, the second group to the prefix (1 0 1 1) of the second row, and the third group to the prefix (1 1 0 1) of the third row.

$$\mid \xi\rangle = (\alpha_0 \mid 000\rangle + \alpha_1 \mid 111\rangle) \otimes (\mid 0\rangle) \otimes \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \otimes \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}}$$

or

$$\begin{array}{lll} \mid \xi\rangle = \frac{1}{\sqrt{8}}[ & \alpha_0( & \mid 0000000\rangle + \mid 0000001\rangle + \mid 0000010\rangle + \mid 0000011\rangle \\ & + & \mid 0000100\rangle + \mid 0000101\rangle + \mid 0000110\rangle + \mid 0000111\rangle) \\ & +\alpha_1( & \mid 1110000\rangle + \mid 1110001\rangle + \mid 1110010\rangle + \mid 1110011\rangle \\ & + & \mid 1110100\rangle + \mid 1110101\rangle + \mid 1110110\rangle + \mid 1110111\rangle)]. \end{array}$$

The three groups of `CNOT`s gates act on $\mid \xi\rangle$ and switch on the parity bits dictated by the parity check matrix $H$; the first group of `CNOT`s corresponds to the prefix (0 1 1 1) of the first row, the second group to the prefix (1 0 1 1) of the second row, and the third group to the prefix (1 1 0 1) of the third row; the first four bits of the Hamming codeword are identical to the four information bits.

**Error correction.** Let $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ and $\mid A\rangle_{env}$ be the states of the qubit and of the environment, respectively. The decoherence is caused by the entanglement of the qubit with the environment thus, the effect of an error is captured by a linear combination of the four states corresponding to: no error, bit-flip, phase-flip, and both bit- and phase-flip [342]:

$$
\begin{aligned}
\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle \quad \mapsto \quad & (\alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle) \otimes \mid A_{no\ error}\rangle_{env} \\
+ \quad & (\alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle) \otimes \mid A_{bit-flip}\rangle_{env} \\
+ \quad & (\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle) \otimes \mid A_{phase-flip}\rangle_{env} \\
+ \quad & (\alpha_0 \mid 1\rangle - \alpha_1 \mid 0\rangle) \otimes \mid A_{bit+phase-flip}\rangle_{env}.
\end{aligned}
$$

The system "qubit+environment" evolves according to a unitary transformation, but we make no assumption regarding the orthogonality and normalization of the states $\mid A_{x\ldots x}\rangle_{env}$.

We perform a measurement which projects the current state of the qubit onto the basis formed by the four states and determine the effect of the environment on the qubit, no error, bit-flip, phase-flip, bit and phase-flip. To correct a bit-flip, phase-flip, or bit and phase-flip error we apply to the current state of the qubit the $X, Z, Y$ transformations, respectively. In this manner we restore the qubit to its original state and disentangle the qubit from its environment. The correction circuit consists of the $I, X, Z, Y$ gates activated by the result of the measurement of the syndrome.

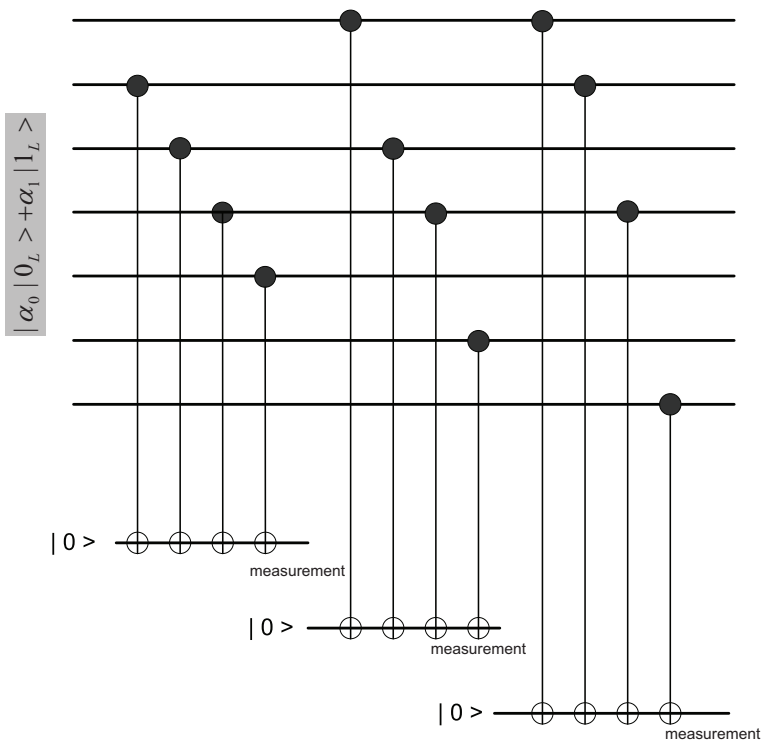



Figure 124: A circuit for calculating the syndrome for bit-flip errors for the 7-qubit Steane code. The input is the encoded information, the codeword $\mid \varphi\rangle = \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$. The 7 qubits of the codeword act as the control qubits of the 12 `CNOT` gates which set the ancilla qubits; then we perform a non-demolition measurement of the syndrome. If a phase error occurs in one of the ancilla qubits used as the target of a `CNOT` gate, then the error can "back-propagate" to the control qubits and alter the state of the subset of the 7 physical qubits which act as control qubits for the ancilla.

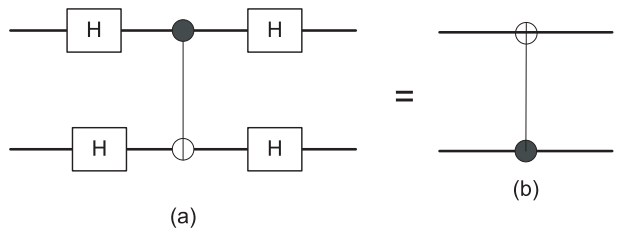



Figure 125: Back error propagation. We can reverse the role of a target and control qubit of a CNOT gate with Hadamard gates. The rotation of the basis performed by the Hadamard gate transforms a bit-flip error to a phase-flip error. When a phase error occurs in one qubit used as the target of a CNOT gate the error can propagate "backwards" to the control qubit.

The circuit in Figure 124 is used for calculating the bit-flip syndrome for the 7-qubit Steane code with the parity check matrix $H$. Each ancilla qubit is determined by the correspond row of the parity check matrix: the first is determined by (0 1 1 1 1 0 0), the control qubits of the CNOTs are the 2th, 3th, 4th, and 5th qubits of the codeword; the second by (1 0 1 0 0 1 0); and the third by (1 0 1 0 1 0 1).

We need at least three ancilla qubits to identify the individual qubits of the block of 7 qubits. If $\mid \xi\rangle$ is the state of the ensemble consisting of the codeword initially in state $\mid \xi\rangle$ and the ancilla qubits, initially in state, $\mid 000\rangle$, then the ensemble codeword+ancilla qubits is transformed as follows:

$$\mid \xi\rangle \otimes \mid 0\rangle_{ancilla} \mapsto \mid \xi\rangle \otimes \mid H\xi\rangle_{ancilla}$$

where $\mid H\xi\rangle_{aniclla}$ is the final state of the three ancilla qubits used to compute the syndrome. If only one of the 7 qubits of a block has suffered a bit-flip the measurement of the three ancilla qubits identifies which one of the qubits of the block is in error, without revealing anything about the quantum information, the state $\mid \psi\rangle$, encoded to the block.

The measurement should be non-destructive, it should preserves the code subspace. To ensure a non-demolition measurement we have to prepare the ancilla qubits in a special state. Then we "write" the parity qubits to the ancilla and then we measure the ancilla. Clearly, the parity bits of a Hamming code are either 0 or 1, thus we can "write" the parity qubits to the ancilla. To diagnose phase-flip errors we have to repeat the process in a rotated base.

It is now a good time to observe that we need more than three ancilla qubits, even though 3 bits are sufficient to express integers in the range 1 to 7. The circuit to compute the ancilla qubits for the Steane's code (Figure 124) uses one ancilla qubit as target of multiple CNOT gates. Such a circuit is not fault-tolerant.

If the control qubit of a CNOT gate is in error, then it can propagate "forward" to the target qubit. Recall that we can reverse the role of a target and control qubit of a CNOT gate with Hadamard gates, see Figure 125 (a) [284]. We also know that the rotation of the basis performed by the Hadamard gate transforms a bit-flip error to a phase-flip error. Thus, if a phase error occurs in one qubit used as the target of a CNOT gate, then the error can propagate "backwards" to the control qubit. The result of the CNOT transformation corresponds to a virtual bit-flip error of the control qubit. When a single qubit is the target of multiple CNOT gates *an error of the target qubit will affect all the control qubits.* The fault-tolerant version

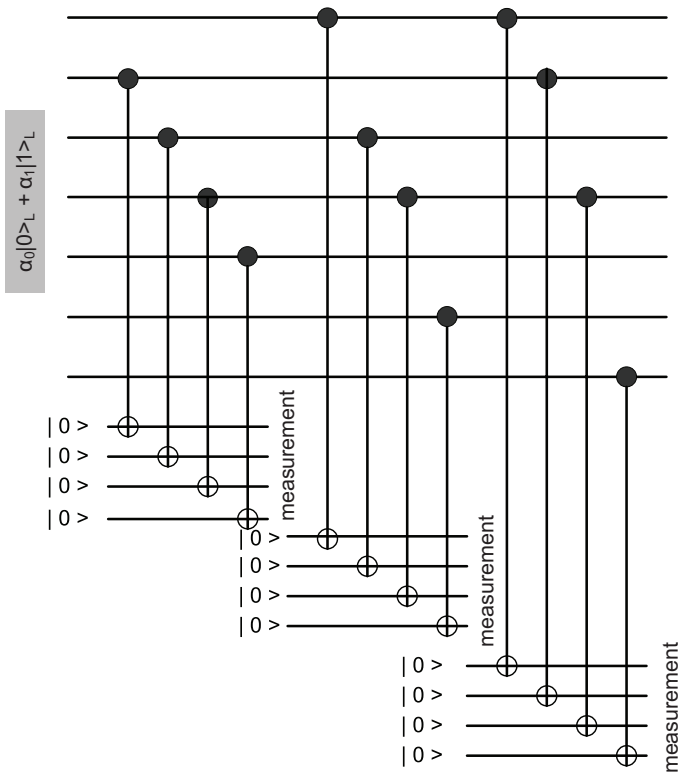



Figure 126: The fault-tolerant version of the circuit for calculating the bit-flip syndrome for the 7-bit Steane code in Figure 124. Each one of the 12 ancilla qubits is the target of a single `CNOT` gate. The circuit projects the state $(\alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L)$ onto one of the basis states, $\mid 0\rangle_L$ and $\mid 1\rangle_L$.

of the circuit for calculating the bit-flip syndrome for the 7-bit Steane code in Figure 126 has 12 rather than 3 ancilla qubits.

In summary, Steane's code uses logical $\mid 0\rangle_L$ and $\mid 1\rangle_L$ qubits which are superpositions of Hamming codewords. In the original and the rotated basis we can thus perform the parity check which allows us to diagnose and then correct a single qubit error.

## 5.9 An Inequality for Representations in Different Bases

Classical error correcting codes satisfy the bounds discussed in Section 4.7; these bounds relate the number of codewords, $\mid \mathcal{C} \mid$ of a code $\mathcal{C}$, the length $n$ of a codeword, and the distance $d$ of a code. From the discussion of Shor and Steane codes we know that a quantum error correcting code requires two classical error correcting codes, one to correct bit-flip and the other to correct phase-flip errors. In this section we discuss the "entropic uncertainty relation" introduced in [56] and [115] that provides an upper bound on the largest distance achievable in the two bases required to correct both bit-flip ( $\mid 0\rangle \leftrightarrow \mid 1\rangle$ ) and phase-flip ($\mid 0\rangle + \mid 1\rangle \leftrightarrow \mid 0\rangle - \mid 1\rangle$) errors.

The transformation performed by a Hadamard gate, $H$, maps basis $B_1 = \{\mid 0\rangle, \mid 1\rangle\}$ to

$B_2 = \{| +\rangle, | -\rangle\}$ with

$$| +\rangle = H \mid 0\rangle = \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \quad \text{and} \quad | -\rangle = H \mid 1\rangle = \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}$$

This rotation of basis transforms phase-flip errors to bit-flip errors and viceversa:

$$(\mid 0\rangle + \mid 1\rangle)_{B_1} \mapsto \mid 0\rangle_{B_2} \quad \text{and} \quad (\mid 0\rangle - \mid 1\rangle)_{B_1} \mapsto \mid 1\rangle_{B_2}.$$

In $\mathcal{H}_{2^n}$ we consider basis $B_1$ with basis vectors $\{| e_0\rangle, | e_1\rangle \ldots | e_{2^n-1}\rangle\}$ and $B_2$ with basis vectors $\{| e'_0\rangle, | e'_1\rangle \ldots | e'_{2^n-1}\rangle\}$. Table 22 shows states $| e'_0 e'_1 \ldots e'_{2^n-1}\rangle_{B_2}$ represented as a superposition of basis vectors $\{| e_0\rangle, | e_1\rangle \ldots | e_{2^n-1}\rangle\}$.

We say that a basis vector $| u_j\rangle$ has an "odd" parity if the number of 1s in the binary representation of integer $u_j$ is an odd number and $| u_j\rangle$ has an "even" parity if the number of 1s in the binary representation of integer $u_j$ is an even number. For example, $| 001\rangle$ has odd parity and $| 101\rangle$ has even parity. It is easy to infer and then verify the following statements:

**Statement 1.** The $| 00\ldots0\rangle_{B_2}$ state is a superposition of all basis vectors in $B_1$ with equal and positive coefficients.

**Statement 2.** The $| 11\ldots1\rangle_{B_2}$ state is a superposition of all basis vectors in $B_1$ with equal coefficients of positive (even parity) and negative (odd parity) sign.

**Statement 3.** The $(| 00\ldots0\rangle + | 11\ldots1\rangle)_{B_2}$ state is a superposition of basis vectors in $B_1$ with even parity. All coefficients are equal and positive.

**Statement 4.** The $(| 00\ldots0\rangle - | 11\ldots1\rangle)_{B_2}$ state is a superposition of basis vectors in $B_1$ with odd parity. All coefficients are equal and positive.

**Statement 5.** If the $i$-th bit in the state $(| 00\ldots0\rangle + | 11\ldots1\rangle)_{B_2}$ is flipped, then the mapping of this state is similar with the original superposition, but the coefficients of the basis states when this bit is equal to 1 are negative. For example, when the first bit is flipped the state becomes $| 100\rangle + | 011\rangle$ and then, instead of the superposition $\frac{1}{\sqrt{2}}(| 000\rangle + | 011\rangle) + | 101\rangle + | 110\rangle)$ we have $\frac{1}{\sqrt{2}}(| 000\rangle + | 011\rangle) - | 101\rangle - | 110\rangle)$.

Let $| \psi\rangle = \alpha_0 \mid e_0\rangle + \alpha_1 \mid e_1\rangle + \ldots + \alpha_{2^n-1} \mid e_{2^n-1}\rangle$ be a vector in a Hilbert space, $\mathcal{H}_{2^n}$. The basis vectors are $| e_i\rangle = | 2^i - 1\rangle, \quad 0 \le i \le 2^n - 1$. Call $w(| \psi\rangle)_i$ the <u>weight</u> of the state vector in basis $B_i, i = 1, 2$. The weight is equal to the number of non-zero complex numbers $\alpha_i, \quad 0 \le i \le 2^n - 1$, or the number of basis vectors used to express $| \psi\rangle$ in $B_i, i = 1, 2$.

**Proposition.** *The state of an n-qubit system can be expressed as a superposition of $w_1$ basis vectors in base $B_1$ and as a superposition of $w_2$ basis vectors in a rotated base, $B_2$. The weights $w_1$ and $w_2$ satisfy the following inequality:*

$$w_1 \cdot w_2 \ge 2^n.$$

*Proof:* Recall that the Hadamard gate performs a Walsh-Hadamard Transform of one qubit

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Table 22: Changes of basis in $\mathcal{H}_{2^n}$ from $B_1 = \{\mid e_0\rangle, \mid e_1\rangle \ldots \mid e_{2^n-1}\rangle\}$ to $B_2 = \{\mid e'_0\rangle, \mid e'_1\rangle \ldots \mid e'_{2^n-1}\rangle\}$. The states $\mid b_0 b_1 \ldots b_{2^n-1}\rangle_{B_2}$ represented as a superposition of bases vectors $\{\mid e_0\rangle, \mid e_1\rangle \ldots \mid e_{2^n-1}\rangle\}$ in $B_1$ for $n = 1, 2, 3$.

| *Base* $B_2$ | | *Base* $B_1$ |
|---|---|---|
| $n = 1$ | | |
| $\mid +\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$ |
| $\mid -\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle)$ |
| $n = 2$ | | |
| $\mid ++\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) = \frac{1}{2}(\mid 00\rangle + \mid 01\rangle + \mid 10\rangle + \mid 11\rangle)$ |
| $\mid --\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) = \frac{1}{2}(\mid 00\rangle - \mid 01\rangle - \mid 10\rangle + \mid 11\rangle)$ |
| $\mid ++\rangle + \mid --\rangle$ | $\mapsto$ | $\mid 00\rangle + \mid 11\rangle$ |
| $\mid ++\rangle - \mid --\rangle$ | $\mapsto$ | $\mid 01\rangle + \mid 10\rangle$ |
| $\mid -+\rangle + \mid +-\rangle$ | $\mapsto$ | $\mid 00\rangle - \mid 11\rangle$ |
| $n = 3$ | | |
| $\mid +++\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{2}(\mid 00\rangle + \mid 01\rangle + \mid 10\rangle + \mid 11\rangle)$ <br> $= \frac{1}{2\sqrt{2}}(\mid 000\rangle + \mid 001\rangle + \mid 010\rangle + \mid 011\rangle + \mid 100\rangle + \mid 101\rangle + \mid 110\rangle + \mid 111\rangle)$ |
| $\mid ---\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{2}(\mid 00\rangle - \mid 01\rangle - \mid 10\rangle + \mid 11\rangle)$ <br> $= \frac{1}{2\sqrt{2}}(\mid 000\rangle - \mid 001\rangle - \mid 010\rangle + \mid 011\rangle - \mid 100\rangle + \mid 101\rangle + \mid 110\rangle - \mid 111\rangle)$ |
| $\mid +++\rangle + \mid ---\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 000\rangle + \mid 011\rangle + \mid 101\rangle + \mid 110\rangle)$ |
| $\mid +++\rangle - \mid ---\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 001\rangle + \mid 010\rangle + \mid 100\rangle + \mid 111\rangle)$ |
| $\mid ++-\rangle + \mid --+\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle)$ <br> $+ \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$ <br> $= \frac{1}{\sqrt{2}}(\mid 000\rangle - \mid 011\rangle + \mid 110\rangle - \mid 101\rangle)$ |
| $\mid -++\rangle + \mid +--\rangle$ | $\mapsto$ | $\frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$ <br> $+ \frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle) \otimes \frac{1}{\sqrt{2}}(\mid 0\rangle - \mid 1\rangle)$ <br> $= \frac{1}{\sqrt{2}}(\mid 000\rangle + \mid 011\rangle - \mid 110\rangle - \mid 110\rangle)$ |

and that $H^{\otimes n}$ denotes the Walsh-Hadamard Transform applied to $n$ qubits. Consider the state:

$$\mid \psi\rangle = \alpha_0 \mid e_0\rangle + \alpha_1 \mid e_1\rangle + \ldots + \alpha_{2^n-1} \mid e_{2^n-1}\rangle$$

expressed in $B_1$ with basis vectors $\mid e_0\rangle, \mid e_1\rangle, \ldots \mid e_{2^n-1}\rangle$. The same state is expressed in base $B_2$ as:

$$\mid \varphi\rangle = H^{\otimes n} \mid \psi\rangle$$

The respective weights are:

$$w_1 = w(\mid \psi\rangle) \qquad \text{and} \qquad w_2 = w(\mid \varphi\rangle)$$

Our proof is by induction. For $n = 1$ the vectors in $B_1$ and $B_2$ are respectively:

$$\mid 0\rangle \;\mapsto\; \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \;\; \mid 1\rangle \;\mapsto\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \;\mapsto\; \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}} \;\mapsto\; \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Thus for $n = 1$ the relationship is true: $w_1 = 1$, $w_2 = 2$, and, $w_1 \cdot w_2 = 1 \cdot 2 = 2 \geq 2$.

Now assume that this is true for $n = k$ and attempt to prove it for $n = k+1$. We represent the integer $k$ by a binary string of length $n$, denoted as $k_n$. For example, for $n = 2$ we write the basis vectors as follows:

$$\mid \psi\rangle = \alpha_0 \mid 00\rangle + \alpha_1 \mid 01\rangle + \alpha_2 \mid 10\rangle + \alpha_3 \mid 11\rangle \quad \mapsto \quad \mid \psi\rangle = \alpha_0 \mid 0_2\rangle + \alpha_1 \mid 1_2\rangle + \alpha_2 \mid 2_2\rangle + \alpha_3 \mid 3_2\rangle.$$

We observe that:

$$\begin{array}{rll} \mid \psi\rangle = & \alpha_0 \mid 0_{n+1}\rangle + \alpha_1 \mid 1_{n+1}\rangle + \ldots + \alpha_{2^{n+1}-1} \mid (2^{n+1}-1)_{n+1}\rangle & = \\ & \mid 0_1\rangle \otimes [\alpha_0 \mid 0_n\rangle + \alpha_1 \mid 1_n\rangle + \ldots + \alpha_{2^n-1} \mid (2^n-1)_n\rangle] & + \\ & \mid 1_1\rangle \otimes [\alpha_{2^n} \mid 0_n\rangle + \alpha_{2^n+1} \mid 1_n\rangle + \ldots + \alpha_{2^{n+1}-1} \mid (2^n-1)_n\rangle] & \end{array}$$

We use the following notations:

$$\mid \psi_F\rangle = \alpha_0 \mid 0_n\rangle + \alpha_1 \mid 1_n\rangle + \ldots + \alpha_{2^n-1} \mid (2^n-1)_n\rangle$$

and

$$\mid \psi_L\rangle = \alpha_{2^n} \mid 0_n\rangle + \alpha_{2^n+1} \mid 1_n\rangle + \ldots + \alpha_{2^{n+1}-1} \mid (2^n-1)_n\rangle.$$

Thus:

$$\mid \psi\rangle = \mid 0_1\rangle \otimes \mid \psi_F\rangle + \mid 1_1\rangle \otimes \mid \psi_L\rangle.$$

Now

$$\begin{array}{rcl} \mid \varphi\rangle & = & H^{\otimes k} \mid \psi\rangle \\ & = & (H \otimes H^{\otimes k})[\mid 0_1\rangle \otimes \mid \psi_F\rangle + \mid 1_1\rangle \otimes \mid \psi_L\rangle] \\ & = & (\mid 0_1\rangle + \mid 1_1\rangle) H^{\otimes k} \mid \psi_F\rangle + (\mid 0_1\rangle - \mid 1_1\rangle) H^{\otimes k} \mid \psi_L\rangle \\ & = & \mid 0_1\rangle (H^{\otimes k} \mid \psi_F\rangle + H^{\otimes k} \mid \psi_L) + \mid 1_1(\rangle H^{\otimes k} \mid \psi_F\rangle - H^{\otimes k} \mid \psi_L\rangle) \end{array}$$

Assume that $\mid \psi_F\rangle = 0$ and, then, by inductive hypothesis, $\mid \psi_L\rangle \neq 0$:

$$w(\mid \psi\rangle) \cdot w(\mid \varphi\rangle) = 2 \cdot w(\mid \psi_L\rangle) \cdot w(H^{\otimes k} \mid \psi_L\rangle) \geq 2 \cdot 2^k = 2^{k+1}.$$

Assume that $\mid \psi_L\rangle = 0$ and $\mid \psi_F\rangle \neq 0$. A similar derivation allows us to prove that in this case $w_1 \cdot w_2 = 2^{k+1}$.

Now consider the case $\mid \psi_F\rangle \neq 0$ and $\mid \psi_L\rangle \neq 0$. Then

$$w(\mid \psi_F\rangle) \cdot w(H^{\otimes k} \mid \psi_F\rangle) \geq 2^k \quad \text{and} \quad w(\mid \psi_L\rangle) \cdot w(H^{\otimes k} \mid \psi_L\rangle) \geq 2^k.$$

Also, if $\alpha_i \neq 0$ then $\alpha_i + \alpha_{i+2^k} \neq 0$, or $\alpha_i - \alpha_{i+2^k} \neq 0$, where we have addition modulo $2^{k+1}$ for the subscripts. Thus,

$$w(\mid \varphi\rangle) \geq w(H^{\otimes k} \mid \psi_F\rangle) \quad \text{and} \quad w(\mid \varphi\rangle) \geq w(H^{\otimes k} \mid \psi_L\rangle).$$

Finally, in this case

$$\begin{aligned} w(\mid \psi\rangle)w(\mid \varphi\rangle) &= [w(\mid \psi_F\rangle) + w(\mid \psi_L\rangle)] \cdot w(\mid \varphi\rangle) \\ &= w(\mid \psi_F\rangle) \cdot w(\mid \varphi\rangle + w(\mid \psi_L\rangle) \cdot w(\mid \varphi\rangle\rangle) \\ &\geq w(\mid \psi_F\rangle) \cdot w(H^{\otimes k} \mid \phi_F\rangle) + w(\mid \psi_L\rangle) \cdot w(H^{\otimes k} \mid \phi_L\rangle) \\ &\geq 2^k + 2^k \\ &\geq 2^{k+1} \end{aligned}$$

This concludes the proof that the weights $w_1$ and $w_2$ satisfy the inequality $w_1 \cdot w_2 \geq 2^n$.

□

## 5.10 Calderbank-Shor-Steane (CSS) Codes

The Calderbank-Shor-Steane (CSS) codes were discovered independently by Calderbank and Shor [81] and Steane [402]. CSS codes are a subclass of the larger class of stabilizer codes discussed in Section 5.12.

First, we review basic concepts from linear coding theory discussed in Section 4.5. Let $\mathcal{C}$ be a linear $[n, k, d]$ code over the binary field $F_2^n$ capable of correcting $t = \lfloor \frac{d-1}{2} \rfloor$ errors; $\mathcal{C}^\perp$, the dual of the code $\mathcal{C}$, is the set of all binary $n$-tuples orthogonal to all codewords of $\mathcal{C}$

$$\mathcal{C}^\perp = \{v \in F_2^n : \quad v \cdot c = 0 \ \ \forall c \in \mathcal{C}\}.$$

The generator matrix G of $\mathcal{C}$ is a $k \times n$ matrix whose row space consists of independent $n$-tuples spanning the code $\mathcal{C}$; the parity check matrix H of code $\mathcal{C}$ is an $(n-k) \times n$ matrix such that $\mathrm{H}c^T = 0, \ \ \forall c \in \mathcal{C}$. Code $\mathcal{C}$ is spanned by $k$ independent $n$-tuples, $\dim(\mathcal{C}) = k$ and $\mathcal{C}^\perp$ is spanned by $n-k$ independent $n$-tuples, $\dim(\mathcal{C}^\perp) = n-k$. If G and H are the generator and, respectively, the parity check matrix of $\mathcal{C}$ then H and G are the generator and, respectively, the parity check matrix of $\mathcal{C}^\perp$. The code $\mathcal{C}$ is weakly self-dual if $\mathcal{C}^\perp \subset \mathcal{C}$.

Calderbank and Shor proved the Quantum Error Correction Theorem relating classical and quantum error-correcting codes [81]:

**Theorem.** *If $\mathcal{C}_1$ and $\mathcal{C}_2$ are both $[n, k, d]$ codes with $\{0\} \subseteq \mathcal{C}_2^\perp \subseteq \mathcal{C}_1 \subseteq \mathcal{F}_2^n$, then the quantum code $\mathcal{Q}_{\mathcal{C}_1,\mathcal{C}_2}$ is a $t$-error correcting code, with $t = \lfloor (d-1)/2 \rfloor$.*

Steane reformulated this theorem based on several properties of $\mathcal{C}_2^\perp$: (i) the dimension of $\mathcal{C}_2^\perp$ is $(n-k)$; (ii) $\mathcal{C}_2^\perp \subseteq \mathcal{C}_1$; and (iii) $\mathcal{C}_2^\perp$ is a $K$-th order subcode of $\mathcal{C}_1$ with $K = (2k-n)$.

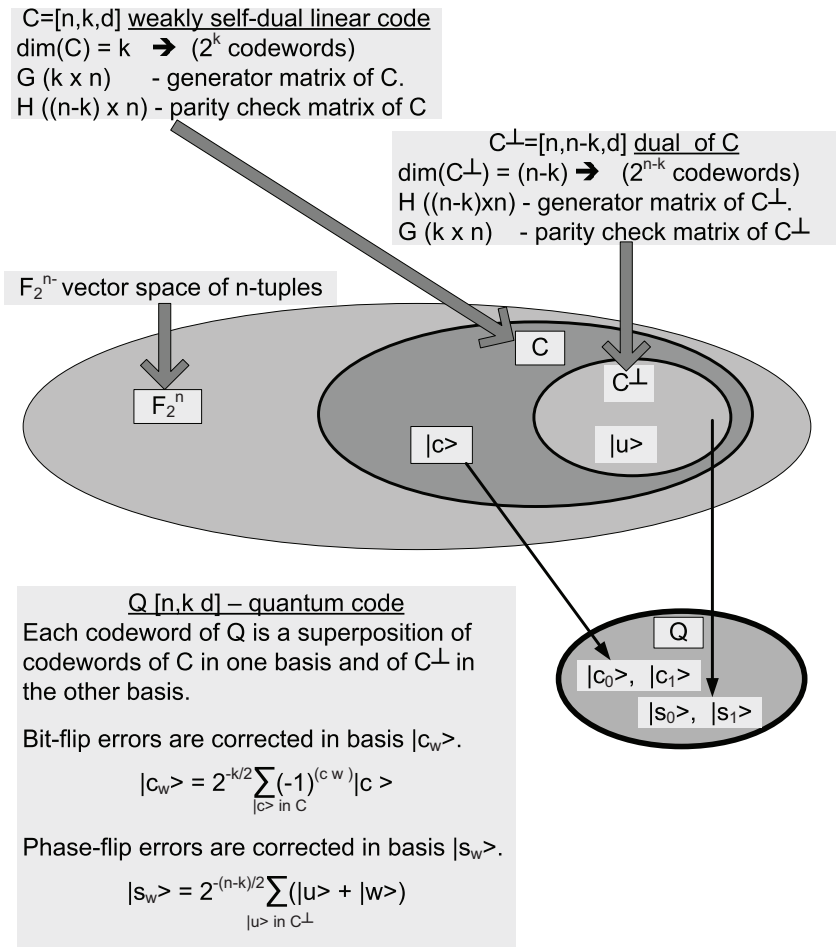



Figure 127: The construction of a CSS code $\mathcal{Q}$ based on a self-dual linear code $\mathcal{C}$.

**Theorem.** *To encode $K$ qubits with minimum distance $d_1$ in one basis and $d_2$ in another, it is sufficient to find a linear code of minimum distance $d_1$, whose $K$-th order subcode is the dual of a distance $d_2$ code.*

A quantum code $\mathcal{Q}$ is denoted by $[n, k, d_1, d_2]$ with $n$ the number of qubits used to store or transmit $k$ bits of information; $\mathcal{Q}$ allows correction of up to $\lfloor (d_1 - 1)/2 \rfloor$ bit-flip and simultaneously up to $\lfloor (d_2 - 1)/2 \rfloor$ phase-flip errors.

**Construction of a quantum code $\mathcal{Q}$ based on a classical weakly self-dual code.** The codewords of $\mathcal{Q}$, see Figure 127, are superpositions of the codewords of $\mathcal{C}$ assembled according to the following rules [81]:

(a) Construct first a $c$-basis. If $v \in \mathcal{C}$ then a quantum state $\mid c_w\rangle$ is defined as

$$\mid c_w\rangle = 2^{-\frac{\dim(\mathcal{C})}{2}} \sum_{v \in \mathcal{C}} (-1)^{v \cdot w} \mid v\rangle, \quad w \in F_2^n.$$

with "$\cdot$" the bitwise addition modulo 2, of two $n$-tuples, $v$ and $w$, $\sum_{i=1}^{n} v_i w_i \mod 2$.

(b) Construct an $s$-basis. According to (a) the codewords of the quantum code $\mathcal{Q}$ are the set of $\mid c_w\rangle \;\; \forall w \in \mathcal{C}^{\perp}$. We rotate each of the $n$ qubits of a codeword:

$$\mid 0\rangle \mapsto \frac{\mid 0\rangle + \mid 1\rangle}{\sqrt{2}} \qquad \text{and} \qquad \mid 1\rangle \mapsto \frac{\mid 0\rangle - \mid 1\rangle}{\sqrt{2}}.$$

Phase errors in the original basis become bit errors in the new basis, $\mid 0\rangle \mapsto \mid 0\rangle$ and $\mid 1\rangle \mapsto - \mid 1\rangle$. After the change of basis applied to a qubit in state $\mid c_w\rangle$ we obtain the state

$$\mid s_w\rangle = 2^{\frac{\dim(\mathcal{C})-n}{2}} \sum_{u \in \mathcal{C}^{\perp}} \mid u + w\rangle, \quad w \in F_2^n.$$

It is left as an exercise for the reader to prove that given $(\mid v\rangle, \mid u\rangle) \in \mathcal{C}$ then two basis vectors are identical $\mid s_v\rangle = \mid s_u\rangle$ and $\mid c_v\rangle = \mid c_u\rangle$ if and only if $\mid v\rangle = \mid u\rangle$. The dimension of the quantum code $\mathcal{Q}$ is

$$\dim(\mathcal{Q}) = \dim(\mathcal{C}) - \dim(\mathcal{C}^{\perp}) = k - (n-k) = 2k - n.$$

Codes $\mathcal{C}$ and $\mathcal{C}^{\perp}$ have $2^k$ and $2^{n-k}$ codewords, respectively, thus, the quantum code $\mathcal{Q}$ has $2^{2k-n}$ codewords.

We first correct the bit-flip errors in the $c$-basis then perform the rotation of basis and correct the bit-flip errors in the $s$-basis; the phase-flip errors in the $c$-basis, if any, become bit-flip errors in the $s$-basis. We can correct at most $t$ bit-flip and phase-flip errors simply because the classical code $\mathcal{C}$ can correct up to $t$ errors. Shor and Calderbank show that the two error correction steps do not interfere with each other [81]. The following example from [81] illustrates the construction of a quantum code.

**Example.** *Construct quantum code $\mathcal{Q}$ based on $\mathcal{C}$ and $\mathcal{C}^{\perp}$ with $\mathcal{C}$ the* $[7, 4, 3]$ *Hamming code*

$$\begin{array}{lllllllll} \mathcal{C} = \{ & 0000000 & 0001011 & 0010110 & 0011101 & 0100111 & 0101100 & 0110001 & 0111010 \\ & 1000101 & 1001110 & 1010011 & 1011000 & 1100010 & 1101001 & 1110100 & 1111111 \; \} \end{array}.$$

The dimension of this code is $\dim(\mathcal{C}) = 4$; the dual code, $\mathcal{C}^{\perp}$, is

$$\mathcal{C}^{\perp} = \{ \;\; 0000000 \;\; 0011101 \;\; 0100111 \;\; 0111010 \;\; 1001110 \;\; 1010011 \;\; 1101001 \;\; 1110100 \;\; \} \; .$$

$\mathcal{C}^{\perp} \subset \mathcal{C}$ thus, $\mathcal{C}$ is a weakly self-dual code; the dimension of $\mathcal{C}^{\perp}$ is

$$\dim(\mathcal{C}^{\perp}) = n - \dim(\mathcal{C}) = 7 - 4 = 3.$$

We see that $2^{\dim(\mathcal{C}) - \dim(\mathcal{C}^{\perp})} = 2$ thus, $\mathcal{Q}$ has two orthogonal vectors, $\mid c_0\rangle$ and $\mid c_1\rangle$ constructed according to the rule

$$\mid c_w\rangle = 2^{-\frac{\dim(\mathcal{C})}{2}} \sum_{c \in \mathcal{C}} (-1)^{c \cdot w} \mid c\rangle, \quad w = 0, 1.$$

When $w = 0$ then $\forall c \in \mathcal{C} \;\; c \cdot w = c \cdot (0000000) = 0$; thus, all the codewords in the sum have positive amplitudes

$$
\begin{array}{rl}
\mid c_0\rangle = & \frac{1}{4}( \quad \mid 0000000\rangle + \mid 0001011\rangle + \mid 0010110\rangle + \mid 0011101\rangle \\
& + \quad \mid 0100111\rangle + \mid 0101100\rangle + \mid 0110001\rangle + \mid 0111010\rangle \\
& + \quad \mid 1000101\rangle + \mid 1001110\rangle + \mid 1010011\rangle + \mid 1011000\rangle \\
& + \quad \mid 1100010\rangle + \mid 1101001\rangle + \mid 1110100\rangle + \mid 1111111\rangle )
\end{array} .
$$

When $w = 1$ then

$$
c \cdot w = c \cdot (1111111) = \begin{cases} +1 & \text{if the weight of c is odd} \\ 0 & \text{if the weight of c is even} \end{cases} .
$$

The codewords with an even weight have positive amplitudes while those with an odd weight have negative amplitudes

$$
\begin{array}{rl}
\mid c_1\rangle = & \frac{1}{4}( \quad \mid 0000000\rangle - \mid 0001011\rangle - \mid 0010110\rangle + \mid 0011101\rangle \\
& + \quad \mid 0100111\rangle - \mid 0101100\rangle - \mid 0110001\rangle + \mid 0111010\rangle \\
& - \quad \mid 1000101\rangle + \mid 1001110\rangle + \mid 1010011\rangle - \mid 1011000\rangle \\
& - \quad \mid 1100010\rangle + \mid 1101001\rangle + \mid 1110100\rangle - \mid 1111111\rangle )
\end{array} .
$$

We use a similar strategy to construct $\mid s_0\rangle$ and $\mid s_1\rangle$

$$
\mid s_w\rangle = 2^{\frac{\dim(\mathcal{C}^\perp)-n}{2}} \sum_{u \in \mathcal{C}^\perp} \mid u + w\rangle, \quad w = 0, 1.
$$

When $w = 0$ then $u + w = u, \quad \forall u \in \mathcal{C}^\perp$ and

$$
\begin{array}{rl}
\mid s_0\rangle = & \frac{1}{2\sqrt{2}}( \quad \mid 0000000\rangle + \mid 0011101\rangle + \mid 0100111\rangle + \mid 0111010\rangle \\
& + \quad \mid 1001110\rangle + \mid 1010011\rangle + \mid 1101001\rangle + \mid 1110100\rangle \quad )
\end{array} .
$$

When $w = 1$ then $u + w = u + (1111111) = \bar{u} \quad \forall u \in \mathcal{C}^\perp$; here $\bar{u}$ is the binary $n$-tuple with every bit complemented and

$$
\begin{array}{rl}
\mid s_1\rangle = & \frac{1}{2\sqrt{2}}( \quad \mid 1111111\rangle + \mid 1100010\rangle + \mid 1011000\rangle + \mid 1000101\rangle \\
& + \quad \mid 0110001\rangle + \mid 0101100\rangle + \mid 0010110\rangle + \mid 0001011\rangle \quad )
\end{array} .
$$

**Steane's version.** We construct a quantum code $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2)$ based on two classical codes, $\mathcal{C}_1$ and $\mathcal{C}_2 \subset \mathcal{C}_1$ [402]. Let $\mathcal{C}_2^\perp$ be the dual code of $\mathcal{C}_2$. Both $\mathcal{C}_1$ and $\mathcal{C}_2^\perp$ are capable of correcting $t = \lfloor \frac{d-1}{2} \rfloor$ errors. $\mathcal{C}_1$ is an $[n, k_1, d]$ code, $\mathcal{C}_2$ is an $[n, k_2, d']$ code, and $\mathcal{C}_2^\perp$ is an $[n, n - k_2, d]$ code. Call $\mathrm{H}_1$ and $\mathrm{H}_2$ the parity check matrices of $\mathcal{C}_1$ and $\mathcal{C}_2^\perp$, respectively.

The quantum code $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2)$ should be capable of correcting $t$ or fewer bit-flip and phase-flip errors. We construct first an $s$-basis. If $u \in \mathcal{C}_1$ then we define the state :

$$
\mid s'_{u+w}\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} \mid u + w\rangle
$$

with $w$ chosen from the cosets $\frac{\mathcal{C}_1}{\mathcal{C}_2}$. We can compute the cosets of $\mathcal{C}_2$ in $\mathcal{C}_1$ because $\mathcal{C}_2$ is an additive subgroup of $\mathcal{C}_1$, $\mathcal{C}_2 \subset \mathcal{C}_1$. There are $2^{k_2}$ elements in $\mathcal{C}_2$, thus as many coset leaders. It follows that each coset consists of:

$$\frac{\mid \mathcal{C}_1 \mid}{\mid \mathcal{C}_2 \mid} = \frac{2^{k_1}}{2^{k_2}} = 2^{k_1-k_2}$$

$n$-tuples. The corresponding table has $2^{k_1}$ elements, $2^{k_2}$ rows and $2^{k_1-k_2}$ columns. It is easy to see that the state $\mid s'_{u+w}\rangle$ depends only on the coset of $\mathcal{C}_1/\mathcal{C}_2$. Indeed, $\mid s'_{u+w}\rangle = \mid s'_{u'+w}\rangle$ if and only if $\mid u-u'\rangle \in \mathcal{C}_2$. Thus, the number of codewords in $\mathcal{Q}(\mathcal{C}_1, \mathcal{C}_2)$ is equal to $2^{k_1-k_2}$. Note that Steane's codewords $\mid s'_w\rangle$ generate the same subspace in $\mathcal{H}_2^{\otimes n}$ as the codewords $\mid c_w\rangle$ of Calderbank and Shor.

**The error correction**. The decoding scheme for CSS codes based on weakly self-dual codes is depicted in Figure 128. Assume that an error is represented by an $n$-tuple $e$ such that:

(i) Both bit-flip and phase-flip errors could be present, $\mid e\rangle = \mid e_{bf}\rangle + \mid e_{pf}\rangle$. The $n$-tuples $e_{bf}$ and $e_{pf}$ have a 1 in a position corresponding to a bit-flip and phase-flip error, respectively.

(ii) There are at most $t$ errors, $w(e) \leq t$, the weight of the error vector is less or equal to $t$, thus $w(e_{bf}) + w(e_{pf}) \leq t$.

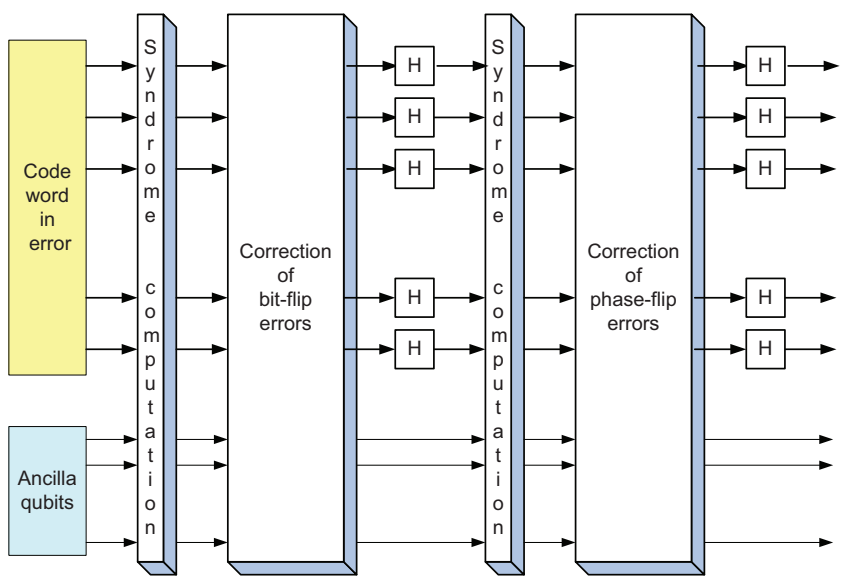



Figure 128: The decoding scheme for CSS codes constructed from weakly self-dual codes.

If the original state $\mid s'_{u+w}\rangle$ is affected by $\mid e_{bf}\rangle$ bit-flip and $\mid e_{pf}\rangle$ phase-flip errors, then the corrupted state is

$$\mid s'_{(u+w),error}\rangle = \frac{1}{2^{\frac{k_2}{2}}} \sum_{u\in\mathcal{C}_1, w\in\mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)+e_{pf}} \mid u+w+e_{bf}\rangle.$$

We use a sufficient number of ancilla qubits initially in the all-zero state $\mid 00\ldots0\rangle$ to store the syndrome for the code $\mathcal{C}_1$ . When we apply the parity check matrix $\mathrm{H}_1$ to the corrupted state we produce the state

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)+e_{pf}} \mid u+w+e_{bf}\rangle \otimes \mid \mathrm{H}_1(e_{bf})\rangle.$$

We write explicitly $\mid ab\rangle$ as $\mid a\rangle \otimes \mid b\rangle$ to emphasize the action of the parity check matrix which calculates the syndrome in the ancilla qubits. The application of $\mathrm{H}_1$ transforms the joint state of an original element of the superposition and the ancilla qubits as follows:

$$\mid u+w+e_{bf}\rangle \otimes \mid 0\rangle \overset{\mathrm{H}_1}{\mapsto} \mid u+w+e_{bf}\rangle \otimes \mid \mathrm{H}_1(u+w+e_{bf})\rangle = \mid u+w+e_{bf}\rangle \otimes \mid \mathrm{H}_1(e_{bf})\rangle.$$

The computation of the syndrome does not affect the state and that for $u+w \in \mathcal{C}_1$, $\mathrm{H}_1(u+w+e_{bf}) = \mathrm{H}_1(e_{bf})$ because $\mathrm{H}_1(u+w) = 0$; after we discard the ancilla qubits the state is

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)\cdot e_{pf}} \mid u+w+e_{bf}\rangle.$$

The ancilla qubits contain a portion of the state in error, namely $\mid \mathrm{H}_1(e_{bf})\rangle$ which allows us to correct the $e_{bf}$ bit-flip errors since code $\mathcal{C}_1$ can correct at most $t$ errors and the weight of the phase-flip error vector is at most $t$, $w(e_{bf}) \leq t$; the state after the correction of bit-flip errors is

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)\cdot e_{pf}} \mid u+w\rangle.$$

Now we correct the residual phase-flip errors, if any; as before, we apply a Hadamard Transform to each qubit and the resulting state is

$$\frac{1}{2^{\frac{n+k_2}{2}}} \sum_{z} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)\cdot(e_{pf}+z)} \mid z\rangle$$

with $z$ an $n$-tuple. This can be rewritten as

$$\frac{1}{2^{\frac{n+k_2}{2}}} \sum_{z'} \sum_{u \in \mathcal{C}_1, w \in \mathcal{C}_1/\mathcal{C}_2} (-1)^{(u+w)\cdot z'} \mid z'+e_{pf}\rangle$$

with $z' = z + e_{pf}$. If $z' \in \mathcal{C}_2^{\perp}$ then $\sum_{w \in \mathcal{C}_2} (-1)^{w \cdot z'} = \mid C_2 \mid$. We can take advantage of the fact that if $\mathcal{C}_2$ is a linear code and $\mathcal{C}_2^{\perp}$ its dual then

$$\sum_{w \in \mathcal{C}_2} (-1)^{w \cdot z'} = \begin{cases} \mid \mathcal{C}_2 \mid = 2^{k_2} & \text{if } z' \in \mathcal{C}_2^{\perp} \\ 0 & \text{if } z' \notin \mathcal{C}_2^{\perp} \end{cases}$$

Indeed, if $z' \in \mathcal{C}_2^{\perp}$ then the two vectors are orthogonal, $w \cdot z' = 0$ and the sum equals $\mid \mathcal{C}_2 \mid$; the state after the Hadamard Transform can be expressed as

$$\frac{1}{2^{\frac{n-k_2}{2}}} \sum_{z' \in \mathcal{C}_2^{\perp}} (-1)^{u \cdot z'} \mid z'+e_{pf}\rangle.$$

We notice that this state resembles a state affected by a bit-flip error vector but in this case the error vector is $e_{pf}$. We use the parity check matrix $H_2$ to correct the errors described by $e_{pf}$ as before; we first compute the syndrome, then correct the errors and reach the state

$$\frac{1}{2^{\frac{n-k_2}{2}}} \sum_{z' \in \mathcal{C}_2^{\perp}} (-1)^{u \cdot z'} \mid z' \rangle$$

Finally, we apply a Hadamard Transform to each qubit and recover the original state:

$$\frac{1}{2^{\frac{k_2}{2}}} \sum_{w \in \mathcal{C}_2} \mid u + w \rangle.$$

In preparation for the discussion of stabilizer codes we present in the next section a group theoretic view of Quantum Error Correction (QEC).

## 5.11 The Pauli Group

The four Pauli operators, $\mathbb{I}, X, Z, Y$ allow us to express the four possible effects of the environment on a qubit in state $\mid \psi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$ : no error (the qubit is unchanged), bit-flip, phase-flip, and bit- and phase-flip:

| | | |
|---|---|---|
| No Error | $\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ | $\mid \varphi \rangle = \mathbb{I} \mid \psi \rangle = \alpha_0 \mid 0 \rangle + \alpha_1 \mid 1 \rangle$ |
| Bit-flip | $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ | $\mid \varphi \rangle = X \mid \psi \rangle = \alpha_1 \mid 0 \rangle + \alpha_0 \mid 1 \rangle$ |
| Phase-flip | $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ | $\mid \varphi \rangle = Z \mid \psi \rangle = \alpha_0 \mid 0 \rangle - \alpha_1 \mid 1 \rangle$ |
| Bit & phase-flip | $Y = iXZ = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ | $\mid \varphi \rangle = Y \mid \psi \rangle = -i(\alpha_1 \mid 0 \rangle - \alpha_0 \mid 1 \rangle.$ |

Pauli operators, $\mathbb{I}$, $X$, $Y$, and $Z$ form a group and have several nice properties:

1. They anticommute:

$$\begin{aligned} \{X, Z\} = 0 &\implies XZ = -ZX \\ \{X, Y\} = 0 &\implies XY = -YX \\ \{Y, Z\} = 0 &\implies YZ = -ZY. \end{aligned}$$

2. The square of each is equal to unity; this means that the transformation they perform on a qubit are reversible.

3. They span the space of $2 \times 2$ matrices; any operator describing the transformation of a single qubit is a rotation $\mathcal{R}_\theta$ with an angle $\theta$ and any rotation operator can be expressed as a linear combination of Pauli operators

$$\mathcal{R}_\theta = e^{-i\frac{\theta}{2}\bar{n}\cdot\bar{\sigma}} = \cos\frac{\theta}{2}\mathbb{I} - i\sin\frac{\theta}{2}(\bar{n}\cdot\bar{\sigma})$$

where $\bar{n} = (x, y, z)$ is a vector about which the rotation is performed and $\bar{\sigma} = (X, Y, Z)$ is a vector of Pauli operators.

Two Pauli matrices $\sigma_i$ and $\sigma_j$ are equivalent if $\sigma_j = c \times (\sigma_i)$ with the constant $c$ equal to $\pm 1$ or $\pm i$, where $i = \sqrt{-1}$:

$$\sigma_i \equiv \sigma_j \quad \Longleftrightarrow \quad \sigma_j = \pm\sigma_i \;\text{ or, }\; \sigma_j = \pm i\sigma_i.$$

Thus, we can define the set of equivalence classes of Pauli operators, $[\mathcal{P}]$. The set of Pauli operators, $\mathcal{P}$, is not an Abelian group,[38] but the set $\Pi$ of equivalence classes $[\mathcal{P}]$ of Pauli operators, also called the projective Pauli group [156], forms an Abelian group. The multiplication tables of the two groups, $\mathcal{P}$ and $[\mathcal{P}]$ are

| $\times$ | $\mathbb{I}$ | $X$ | $Y$ | $Z$ |
|---|---|---|---|---|
| $\mathbb{I}$ | $\mathbb{I}$ | $X$ | $Y$ | $Z$ |
| $X$ | $X$ | $\mathbb{I}$ | $iZ$ | $-iY$ |
| $Y$ | $Y$ | $-iZ$ | $\mathbb{I}$ | $iX$ |
| $Z$ | $Z$ | $iY$ | $-iX$ | $\mathbb{I}$ |

and

| $\times$ | $[\mathbb{I}]$ | $[X]$ | $[Y]$ | $[Z]$ |
|---|---|---|---|---|
| $[\mathbb{I}]$ | $[\mathbb{I}]$ | $[X]$ | $[Y]$ | $[Z]$ |
| $[X]$ | $[X]$ | $[\mathbb{I}]$ | $[\sigma_z]$ | $[Y]$ |
| $[Y]$ | $[Y]$ | $[Z]$ | $[\mathbb{I}]$ | $[X]$ |
| $[Z]$ | $[Z]$ | $[Y]$ | $[X]$ | $[\mathbb{I}]$ |

.

Indeed, from the multiplication table of $\mathcal{P}$ we see that

$$XZ = -iY \quad \text{and} \quad ZX = iY$$

while, from the multiplication table of $[\mathcal{P}]$ we see that

$$[X][Z] = [Z][X] = [Y].$$

The 1-qubit Pauli group $\mathcal{P}_1$ consists of the Pauli operators, $\mathbb{I}$, $X$, $Y$, and $Z$ together with the multiplicative factors $\pm 1$ and $\pm i$

$$\mathcal{P}_1 \equiv \{\pm\mathbb{I}, \pm i\mathbb{I}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}.$$

The cardinality of $\mathcal{P}_1$ is $\mid \mathcal{P}_1 \mid = 16 = 2^4$.

**Proposition.** *The members of the* 1*-qubit Pauli group* $\mathcal{P}_1$

1. *are unitary;*
2. *either commute or anticommute;*
3. *are either Hermitian, or anti-Hermitian.*

*Proof:* Clearly all members are unitary because the four Pauli operators are multiplied with $\pm 1$ or $\pm i$. $X, Y, Z$ anticommute but $\mathbb{I}$ commutes with each of them. $\mathbb{I}, X, Y$ and $Z$ are Hermitian, but

$$(i\mathbb{I})^\dagger = -i\mathbb{I}.$$

[38] A multiplicative group $G$ is Abelian iff $\forall a, b \in G \;\; a \times b = b \times a$.

We introduce now the generator of a group, a concept that allows us to describe succinctly the group $G$; instead of listing a possibly very large number of elements of a group we only specify the much smaller number of generators of the group.

A set of $\mu$ elements $\langle g_1, g_2, \ldots, g_\mu \rangle$ is called the generator of the group $G$ if the element $g_i$ can be written as a product of (possibly repeated) elements from the list $\langle g_1, g_2, \ldots, g_\mu \rangle$, $\forall g_i \in G$; we write $G = \langle g_1, g_2, \ldots, g_\mu \rangle$.

**Proposition.** *The number $\mu$ of generators of the group $G$ satisfies the inequality*

$$\mu \leq \log_2(\mid G \mid).$$

Consider two elements $f, g \in G$ such that $f \in \langle g_1, g_2, \ldots, g_\mu \rangle$ and $g \notin \langle g_1, g_2, \ldots, g_\mu \rangle$; then $fg \notin \langle g_1, g_2, \ldots, g_\mu \rangle$. If we assume

$$fg \in \langle g_1, g_2, \ldots, g_\mu \rangle \text{ then } f^{-1}fg \in \langle g_1, g_2, \ldots, g_\mu \rangle \quad \Rightarrow \quad g \in \langle g_1, g_2, \ldots, g_\mu \rangle$$

which contradicts our assumption $g \notin \langle g_1, g_2, \ldots, g_\mu \rangle$; thus, $\forall f \in \langle g_1, g_2, \ldots, g_\mu \rangle$ we have

$$fg \in \langle g_1, g_2, \ldots, g_\mu, g \rangle \quad \text{and} \quad fg \notin \langle g_1, g_2, \ldots, g_\mu \rangle.$$

It follows that by adding the element $g$ to the set of generators we double the number of elements of the group described by the generator; we conclude that the cardinality of the set of generators is at most the logarithm base 2 of the cardinality of the group.

**Example.** *The generators of $\mathcal{P}_1$ are*

$$\mathcal{P}_1 = \langle X, Z, i\mathbb{I} \rangle.$$

Indeed, $i^2 = -1$, $i^3 = -i$, and $i^4 = 1$; every element of $\mathcal{P}_1$ can be expressed as a product of a finite number of generators:

$$+\mathbb{I} = i\mathbb{I}i\mathbb{I}i\mathbb{I}i\mathbb{I}, \qquad -\mathbb{I} = i\mathbb{I}i\mathbb{I}, \qquad -i\mathbb{I} = i\mathbb{I}i\mathbb{I}i\mathbb{I};$$

while,

$$-X = i\mathbb{I}i\mathbb{I}X, \qquad +iX = i\mathbb{I}X, \qquad -iX = i\mathbb{I}i\mathbb{I}i\mathbb{I}X,$$

and

$$+Y = i\mathbb{I}XZ, \qquad -Y = i\mathbb{I}i\mathbb{I}i\mathbb{I}XZ,$$

$$+iY = i\mathbb{I}i\mathbb{I}XZ, \quad -iY = i\mathbb{I}i\mathbb{I}XZ,$$

and, finally,

$$-Z = i\mathbb{I}i\mathbb{I}Z, \qquad +iZ = i\mathbb{I}Z, \qquad -iZ = i\mathbb{I}i\mathbb{I}i\mathbb{I}Z.$$

We conclude that the three generators $\langle X, Z, i\mathbb{I} \rangle$ allow us to express any of the $16 = 2^4$ operators in $\mathcal{P}_1$. The cardinality of the set of generators satisfies the condition

$$| \langle X, Z, i\mathbb{I} \rangle | \leq \log_2 | \mathcal{P}_1 |, \qquad \text{or} \qquad 3 < 4.$$

The $n$-qubit Pauli group $\mathcal{P}_n$ consists of the $4^n$ tensor products of $\mathbb{I}$, $X$, $Y$, and $Z$ and an overall phase of $\pm 1$ or $\pm i$; the group has $4^{n+1}$ elements.

One element of the group is an $n$-tuple, the tensor product of $n$ one-qubit Pauli operators and can be used to describe the error operator applied to an $n$-qubit register; the weight of such an operator in $\mathcal{P}_n$ is equal to the number of tensor factors which are not equal to $\mathbb{I}$.

**Example.** *Let $n = 5$; the 5-qubit Pauli group consists of the tensor products of the form:*

$$\mathbf{E}^{(1)} \otimes \mathbf{E}^{(2)} \otimes \mathbf{E}^{(3)} \otimes \mathbf{E}^{(4)} \otimes \mathbf{E}^{(5)} \quad \text{with} \quad \mathbf{E}^{(i)} \in \mathcal{P}_1, \quad 1 \leq i \leq 5.$$

The operator $\mathbf{E}_\alpha = \mathbb{I} \otimes X \otimes \mathbb{I} \otimes Z \otimes \mathbb{I}$ has a weight of 2 and represents: no error for qubits 1, 3, and 5, a bit-flip of qubit 2 and a phase-flip of qubit 4; the qubits are numbered from left to right.

## 5.12 Stabilizer Codes

Classical codes use simple tests to determine if an $n$-tuple is a codeword:

- If H is the parity check matrix of the linear code $\mathcal{C}$, $v \in \mathcal{C}$ is a codeword, and $w = v + e$ is an $n$-tuple in error then $\mathrm{H}v^T = 0$ and $\Sigma = \mathrm{H}w^T = \mathrm{H}e^T \neq 0$; we use the syndrome $\Sigma$ to locate the error.
- A cyclic code $\mathcal{C}$ is an ideal of polynomials with the generator polynomial $g(x)$ and we conclude that the received polynomial is a codeword, $r(x) \in \mathcal{C}$, if $g(x)$ divides $r(x)$; otherwise we use the remainder of the division of $r(x)$ by $g(x)$ to locate the error(s).

We wish to construct some tests for quantum codes that resemble the effect of the parity check matrix or of the generator polynomials and allow us to detect if errors have occurred, and if so, to correct them. Assume the $m$ codewords of a code are described by the vectors $| \psi_i \rangle$ and that we have identified a group of $q$ operators $\mathbf{M}_j$ that allow us to detect the errors that may affect any codeword. In this context the word "detect" means that we carry out a measurement which does not reveal any information about the actual state $| \psi_i \rangle$, but indicates if the codeword is affected by error(s). This line of thought leads us to an alternative way of specifying a quantum code as *the set of quantum operators that correct all possible errors affecting all codewords.*

The stabilizer formalism is a succinct manner of describing a quantum error correcting code by a set of quantum operators. In this section we first describe the basic idea and illustrate its application to the nine qubit code of Shor, then we discuss the general stabilizer formalism.

**Basic concepts.** Consider a set of $n$-qubit registers, in state $| \psi_i \rangle$, $1 \leq i \leq m$, each one representing a word of a quantum code $\mathcal{Q}$. Informally, the stabilizer $\mathcal{S}$ of $\mathcal{Q}$ is a subgroup of the n-qubit Pauli group,

$$\mathcal{S} \subset \mathcal{P}_n.$$

The generators of the subgroup $\mathcal{S}$ are

$$M = \{\mathbf{M}_1, \mathbf{M}_2 \ldots \mathbf{M}_q\}.$$

The eigenvectors of the generators $\{\mathbf{M}_1, \mathbf{M}_2 \ldots \mathbf{M}_q\}$ have special properties: those corresponding to eigenvalues of $+1$ are the codewords of $\mathcal{Q}$ and those corresponding to eigenvalues of $-1$ are codewords affected by errors. If $\mathbf{M}_j \in M$ and if

$$\mathbf{M}_j \mid \psi_i\rangle = (+1) \mid \psi_i\rangle$$

then $\mid \psi_i\rangle$ is a codeword, $\mid \psi_i\rangle \in \mathcal{Q}$. This justifies the name "stabilizer" given to the set $\mathcal{S}$, any operator in $\mathcal{S}$ leaves the state of a word unchanged. On the other hand if

$$\mathbf{M}_j \mid \varphi_k\rangle = (-1) \mid \varphi_k\rangle.$$

then $\mid \varphi_k\rangle = \mathbf{E}_\alpha \mid \psi_k\rangle$, the state $\mid \varphi_k\rangle$ is the result of an error affecting the codeword $\mid \psi_k\rangle \in \mathcal{Q}$.

The errors $E = \{\mathbf{E}_1, \mathbf{E}_2 \ldots\}$ affecting a codeword in $\mathcal{Q}$, are also a subgroup of the n-qubit Pauli group $E \subset \mathcal{P}_n$; each error operator $\mathbf{E}_i$ is a tensor product of $n$ Pauli matrices. The weight of an error operator is equal to the number of errors affecting a quantum word, thus the number of Pauli operators other than $\mathbb{I}$ in this $n$-dimensional tensor product.

We show that the error operators anticommute with the generators of the stabilizer group, $\mathcal{S}$. Therefore, to detect errors we have to compute the eigenvectors of the generators and identify those with an eigenvalue of $-1$.

**Example.** *A pair of entangled qubits.* Their state is

$$\mid \beta_{00}\rangle = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}} \quad \text{or} \quad \mid \beta_{00}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

The following transformations are applied to the qubits

$$\mathbf{M}_x = X \otimes X = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{M}_z = Z \otimes Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We see that

$$\mathbf{M}_x \mid \beta_{00}\rangle = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus,

$$\mathbf{M}_x \mid \beta_{00}\rangle = \mid \beta_{00}\rangle \quad \text{or} \quad \mathbf{M}_x \mid \beta_{00}\rangle = (+1) \mid \beta_{00}\rangle.$$

Similarly,

$$\mathbf{M}_z \mid \beta_{00}\rangle = \mid \beta_{00}\rangle \quad \text{or} \quad \mathbf{M}_z \mid \beta_{00}\rangle = (+1) \mid \beta_{00}\rangle.$$

These two equalities show that $\mid \beta_{00}\rangle$ is an eigenvector of the operators $\mathbf{M}_x$ and $\mathbf{M}_z$, respectively, and that the corresponding eigenvalues are equal to $+1$. The generators commute, thus $\mathcal{S}$ is an Abelian group.

**Example.** *The nine-qubit error correcting code of Shor.* The logical qubits are

$$\mid 0\rangle_L = \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle+|111\rangle}{\sqrt{2}}$$

$$\mid 1\rangle_L = \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}} \otimes \frac{|000\rangle-|111\rangle}{\sqrt{2}}$$

Consider the following set of operators acting on nine-qubit codewords:

| Operator/Qubit | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{M}_1$ | $Z$ | $Z$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ |
| $\mathbf{M}_2$ | $Z$ | $\mathbb{I}$ | $Z$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ |
| $\mathbf{M}_3$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $Z$ | $Z$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ |
| $\mathbf{M}_4$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $Z$ | $\mathbb{I}$ | $Z$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ |
| $\mathbf{M}_5$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $Z$ | $Z$ | $\mathbb{I}$ |
| $\mathbf{M}_6$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $Z$ | $\mathbb{I}$ | $Z$ |
| $\mathbf{M}_7$ | $X$ | $X$ | $X$ | $X$ | $X$ | $X$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ |
| $\mathbf{M}_8$ | $X$ | $X$ | $X$ | $\mathbb{I}$ | $\mathbb{I}$ | $\mathbb{I}$ | $X$ | $X$ | $X$ |

The eight operators $\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_8$ are the generators of the group $\mathcal{P}$ of operators that stabilize the nine qubit error correcting code. The logical qubits $\mid 0\rangle_L$ and $\mid 1\rangle_L$ are eigenvectors of the eight operators and the corresponding eigenvalues are equal to $+1$

$$\mathbf{M}_j \mid 0\rangle_L = (+1) \mid 0\rangle_L \quad \text{and} \quad \mathbf{M}_j \mid 1\rangle_L = (+1) \mid 1\rangle_L, \quad 1 \leq j \leq 8.$$

For example, $\mid 0\rangle_L$ is an eigenvector of $\mathbf{M}_7$ corresponding to an eigenvalue $\lambda = +1$,

$$\mathbf{M}_7 \mid 0\rangle_L = (+1) \mid 0\rangle_L,$$

where,

$$\begin{aligned} \mid 0\rangle_L = & \quad \mid 000000000\rangle + \mid 000000111\rangle + \mid 000111000\rangle + \mid 000111111\rangle \\ + & \quad \mid 111000000\rangle + \mid 111000111\rangle + \mid 111111000\rangle + \mid 111111111\rangle. \end{aligned}$$

The tensor product of two operators $\sigma_a^{(i)}$ and $\sigma_b^{(j)}$ acting on qubits $i$ and $j$ of an $n$ qubit word is $\sigma_a^{(i)} \otimes \sigma_b^{(j)}$. According to its definition,

$$\mathbf{M}_7 = X^{(1)} \otimes X^{(2)} \otimes X^{(3)} \otimes X^{(4)} \otimes X^{(5)} \otimes X^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)},$$

the operator $\mathbf{M}_7$ flips the first six qubits of a nine-qubit word and leaves the remanding three unchanged, thus,

$$\begin{aligned}\mathbf{M}_7 \mid 0\rangle_L = & \quad \mid 111111000\rangle + \mid 111111111\rangle + \mid 111000000\rangle + \mid 111000111\rangle \\ + & \quad \mid 000111000\rangle + \mid 000111111\rangle + \mid 000000000\rangle + \mid 000000111\rangle.\end{aligned}$$

Let $\mathbf{E}_x^{(1)}$ and $\mathbf{E}_x^{(2)}$ be the two operators describing a single bit-flip error of the first and the second qubit of a nine-qubit codeword, respectively,

$$\begin{aligned}\mathbf{E}_x^{(1)} = & \quad X^{(1)} \otimes \mathbb{I}^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)} \\ \mathbf{E}_x^{(2)} = & \quad \mathbb{I}^{(1)} \otimes X^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)}.\end{aligned}$$

Both operators anticommute with $\mathbf{M}_1$

$$\mathbf{M}_1 = Z^{(1)} \otimes Z^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)}.$$

For example,

$$\begin{aligned}\mathbf{M}_1\mathbf{E}_x^{(1)} = & \quad (Z^{(1)} \otimes Z^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)}) \\ & (X^{(1)} \otimes \mathbb{I}^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)}) \\ = & \quad (Z^{(1)}X^{(1)}) \otimes (Z^{(2)}\mathbb{I}^{(2)}) \otimes (\mathbb{I}^{(3)}\mathbb{I}^{(3)}) \otimes (\mathbb{I}^{(4)}\mathbb{I}^{(4)}) \\ & \otimes(\mathbb{I}^{(5)}\mathbb{I}^{(5)}) \otimes (\mathbb{I}^{(6)}\mathbb{I}^{(6)}) \otimes (\mathbb{I}^{(7)}\mathbb{I}^{(7)}) \otimes (\mathbb{I}^{(8)}\mathbb{I}^{(8)}) \otimes (\mathbb{I}^{(9)}\mathbb{I}^{(9)}) \\ = & \quad -(X^{(1)}Z^{(1)}) \otimes Z^{(2)} \otimes \mathbb{I}^{(3)} \otimes \mathbb{I}^{(4)} \otimes \mathbb{I}^{(5)} \otimes \mathbb{I}^{(6)} \otimes \mathbb{I}^{(7)} \otimes \mathbb{I}^{(8)} \otimes \mathbb{I}^{(9)} \\ = & \quad -\mathbf{E}_x^{(1)}\mathbf{M}_1.\end{aligned}$$

From our previous derivation it follows that

$$\mathbf{M}_1\mathbf{E}_x^{(1)} \mid 0\rangle_L = -\mathbf{E}_x^{(1)}\mathbf{M}_1 \mid 0\rangle_L = -\mathbf{E}_x^{(1)}[\mathbf{M}_1 \mid 0\rangle_L] = (-1)\mathbf{E}_x^{(1)} \mid 0\rangle_L.$$

This means that $\mathbf{E}_x^{(1)} \mid 0\rangle_L$ is an eigenvector of the stabilizer $\mathbf{M}_1$ with an eigenvalue of $\lambda = -1$. Similarly, $\mathbf{E}_x^{(1)} \mid 1\rangle_L$, $\mathbf{E}_x^{(2)} \mid 0\rangle_L$, and $\mathbf{E}_x^{(2)} \mid 1\rangle_L$ are eigenvectors of the stabilizer $\mathbf{M}_1$ with eigenvalues equal to $-1$. Thus, $\mathbf{M}_1$ can be used to detect a single bit-flip error of qubits 1 and 2 of the a nine qubit word of Shor's code. The operators corresponding to a single bit-flip error of qubits 3 to 9 commute with $\mathbf{M}_1$ thus, $\mathbf{M}_1$ cannot be used to detect a single bit-flip error of qubits 3 to 9. Indeed,

$$\mathbf{M}_1\mathbf{E}_x^{(i)} \mid 0\rangle_L = \mathbf{E}_x^{(i)}\mathbf{M}_1 \mid 0\rangle_L \qquad \text{and} \qquad \mathbf{M}_1\mathbf{E}_x^{(i)} \mid 1\rangle_L = \mathbf{E}_x^{(i)}\mathbf{M}_1 \mid 1\rangle_L \qquad 3 \leq i \leq 9.$$

It is left as an exercise to show that any error operator $\mathbf{E}^i_{x|y|z}$, $1 \leq i \leq 9$ anticommutes with at least one of the eight operators $\mathbf{M}_j$, $1 \leq j \leq 8$; thus, any single qubit error can be detected.

This example shows that it is sufficient only to measure the eigenvalues of the stabilizers $\mathbf{M}_i$ to detect errors; an eigenvalue of $-1$ shows that the corresponding codeword was affected by the error $E$. We can detect the occurrence of an error if the corresponding error operator anticommutes with the generators of a code regarded as a subgroup.

**The general stabilizer formalism.** The stabilizer formalism provides an alternative description of an $[n, k, d]$ quantum code; we use the following notations:

- $\mid \psi_i\rangle$ is a codeword,
- $\mathbf{M}$ is a member of the group of operators that stabilize codewords, $\mathbf{M} \mid \psi_i\rangle = (+1) \mid \psi_i\rangle$,
- $\mathbf{E}$ is an error operator. Sometimes we denote the error operator by $\mathbf{E}^k_{x|y|z}$ to specify a bit-flip error, $X$, a phase-flip error, $Z$, or a bit and phase-flip error, $Y$, of qubit k.

We show that if the two operators, $\mathbf{M}$ and $\mathbf{E}$ anticommute, $\{\mathbf{M}, \mathbf{E}\} = 0$, then

$$\mathbf{M}(\mathbf{E} \mid \psi_i\rangle) = (-1)\ (\mathbf{E} \mid \psi_i\rangle).$$

This means that $\mathbf{E} \mid \psi_i\rangle$ is an eigenvector of $\mathbf{M}$ corresponding to the eigenvalue $-1$. We detect the error $\mathbf{E}$ by measuring $\mathbf{M}$ on the original codeword $\mid \psi_i\rangle$ after error $\mathbf{E}$ forced it to state $\mathbf{E} \mid \psi_i\rangle$.

We review first a few properties of the Pauli group. All elements of $\mathcal{P}$ are unitary. $\mathbb{I}, X, Y$ and $Z$ are Hermitian, but $(i\mathbb{I})^\dagger = -i\mathbb{I}$, thus all elements of $\mathcal{P}$ are either Hermitian or anti-Hermitian. All elements of $\mathcal{P}$ either commute or anticommute. Recall that the *commutator of two operators* is

$$[\mathbf{M}_\alpha, \mathbf{M}_\beta] = \mathbf{M}_\alpha \mathbf{M}_\beta - \mathbf{M}_\beta \mathbf{M}_\alpha.$$

If the two operators commute, $\mathbf{M}_\alpha \mathbf{M}_\beta = \mathbf{M}_\beta \mathbf{M}_\alpha$ then $[\mathbf{M}_\alpha, \mathbf{M}_\beta] = 0$. The *anti-commutator of two operators* is

$$\{\mathbf{M}_\alpha, \mathbf{M}_\beta\} = \mathbf{M}_\alpha \mathbf{M}_\beta + \mathbf{M}_\beta \mathbf{M}_\alpha.$$

If the two operators anticommute, $\mathbf{M}_\alpha \mathbf{M}_\beta = -\mathbf{M}_\beta \mathbf{M}_\alpha$ then $\{\mathbf{M}_\alpha, \mathbf{M}_\beta\} = 0$.

We continue this discussion with a few definitions.

The stabilizer $\mathcal{S}$ of an $[n, k, d]$ code is an Abelian subgroup of the Pauli group $\mathcal{P}$. The generators of $\mathcal{S}$ are: $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_{n-k}$. The weight of $\mathbf{M}_i \in \mathcal{S}$ is the number of elements of $\mathbf{M}_i$ different from $\mathbb{I}$.

A nondegenerate stabilizer code $[n, k, d]$ is one where none of the elements of the stabilizer $\mathcal{S}$ (except the identity) has weight less than $d$; in case of a degenerate stabilizer code at least one element of the stabilizer $\mathcal{S}$ other than the identity has a weight less than $d$.

**Example.** *Shor's code.* The weight of $\mathbf{M}_1$, the first generator for Shor's code, is equal to 2, thus, the nine-qubit code is degenerate.

The coding space

$$T = \{\mid \psi\rangle \ \text{ such that } \ \mathbf{M} \mid \psi\rangle = (+1) \mid \psi\rangle \qquad \mathbf{M} \in \mathcal{S}\}$$

is the space of all vectors $\mid \psi\rangle$ fixed by $\mathcal{S}$.

Recall that *two commuting matrices can be simultaneously diagonalized, thus we can measure the eigenvalue of one of them without disturbing the eigenvector of the other. Thus, $\mathcal{S}$ must be Abelian* because only commuting operators may have simultaneous eigenvectors and the code consist of eigenvectors of the generators of $\mathcal{S}$. Due to the fact that $\mathcal{S}$ is Abelian and because neither $-1$ nor $+i$ is in $\mathcal{S}$ it follows that the coding space $T$ has dimension $2^k$. Indeed, an $(n, k, d)$ code should be able to encode $2^k$ messages thus must have $2^k$ codewords.

**Proposition.** *$\mathcal{S}$ is a stabilizer of a non-trivial Hilbert subspace $V_{2^n} \subset \mathcal{H}_{2^n}$ if and only if:*

1. $\mathcal{S} = \{\mathbf{S}_1, \mathbf{S}_2, \ldots\}$ is an Abelian group:

$$\mathbf{S}_i\mathbf{S}_j = \mathbf{S}_j\mathbf{S}_i, \quad \forall \mathbf{S}_i, \mathbf{S}_j \in S \quad i \neq j.$$

2. The identity matrix multiplied by $-1$ is not in $\mathcal{S}$:

$$-\mathbb{I}^{\otimes n} \notin \mathcal{S}.$$

*Proof:* We show only that the two conditions are necessary; first, we note that $\mathbf{S}_i, \mathbf{S}_j \in S$ are tensor products of Pauli operators thus, they either commute or anticommute. Then

$$\mathbf{S}_i\mathbf{S}_j \mid \psi\rangle = \mid \psi\rangle, \quad \forall \mid \psi\rangle \in V_{2^n}.$$

Indeed, $\mathbf{S}_i$ and $\mathbf{S}_j$ are both stabilizers, $\mathbf{S}_i \mid \psi\rangle = \mid \psi\rangle$ and $\mathbf{S}_j \mid \psi\rangle = \mid \psi\rangle$ thus,

$$\mathbf{S}_i\mathbf{S}_j \mid \psi\rangle = \mathbf{S}_i(\mathbf{S}_j \mid \psi\rangle) = \mathbf{S}_i \mid \psi\rangle = \mid \psi\rangle.$$

If $\mathbf{S}_i$ and $\mathbf{S}_j$ anticommute then

$$-\mathbf{S}_j\mathbf{S}_i \mid \psi\rangle = - \mid \psi\rangle, \quad \forall \mid \psi\rangle \in V_{2^n}.$$

We see that

$$-\mathbf{S}_j\mathbf{S}_i \mid \psi\rangle = -\mathbf{S}_j\left(\mathbf{S}_i \mid \psi\rangle\right) = -\mathbf{S}_j \mid \psi\rangle = - \mid \psi\rangle.$$

We assumed that the two operators anticommute, $\mathbf{S}_i\mathbf{S}_j = -\mathbf{S}_j\mathbf{S}_i$ thus

$$\mathbf{S}_i\mathbf{S}_j \mid \psi\rangle = -\mathbf{S}_j\mathbf{S}_i \mid \psi\rangle \quad \Longrightarrow \quad \mid \psi\rangle = - \mid \psi\rangle \quad \forall \mid \psi\rangle \in V_{2^n}$$

which is a contradiction, since we assumed that $V_{2^n}$ a non-trivial Hilbert subspace. Therefore we conclude that all members of the stabilizer group commute and $\mathcal{S}$ is an Abelian group.

Now, assume that $-\mathbb{I}^{\otimes n} \in \mathcal{S}$; then

$$-\mathbb{I}^{\otimes n} \mid \psi\rangle = \mid \psi\rangle \quad \text{or} \quad \mathbb{I}^{\otimes n} \mid \psi\rangle = - \mid \psi\rangle, \quad \forall \mid \psi\rangle \in V_{2^n}.$$

This leads to a contradiction because $\mathbb{I}^{\otimes n} \mid \psi\rangle = \mid \psi\rangle$; we conclude that $-\mathbb{I}^{\otimes n} \notin S$. $\square$

**Proposition.** *Let $\mathbf{E}$ be an error operator. If $\mathbf{M} \in \mathcal{S}$ and $\mathbf{M}$ and $\mathbf{E}$ anticommute then,*

$$\mathbf{ME} \mid \psi_i\rangle = -\mathbf{E} \mid \psi_i\rangle \quad \forall \mid \psi_i\rangle \in T.$$

We can detect the error $\mathbf{E}$ by measuring $\mathbf{M}$; indeed, $\mathbf{E} \mid \psi_i\rangle$ is an eigenvector of the generator $\mathbf{M}$ corresponding to the eigenvalue $-1$.

*Proof:* We know that $\mathbf{M}$ and $\mathbf{E}$ anticommute, $\mathbf{ME} = -\mathbf{EM}$, and that $\mathbf{M}$ is a stabilizer of the code, $\mathbf{M} \mid \psi_i\rangle = \mid \psi_i\rangle \;\; \forall \mid \psi_i\rangle \in T$. Thus:

$$\mathbf{M}(\mathbf{E} \mid \psi_i\rangle) = \mathbf{ME} \mid \psi_i\rangle = -\mathbf{EM} \mid \psi_i\rangle = -\mathbf{E}(\mathbf{M} \mid \psi_i\rangle) = (-1)\ \mathbf{E} \mid \psi_i\rangle.$$

We also observe that given any codewords, $\forall \mid \psi_i\rangle, \mid \psi_j\rangle \in T$:

$$\langle\psi_i \mid \mathbf{E} \mid \psi_j\rangle = \langle\psi_i \mid \mathbf{EM} \mid \psi_j\rangle = -\langle\psi_i \mid \mathbf{ME} \mid \psi_j\rangle = -\langle\psi_i \mid \mathbf{E} \mid \psi_j\rangle = 0.$$

Thus, the code satisfies the necessary condition for the existence of a quantum code discussed in Section 5.1.

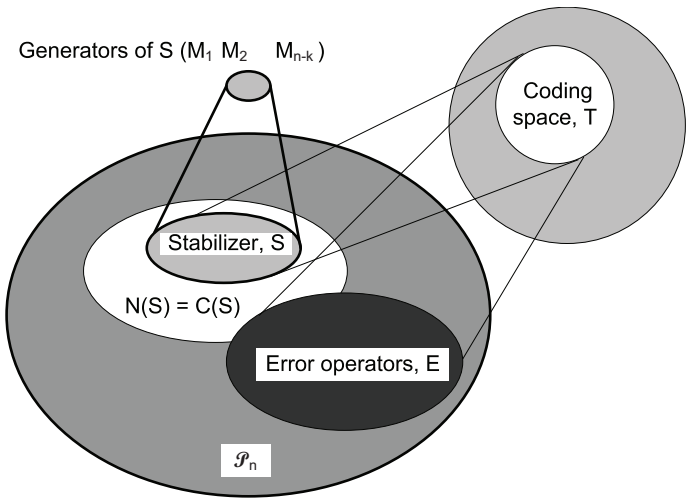



Figure 129: A quantum code with stabilizer $\mathcal{S}$ and normalizer $N(S)$ detects all errors $\mathbf{E}$ that are either in $\mathcal{S}$, or anticommute with some element of $\mathcal{S}$. The stabilizer $\mathcal{S}$ of an $[n, k, d]$ code has $2^{n-k}$ elements; the coding space, $T$, has $2^k$ elements and is a subspace of the larger $2^n$ space. The code has distance $d$ if and only if $N(S) - S$ contains no elements with weight less than $d$. $\mathbf{M}_1, \mathbf{M}_2, \ldots \mathbf{M}_{n-k}$ are the $n-k$ generators of $\mathcal{S}$. $N(S) = C(S)$ and $E \notin S$. The group $N(S)$ has $4 \cdot 2^{n-k}$ elements. If $\mathbf{E} \notin N(S) - S$, then $\mathbf{E}$ rearranges elements of $T$, it does not take codewords outside $T$. $\mathcal{P}_n$ is the $n$-qubit Pauli group.

**Proposition.** *Given an $[n, k, d]$ stabilizer code (with $n$ the length of a codeword, $k$ the number of information symbols, and $d$ the distance of the code) the cardinality of the stabilizer $\mathcal{S}$ and the cardinality of its generator $M$ are respectively:*

$$\mid \mathcal{S} \mid = 2^{n-k}, \qquad \mid M \mid = n - k.$$

The <u>normalizer</u>, $N(\mathcal{S})$, $\mathcal{S} \in \mathcal{P}$, is the set of elements in $\mathcal{P}$ that fix the stabilizer $\mathcal{S}$ under conjugation. It is left as an exercise to show that $\mathcal{S}$ is a normal subgroup of $N(\mathcal{S})$:

$$\mathcal{S} \subset N(\mathcal{S}).$$

Since the elements of the normalizer $N(\mathcal{S})$ move codewords around in the coding space $T$, they are in fact the *encoding operators*. Only the elements $\mathbf{A} \in N(\mathcal{S})/\mathcal{S}$ act on the codewords of $T$ nontrivially; if $\mathbf{B} \in \mathcal{S}$ and $\mid \psi_i\rangle \in T$ then $\mathbf{B} \mid \psi_i\rangle = \mid \psi_i\rangle$.

The centralizer, $C(\mathcal{S})$, $\mathcal{S} \in \mathcal{P}$, is the set of elements in $\mathcal{P}$ that commute with all the elements of the stabilizer $\mathcal{S}$. Figure 129 shows the relationship between groups of operators, errors operators, and the coding space for stabilizer codes.

**Proposition.** *The normalizer is equal to the centralizer $N(\mathcal{S}) = C(\mathcal{S})$ and they contain $4 \cdot 2^{n+k}$ elements.*

$\mathcal{S}$ is an Abelian subgroup of $\mathcal{P}$ thus, $\forall\ \mathbf{A} \in \mathcal{P}$ and $\forall\ \mathbf{M} \in \mathcal{S}$, $\mathbf{A}$ and $\mathbf{M}$ either commute or anticommute, $\mathbf{MA} = \pm\mathbf{AM}$; moreover, $\mathbf{M}$ is an element of the stabilizer, $\mathbf{M} \in \mathcal{S}$, and this implies that $\mathbf{AM} = \mathbf{M}$ thus,

$$\mathbf{A}^\dagger\mathbf{MA} = \pm\mathbf{A}^\dagger\mathbf{AM} = \pm\mathbf{M}.$$

But, $-1 \notin \mathcal{S}$ thus, $\mathbf{A} \in N(\mathcal{S})$ if and only if $\mathbf{A} \in C(\mathcal{S})$; we conclude that $N(\mathcal{S}) = C(\mathcal{S})$. It is left as an exercise to prove that $\mid N(\mathcal{S}) \mid = 4 \cdot 2^{n+k}$ using the fact that the Pauli group $\mathcal{P}_n$ has $4^{n+1}$ elements including an overall phase of $\pm 1$ and $\pm i$.

**Proposition.** *A code with stabilizer $\mathcal{S}$ and normalizer $N(\mathcal{S})$ will have distance $d$ if and only if $N(\mathcal{S}) - \mathcal{S}$ contains no elements with weight less than $d$.*

The error syndrome corresponding to stabilizer $\mathbf{M}$ is a function of $\mathbf{E}$, the error operator, $f_{\mathbf{M}}(\mathbf{E}) : \mathcal{P} \mapsto \mathbb{Z}_2$ defined as

$$f_{\mathbf{M}}(\mathbf{E}) = \begin{cases} 0 & \text{if } [\mathbf{M}, \mathbf{E}] = 0 \\ 1 & \text{if } \{\mathbf{M}, \mathbf{E}\} = 0. \end{cases}$$

Let $f(\mathbf{E})$ be the $(n-k)$-bit integer given by the binary vector

$$f(\mathbf{E}) = (f_{\mathbf{M}_1}(\mathbf{E}), f_{\mathbf{M}_2}(\mathbf{E}) \ldots f_{\mathbf{M}_{n-k}}(\mathbf{E})).$$

If $\mathbf{E} \in N(\mathcal{S})$ then $f(\mathbf{E}) = 0$.

To "detect" an error for a stabilizer code we have to measure the eigenvalues of each generator of the stabilizer of the code. The eigenvalue of $\mathbf{M}_i$ is equal to $(-1)^{f_{\mathbf{M}_i}(\mathbf{E})}$ and gives us the error syndrome. For a nondegenerate code the error syndrome reveals position and type of error and for a degenerate code the error syndrome reveals the set of degenerate errors.

**Proposition.** *The error syndrome uniquely identifies the qubit(s) in error if and only if the subsets of the stabilizer group which anti-commute with the error operators are distinct.*

An error can be identified and corrected only if it can be distinguished from any other error in the error set. Let $Q(\mathcal{S})$ be the stabilizer code with stabilizer $\mathcal{S}$. The *Correctable Set of Errors* for $Q(\mathcal{S})$ includes all errors which can be detected by $\mathcal{S}$ and have distinct error syndromes.

**Corollary.** *Given a quantum error correcting code $Q$ capable to correct $e_u$ errors, the syndrome does not allow us to distinguish the case when more than $e_u$ qubits are in error. When we have exact prior knowledge about $e_c$ correlated errors the code is capable of correcting these $e_u + e_c$ errors.*

*Proof:* Assume that $F_1$, $F_2$ cause at most $e_u$ qubits to be in error, thus $F_1$, $F_2$ are included in the *Correctable Set of Errors* of $Q$. Since errors $F_1$ and $F_2$ are distinguishable, there must exist some operator $M \in S$ which commutes with one of them, and anti-commutes with the other:

$$F_1^T F_2 M = -M F_1^T F_2$$

If we know the exact correlated errors $E$ in the system, then:
$(E^T F_1)^T (E^T F_2) M = (F_1^T E E^T F_2) M = F_1^T F_2 M = -M F_1^T F_2 = -M (E^T F_1)^T (E^T F_2)$
which means that the stabilizer $M$ commutes with one of the two errors $E^T F_1$, $E^T F_2$ and anti-commutes with the other. So error $E^T F_1$ is distinguishable from error $E^T F_2$. Therefore, if we know the exact prior errors $E$, we can identify and correct any $E^T F_i$ errors with the weight of $F_i$ equal to or less than $e_u$.

Table 23: Single error operators for the Steane seven qubit code and the generator(s) which anti-commute with each operator.

| Error | Generator(s) | Error | Generator(s) | Error | Generator(s) |
|---|---|---|---|---|---|
| $\mathbf{X}_1$ | $\mathbf{M}_6$ | $\mathbf{Z}_1$ | $\mathbf{M}_3$ | $\mathbf{Y}_1$ | $\mathbf{M}_3,\ \mathbf{M}_6$ |
| $\mathbf{X}_2$ | $\mathbf{M}_5$ | $\mathbf{Z}_2$ | $\mathbf{M}_2$ | $\mathbf{Y}_2$ | $\mathbf{M}_2,\ \mathbf{M}_5$ |
| $\mathbf{X}_3$ | $\mathbf{M}_5,\ \mathbf{M}_6$ | $\mathbf{Z}_3$ | $\mathbf{M}_2,\ \mathbf{M}_3$ | $\mathbf{Y}_3$ | $\mathbf{M}_2,\ \mathbf{M}_3,\ \mathbf{M}_5,\ \mathbf{M}_6$ |
| $\mathbf{X}_4$ | $\mathbf{M}_4$ | $\mathbf{Z}_4$ | $\mathbf{M}_1$ | $\mathbf{Y}_4$ | $\mathbf{M}_1,\ \mathbf{M}_4$ |
| $\mathbf{X}_5$ | $\mathbf{M}_4,\ \mathbf{M}_6$ | $\mathbf{Z}_5$ | $\mathbf{M}_1,\ \mathbf{M}_3$ | $\mathbf{Y}_5$ | $\mathbf{M}_1,\ \mathbf{M}_3,\ \mathbf{M}_5,\ \mathbf{M}_6$ |
| $\mathbf{X}_6$ | $\mathbf{M}_4,\ \mathbf{M}_5$ | $\mathbf{Z}_6$ | $\mathbf{M}_1,\ \mathbf{M}_2$ | $\mathbf{Y}_6$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_4,\ \mathbf{M}_5$ |
| $\mathbf{X}_7$ | $\mathbf{M}_4,\ \mathbf{M}_5,\ \mathbf{M}_6$ | $\mathbf{Z}_7$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_3$ | $\mathbf{Y}_7$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_3,\ \mathbf{M}_4,\ \mathbf{M}_5,\ \mathbf{M}_6$ |

Table 24: The single errors and the corresponding syndromes for the Steane seven qubit code.

| Error | Syndrome $a_{1-6}$ | Error | Syndrome $a_{1-6}$ | Error | Syndrome $a_{1-6}$ |
|---|---|---|---|---|---|
| $X_1$ | 000001 | $Z_1$ | 001000 | $Y_1$ | 001001 |
| $X_2$ | 000010 | $Z_2$ | 010000 | $Y_2$ | 010010 |
| $X_3$ | 000011 | $Z_3$ | 011000 | $Y_3$ | 011011 |
| $X_4$ | 000100 | $Z_4$ | 100000 | $Y_4$ | 100100 |
| $X_5$ | 000101 | $Z_5$ | 101000 | $Y_5$ | 101101 |
| $X_6$ | 000110 | $Z_6$ | 110000 | $Y_6$ | 110110 |
| $X_7$ | 000111 | $Z_7$ | 111000 | $Y_7$ | 111111 |

**Example.** *Steane's code.* The generators of the stabilizer group for Steane's code are

$$\begin{aligned}
\mathbf{M}_1 &= \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes X \otimes X \otimes X \otimes X, & \mathbf{M}_2 &= \mathbb{I} \otimes X \otimes X \otimes \mathbb{I} \otimes \mathbb{I} \otimes X \otimes X, \\
\mathbf{M}_3 &= X \otimes \mathbb{I} \otimes X \otimes \mathbb{I} \otimes X \otimes \mathbb{I} \otimes X, & \mathbf{M}_4 &= \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes Z \otimes Z \otimes Z, \\
\mathbf{M}_5 &= \mathbb{I} \otimes Z \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes Z, & \mathbf{M}_6 &= Z \otimes \mathbb{I} \otimes Z \otimes \mathbb{I} \otimes Z \otimes \mathbb{I} \otimes Z,
\end{aligned}$$

The weight of all stabilizers is four, $w(\mathbf{M}_i) = 4,\ 1 \leq i \leq 6$ and the distance is $d = 3$ thus, the code is a nondegenerate stabilizer code.

Table 23 summarizes the single error operators and generators that anticommute with each operator for Steane's code; Table 24 presents the single errors and the corresponding syndromes for the same code. Subscripts indicate positions of errors; one bit-flip error is denoted as $X_{1-7}$; one phase-flip error as $Z_{1-7}$; the combination of one bit-flip and one phase-flip as $Y_{1-7}$.

Next we discuss quantum codes which use the smallest number of qubits for encoding the quantum information.

## 5.13 Stabilizers for Perfect Quantum Codes

Recall from the discussion of the quantum Hamming bound in Section 5.3 that we need at least five physical qubits to encode one logical qubit. A perfect quantum error correcting code is one that uses the smallest dimension Hilbert space to encode $\mid 0\rangle_L$ and $\mid 1\rangle_L$ as entangled states in $\mathcal{H}_{2^5}$.

A perfect quantum error correcting codes is described by Laflamme *et al.* [264]. This code can be described in different ways all related by a change of base of one of the five physical qubits. We start with the definition of the code in [264], then the one in [308], and in the next section we use a third definition from DiVincenzo and Shor [129].

To introduce the two logical qubits $\mid 0\rangle_L$ and $\mid 1\rangle_L$ according to [264] we first define

$$\begin{array}{llll} \mid \varphi_1\rangle = \mid 000\rangle + \mid 111\rangle, & \mid \varphi_3\rangle = \mid 100\rangle + \mid 011\rangle, & \mid \varphi_5\rangle = \mid 010\rangle + \mid 101\rangle, & \mid \varphi_7\rangle = \mid 110\rangle + \mid 001\rangle, \\ \mid \varphi_2\rangle = \mid 000\rangle - \mid 111\rangle, & \mid \varphi_4\rangle = \mid 100\rangle - \mid 011\rangle, & \mid \varphi_6\rangle = \mid 010\rangle - \mid 101\rangle, & \mid \varphi_8\rangle = \mid 110\rangle - \mid 001\rangle. \end{array}$$

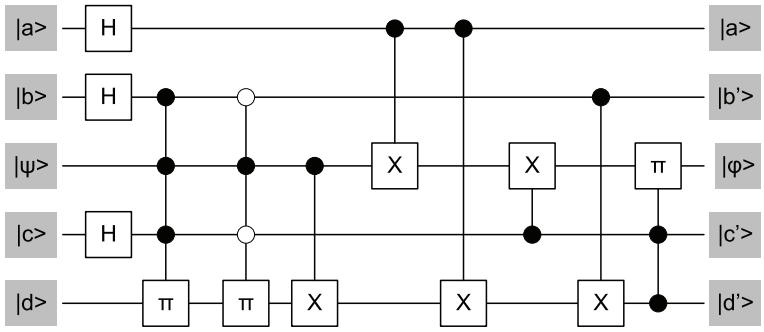



Figure 130: The encoder for the perfect quantum code of Laflamme *et al.* [264].

The two logical qubits for one of the 5-qubit codes are

$$\begin{array}{rcl} \mid 0\rangle_L & = & \mid \varphi_1\rangle \mid 00\rangle - \mid \varphi_3\rangle \mid 11\rangle + \mid \varphi_5\rangle \mid 01\rangle + \mid \varphi_7\rangle \mid 10\rangle \\ \mid 1\rangle_L & = & -\mid \varphi_2\rangle \mid 11\rangle - \mid \varphi_4\rangle \mid 00\rangle - \mid \varphi_6\rangle \mid 10\rangle + \mid \varphi_8\rangle \mid 01\rangle \end{array}$$

or

Table 25: The single errors and the corresponding syndromes for 5-qubit code.

| Error Operator | Syndrome $a_1a_2a_3a_4$ | Error Operator | Syndrome $a_1a_2a_3a_4$ | Error Operator | Syndrome $a_1a_2a_3a_4$ |
|---|---|---|---|---|---|
| $\mathbf{X}_1$ | 0110 | $\mathbf{Z}_1$ | 1000 | $\mathbf{Y}_1$ | 1110 |
| $\mathbf{X}_2$ | 0001 | $\mathbf{Z}_2$ | 0100 | $\mathbf{Y}_2$ | 0101 |
| $\mathbf{X}_3$ | 0111 | $\mathbf{Z}_3$ | 1010 | $\mathbf{Y}_3$ | 1101 |
| $\mathbf{Y}_4$ | 1011 | $\mathbf{Z}_4$ | 0010 | $\mathbf{Y}_4$ | 1001 |
| $\mathbf{X}_5$ | 0011 | $\mathbf{Z}_5$ | 1100 | $\mathbf{Y}_5$ | 1111 |

$$
\begin{aligned}
\mid 0\rangle_L &= \quad \mid 00000\rangle + \mid 11100\rangle - \mid 10011\rangle - \mid 01111\rangle + \mid 01001\rangle + \mid 10101\rangle + \mid 11010\rangle + \mid 00110\rangle \\
\mid 1\rangle_L &= -\mid 00011\rangle + \mid 11111\rangle - \mid 10000\rangle + \mid 01100\rangle - \mid 01010\rangle + \mid 10110\rangle + \mid 11001\rangle - \mid 00101\rangle
\end{aligned}
$$

Table 26: The resulting state $\mid \varphi\rangle$ after an error.

| Error(s) | $\mid \varphi\rangle$ |
|---|---|
| $\mathbf{Y}_3$ | $-\alpha_0 \mid 1\rangle + \alpha_1 \mid 0\rangle$ |
| $\mathbf{Y}_5$ | $-\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle$ |
| $\mathbf{X}_2, \mathbf{Z}_3, \mathbf{Z}_5, \mathbf{Y}_2$ | $\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle$ |
| $\mathbf{X}_5, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_4$ | $-\alpha_0 \mid 0\rangle - \alpha_1 \mid 1\rangle$ |
| $\mathbf{X}_1, \mathbf{X}_3, \mathbf{X}_4, \mathbf{Y}_1, \mathbf{Y}_4$ | $-\alpha_0 \mid 1\rangle - \alpha_1 \mid 0\rangle$ |

The encoding circuit for this code is shown in Figure 130. The input state $\mid \psi\rangle$ is identical with the output state, $\mid \varphi\rangle$ when there is no error. The circuit uses single-qubit Hadamard gates, two-qubit `CNOT` gates, as well as two- and three-qubit controlled $\pi$ gates which rotate the phase of a qubit by $\pi$; the leftmost $\pi$ gate rotates the phase of $\mid d\rangle$ when all control qubits are 1; the second to the right $\pi$ gate rotates the phase of $\mid d\rangle$ when $\mid b\rangle$ and $\mid c\rangle$ are 0 and $\mid \psi\rangle$ is 1. The error operators and the syndromes are summarized in Table 25; the resulting state $\mid \varphi\rangle$ after an error occurs in the original state $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ is shown in Table 26. As before, subscripts indicate positions of errors: one bit-flip error is denoted as $X_{1-5}$; one phase-flip error as $Z_{1-5}$; the combination of one bit-flip and one phase-flip as $Y_{1-5}$.

Another version of the perfect code [308] encodes the logical qubits as

$$
\begin{aligned}
\mid 0\rangle_L = \; & \tfrac{1}{4}\,[\mid 00000\rangle + \mid 10010\rangle + \mid 01001\rangle + \mid 10100\rangle + \mid 01010\rangle - \mid 11011\rangle - \mid 00110\rangle - \mid 11000\rangle \\
& - \mid 11101\rangle - \mid 00011\rangle - \mid 11110\rangle - \mid 01111\rangle - \mid 10001\rangle - \mid 01100\rangle - \mid 10111\rangle + \mid 01101\rangle] \\
\mid 1\rangle_L = \; & \tfrac{1}{4}\,[\mid 11111\rangle + \mid 01101\rangle + \mid 10110\rangle + \mid 01011\rangle + \mid 10101\rangle - \mid 00100\rangle - \mid 11001\rangle - \mid 00111\rangle \\
& - \mid 00010\rangle - \mid 11100\rangle - \mid 00001\rangle - \mid 10000\rangle - \mid 01110\rangle - \mid 10011\rangle - \mid 01000\rangle + \mid 11010\rangle]\,.
\end{aligned}
$$

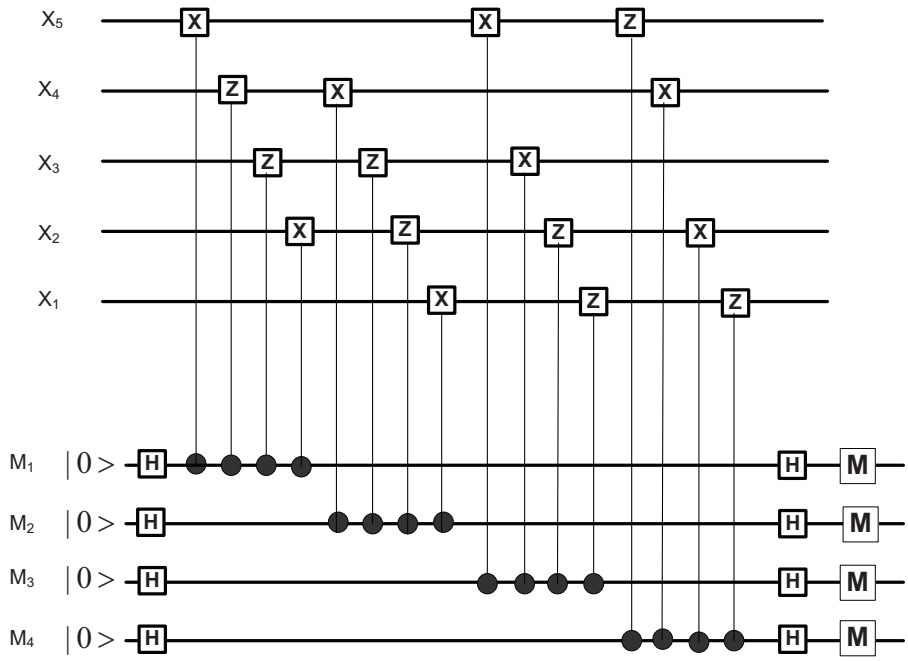



Figure 131: The quantum circuit for the syndrome measurement of the perfect quantum code defined in [308]. Ancilla qubit $i$, $1 \leq i \leq 4$ acts as control qubit for the four gates used by the generators $\mathbf{M}_i$ $1 \leq i \leq 4$.

The stabilizer $\mathcal{S}$ of this code is described by a group $M$ of 4 generators

$$M = \{\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4\}$$
$$\mathbf{M}_1 = X \otimes Z \otimes Z \otimes X \otimes \mathbb{I}, \quad \mathbf{M}_2 = \mathbb{I} \otimes X \otimes Z \otimes Z \otimes X,$$
$$\mathbf{M}_3 = X \otimes \mathbb{I} \otimes X \otimes Z \otimes Z, \quad \mathbf{M}_4 = Z \otimes X \otimes \mathbb{I} \otimes X \otimes Z.$$

The codewords, $\mid 0\rangle_L$ and $\mid 1\rangle_L$ are the eigenvectors with an eigenvalue of $(+1)$ of the stabilizers

$$\mathbf{M}_j \mid 0\rangle_L = (+1) \mid 0\rangle_L \quad \text{and} \quad \mathbf{M}_j \mid 1\rangle_L = (+1) \mid 1\rangle_L, \quad 1 \leq j \leq 4.$$

$M$ is an Abelian subgroup, each generator commutes with all the others; for example,

$$\mathbf{M}_1\mathbf{M}_3 = (X \otimes Z \otimes Z \otimes X \otimes \mathbb{I})(X \otimes \mathbb{I} \otimes X \otimes Z \otimes Z) = \mathbb{I} \otimes Z \otimes (ZX) \otimes (XZ) \otimes Z$$

$$\mathbf{M}_3\mathbf{M}_1 = (X \otimes \mathbb{I} \otimes X \otimes Z \otimes Z)(X \otimes Z \otimes Z \otimes X \otimes \mathbb{I}) = \mathbb{I} \otimes Z \otimes (XZ) \otimes (ZX) \otimes Z.$$

The Pauli operators anti-commute, $XZ = -ZX$, thus and $\mathbf{M}_1\mathbf{M}_3 = \mathbf{M}_3\mathbf{M}_1$.

The circuit for computing the syndrome is shown in Figure 131. The syndrome $\Sigma$ consists of four ancilla qubits $a_1a_2a_3a_4$, each one measured using one of the four generators, $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \mathbf{M}_4$, see Table 27. Table 28 lists single error operators and the generator(s)

Table 27: The single errors and the corresponding syndromes for the second definition of the perfect quantum code.

| Error Operator | Syndrome $a_1a_2a_3a_4$ | Error Operator | Syndrome $a_1a_2a_3a_4$ | Error Operator | Syndrome $a_1a_2a_3a_4$ |
|---|---|---|---|---|---|
| $\mathbf{X}_1$ | 0001 | $\mathbf{Z}_1$ | 1010 | $\mathbf{Y}_1$ | 1011 |
| $\mathbf{X}_2$ | 1000 | $\mathbf{Z}_2$ | 0101 | $\mathbf{Y}_2$ | 1101 |
| $\mathbf{X}_3$ | 1100 | $\mathbf{Z}_3$ | 0010 | $\mathbf{Y}_3$ | 1110 |
| $\mathbf{Y}_4$ | 0110 | $\mathbf{Z}_4$ | 1001 | $\mathbf{Y}_4$ | 1111 |
| $\mathbf{X}_5$ | 0011 | $\mathbf{Z}_5$ | 0100 | $\mathbf{Y}_5$ | 0111 |

Table 28: Single error operators for the second definition of the perfect quantum code and the generator(s) that anti-commute with each operator.

| Error Operator | Generator(s) | Error Operator | Generator(s) | Error Operator | Generator(s) |
|---|---|---|---|---|---|
| $\mathbf{X}_1$ | $\mathbf{M}_4$ | $\mathbf{Z}_1$ | $\mathbf{M}_1,\ \mathbf{M}_3$ | $\mathbf{Y}_1$ | $\mathbf{M}_1,\ \mathbf{M}_3,\ \mathbf{M}_4$ |
| $\mathbf{X}_2$ | $\mathbf{M}_1$ | $\mathbf{Z}_2$ | $\mathbf{M}_2,\ \mathbf{M}_4$ | $\mathbf{Y}_2$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_4$ |
| $\mathbf{X}_3$ | $\mathbf{M}_1,\ \mathbf{M}_2$ | $\mathbf{Z}_3$ | $\mathbf{M}_3$ | $\mathbf{Y}_3$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_3$ |
| $\mathbf{X}_4$ | $\mathbf{M}_2,\ \mathbf{M}_3$ | $\mathbf{Z}_4$ | $\mathbf{M}_1,\ \mathbf{M}_4$ | $\mathbf{Y}_4$ | $\mathbf{M}_1,\ \mathbf{M}_2,\ \mathbf{M}_3,\ \mathbf{M}_4$ |
| $\mathbf{X}_5$ | $\mathbf{M}_3,\ \mathbf{M}_4$ | $\mathbf{Z}_5$ | $\mathbf{M}_2$ | $\mathbf{Y}_5$ | $\mathbf{M}_2,\ \mathbf{M}_3,\ \mathbf{M}_4$ |

which anti-commute with each operator; for example, $\mathbf{X}_1$ (where $\mathbf{X}_1 = X \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}$) anti-commutes with $\mathbf{M}_4$, thus a bit-flip on the first qubit can be detected; $\mathbf{Z}_1$ anti-commutes with $\mathbf{M}_1$ and $\mathbf{M}_3$ thus, a phase-flip of the first qubit can also be detected. Since each of these 15 errors anti-commute with distinct subsets of $\mathcal{S}$ we are able to distinguish individual errors and then correct them.

The code cannot detect two qubit errors; for example, the two bit-flip error $\mathbf{X}_1\mathbf{X}_2 = X \otimes X \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}$ is indistinguishable from $\mathbf{Z}_4 = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes \mathbb{I}$ because both $\mathbf{X}_1\mathbf{X}_2$ and $\mathbf{Z}_4$ anti-commute with the same subset of stabilizers, $\{\mathbf{M}_1, \mathbf{M}_4\}$, and give the same error syndrome. We conclude that *the 5-qubit code can correct any single qubit error, but cannot correct two qubit errors.*

## 5.14 Quantum Restoration Circuits

The parity of a qubit is given by the eigenvalue of the Pauli operator $Z$ applied to the qubit; the parity of an $n$-qubit codeword is given by the eigenvalue of the product operator $\sigma_{z,1}\sigma_{z,2}, \ldots, \sigma_{z,n}$; an eigenvalue of $+1$ corresponds to even parity and $-1$ to odd parity of the qubit or of the codeword.

Now we discuss a family of circuits DiVincenzo calls quantum restoration circuits [130]. We examine a circuit described in [129] which restores a five-qubit codeword to its original

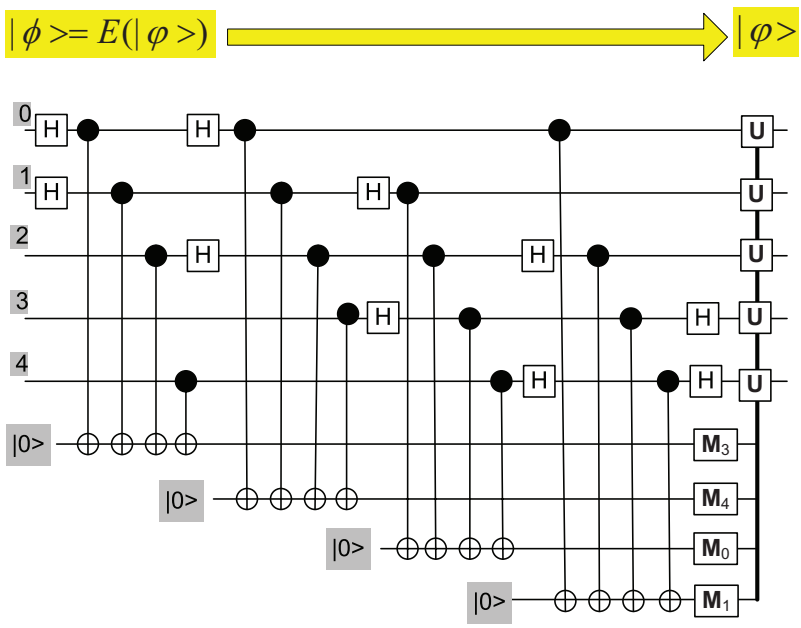



Figure 132: A circuit for the restoration of the qubit $\mid \varphi\rangle = \alpha_0 \mid 0\rangle_L + \alpha_1 \mid 1\rangle_L$ after a single error; the original qubit is encoded using 5 physical qubits, labelled as 0, 1, 2, 3, and 4. The eigenvalues of four measurement operators, $\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_3$ and $\mathbf{M}_4$ are stored in four ancilla qubits initially in state $\mid 0\rangle$. The $\mathbf{M}_3 = X_0X_1Z_2Z_4$ operator uses the first group of four `CNOT gates` with qubits $0, 1, 2$ and 4 as source and the first ancilla qubit as target; the $\mathbf{M}_4 = X_1X_2Z_3Z_0$ operator uses the second group of four `CNOT gates` with qubits $0, 1, 2$ and 3 as source and the second ancilla qubit as target, and so on. When a single error occurs, the non-demolition measurements uniquely identify the qubit in error, as well as the type of error and drive unitary transformations $\mathbf{U}$ used to restore the physical qubits to their original state. Two Hadamard gates $H$ on each line representing a physical qubit rotate the basis to measure the eigenvalues of $X$ when necessary; the state of the qubit is not affected as $\mathbf{HH} = \mathbf{I}$.

state after determining the error syndrome which identifies the location and the type of a single-qubit error, Figure 132. When both the control and the target qubit of a `CNOT` gate are in a rotated base then the role of the control and target are reversed, Figure 12 (d).

The encoding of the logical basis states ($\mid 0\rangle_L$ and ($\mid 1\rangle_L$) is

$$
\begin{aligned}
\mid 0\rangle_L \;=& \mid 00000\rangle \\
&+ \mid 11000\rangle + \mid 01100\rangle + \mid 00110\rangle + \mid 00011\rangle + \mid 10001\rangle \\
&- \mid 10100\rangle - \mid 01010\rangle - \mid 00101\rangle - \mid 10010\rangle - \mid 01001\rangle \\
&- \mid 11110\rangle - \mid 01111\rangle - \mid 10111\rangle - \mid 11011\rangle - \mid 11101\rangle
\end{aligned}
$$

$$
\begin{aligned}
\mid 1\rangle_L \;=& \mid 11111\rangle \\
&+ \mid 00111\rangle + \mid 10011\rangle + \mid 11001\rangle + \mid 11100\rangle + \mid 01110\rangle \\
&- \mid 01011\rangle - \mid 10101\rangle - \mid 11010\rangle - \mid 01101\rangle - \mid 10110\rangle \\
&- \mid 00001\rangle - \mid 10000\rangle - \mid 01000\rangle - \mid 00100\rangle - \mid 00010\rangle.
\end{aligned}
$$

Four measurement operators, $\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_3$ and $\mathbf{M}_4$, allow us to identify the qubit in er-

Table 29: The syndrome and the error for the restoration circuit.

| $\mathbf{M}_3$ | $\mathbf{M}_4$ | $\mathbf{M}_0$ | $\mathbf{M}_1$ | $\mathbf{E}_i$ | $\mathbf{M}_3$ | $M_4$ | $M_0$ | $\mathbf{M}_1$ | $\mathbf{E}_i$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | $I_0I_1I_2I_3I_4$ | 1 | 0 | 0 | 0 | $\mathbf{Z}_0I_1I_2I_3I_4$ |
| 0 | 0 | 0 | 1 | $I_0I_1I_2I_3\mathbf{Z}_4$ | 1 | 0 | 0 | 1 | $I_0I_1\mathbf{X}_2I_3I_4$ |
| 0 | 0 | 1 | 0 | $I_0\mathbf{X}_1I_2I_3I_4$ | 1 | 0 | 1 | 0 | $I_0I_1I_2I_3\mathbf{X}_4$ |
| 0 | 0 | 1 | 1 | $I_0I_1I_2\mathbf{Z}_3I_4$ | 1 | 0 | 1 | 1 | $I_0I_1I_2I_3\mathbf{Y}_4$ |
| 0 | 1 | 0 | 0 | $I_0I_1I_2\mathbf{X}_3I_4$ | 1 | 1 | 0 | 0 | $I_0\mathbf{Z}_1I_2I_3I_4$ |
| 0 | 1 | 0 | 1 | $\mathbf{X}_0I_1I_2I_3I_4$ | 1 | 1 | 0 | 1 | $\mathbf{Y}_0I_1I_2I_3I_4$ |
| 0 | 1 | 1 | 0 | $I_0I_1\mathbf{Z}_2I_3I_4$ | 1 | 1 | 1 | 0 | $I_0\mathbf{Y}_1I_2I_3I_4$ |
| 0 | 1 | 1 | 1 | $I_0I_1I_2\mathbf{Y}_3I_4$ | 1 | 1 | 1 | 1 | $I_0I_1\mathbf{Y}_2I_3I_4$ |

ror, they are tensor products of Pauli operators applied to the qubits of a codeword; the measurement operators are

$$\mathbf{M}_3 = X_0X_1Z_2Z_4, \quad \mathbf{M}_4 = X_1X_2Z_3Z_0, \quad \mathbf{M}_0 = X_2X_3Z_4Z_1, \quad \text{and} \quad \mathbf{M}_1 = X_3X_4Z_0Z_2.$$

The circuit in Figure 132 uses four groups, each consisting of four `CNOT` gates, to compute the eigenvalues of the measurement operators and then store the result in four ancilla qubits initially in state $\mid 0\rangle$. Table 29 lists the 16 possible errors, no-error $I_i$ included. When a single error occurs, the non-demolition measurements uniquely identify the qubit in error and the type of error and drive unitary transformations $\mathbf{U}$ used to restore the qubit affected by error to its original state.

The analysis of this circuit in [129] is important because it establishes that errors in "all known quantum error-correcting codes can be corrected in a fault-tolerant way."

Error correction requires an increased number of physical qubits and quantum gates as well as an increased number of computational steps, thus it increases the vulnerability of the system. For example, we have seen that to restore one logical qubit we need 9 (nine) physical qubits as well as 12 `CNOT` gates, 12 `Hadamard` gates and several additional gates implementing Pauli transformations. This simple example shows that the overhead of error correction discussed in Section 5.3 cannot be ignored and raises the question if error correction is really feasible, if the error correction does not introduce more errors and extends the number of steps and implicitly the computation time beyond the ability of physical qubits to maintain the state required by the computational steps of an algorithm.

A critical question addressed early on was if arbitrary long computations can be carried out with a quantum computing device whose state is subject to decoherence. The answer to this question was provided by the quantum accuracy threshold theorem [5, 240, 247, 343], which states that an arbitrarily long computation can be carried out reliably provided that all sources of "noise" affecting a quantum computer are weaker than the so called "accuracy threshold." A lower bound on this threshold, $\epsilon > 1.9 \times 10^{-4}$ was derived in [9].

## 5.15 Quantum Codes over $GF(p^k)$

The quantum codes we have constructed so far have been over $GF(2)$, the binary field with two elements. Codes over extension fields $GF(p^k)$ have distinct advantages, see Section 4.15; this motivates us to explore the properties of linear codes over the finite field $GF(p^k)$ [179].

The <u>trace</u> $\mathrm{tr} : GF(p^k) \mapsto GF(p)$ is a linear mapping from $GF(p^k)$ to $GF(p)$

$$\mathrm{tr}(\alpha) = \alpha_1^{p^0} + \alpha_2^{p^1} + \ldots + \alpha_k^{p^{k-1}}.$$

The elements $\alpha_1, \alpha_2, \ldots, \alpha_k$ form a basis of $GF(p^k)$ over $GF(p)$ if and only if

$$\det \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \\ \alpha_1^p & \alpha_2^p & \cdots & \alpha_k^p \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{p^{k-1}} & \alpha_2^{p^{k-1}} & \cdots & \alpha_k^{p^{k-1}} \end{pmatrix} \neq 0.$$

The set $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a <u>trace orthogonal basis</u> if

$$\mathrm{tr}(\alpha_i \alpha_j) = 0, \quad 1 \leq i, j \leq k, i \neq j.$$

If in addition $\mathrm{tr}(\alpha_i^2) = 1$ then $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a <u>self-dual basis</u>.

The finite field $GF(p^k)$ is a $k$-dimensional vector space over $GF(p)$; once we choose a basis $\mathcal{B}$ there is a homomorphism $\mathcal{B} : GF(p^k) \mapsto [GF(p)]^k$

$$v \mapsto \mathcal{B}(v), \quad \forall v \in GF(q^k).$$

This homomorphism allows us to express any vector $v \in GF(q^k)$ in terms of the basis $\mathcal{B}$; this mapping is described by an $k \times k$ matrix. Any linear transformation $A \in GL(n, GF(p^k))$, described by the $n \times n$ matrix $A = [a_{ij}]$, $1 \leq i, j \leq n$, can be expressed by substituting $a_{ij}$ with $\mathcal{B}(a_{ij})$ and the change in the ground field, $GF(p)$ can be carried out before the mapping $A$, or after the mapping, as shown by the commutative diagram with $q = p^k$

$$\begin{array}{ccc} GF(q^n) & \stackrel{A}{\longrightarrow} & GF(q^n) \\ \mathcal{B} \downarrow & & \downarrow \mathcal{B} \\ GF(p^{kn}) & \stackrel{\mathcal{B}(A)}{\longrightarrow} & GF(p^{kn}) \end{array}.$$

If we consider two vectors $v, w \in GF(q^k)$ then the product $v \cdot w$ can be represented using a linear mapping $\mathcal{M}_{\mathcal{B}}$ as:

$$\mathcal{B}(v \cdot w) = \mathcal{B}(w) \cdot \mathcal{M}_{\mathcal{B}}(v).$$

Let $\mathcal{B}^{\perp} = (b'_1, b'_2, \ldots, b'_k)$ be the dual of the basis $\mathcal{B} = (b_1, b_2, \ldots, b_k)$; by definition

$$\mathrm{tr}(b_i b'_j) = \delta_{ij},\ 1 \leq i, j \leq k.$$

**Proposition.** *Let $\mathcal{C}^\perp$ be the dual of the $[N, K]$ linear code $\mathcal{C}$ over the finite field $GF(2^k)$. Let $\mathcal{B}(\mathcal{C})$ be the binary expansion of the code $\mathcal{C}$ with respect to basis $\mathcal{B}$ and $\mathcal{B}^\perp(\mathcal{C}^\perp)$ the dual of $\mathcal{C}$ with respect to the dual basis $\mathcal{B}^\perp$; then*

$$\mathcal{B}^\perp(\mathcal{C}^\perp) = [\mathcal{B}(\mathcal{C})]^\perp.$$

*Proof:* The proof follows the one in [179]. Two arbitrary codewords, $c = (c_1 c_2 \ldots c_N) \in \mathcal{C}$ and $d = (d_1 d_2 \ldots d_N) \in \mathcal{C}^\perp$ are orthogonal

$$c \cdot d = 0 \quad \Longrightarrow \quad \sum_{i=1}^{N} c_i d_i = \sum_{i=1}^{N} \left( \sum_{j=1}^{k} c_{ij} b_j \right) \left( \sum_{l=1}^{k} d_{il} b'_l \right) = 0$$

with $c_{ij}, 1 \leq j \leq k$ the binary expression of $c_i \in GF(p^k)$ and $d_{il}, 1 \leq l \leq k$ the binary expression of $d_i \in GF(p^k)$; we wish to show that

$$\sum_{i=1}^{N} \sum_{j=1}^{k} c_{ij} d_{ij} = 0.$$

If we take the trace of the transformations $\mathcal{B}$ and $\mathcal{B}^\perp$ in the expression of $c \cdot d$ we observe that

$$\sum_{i=1}^{N} \sum_{j=1}^{k} \sum_{l=1}^{k} c_{ij} d_{il} \mathrm{tr}\,(b_j b'_l) = \sum_{i=1}^{N} \sum_{l=1}^{k} c_{ij} d_{ij} = 0$$

because $\mathrm{tr}\,(b_j b'_l) = \delta_{jl}$. □

This proposition allows us to conclude that the following diagram is commutative

$$\begin{array}{ccc} \mathcal{C} & \longrightarrow & \mathcal{C}^\perp \\ \text{basis } \mathcal{B} \downarrow & & \downarrow \text{ dual basis } \mathcal{B}^\perp \\ \mathcal{B}(\mathcal{C}) & \longrightarrow & \mathcal{B}^\perp(\mathcal{C}^\perp) \end{array}$$

**Quantum circuits for invertible linear transformations in $GF(2^n)$.** We have seen that a linear mapping in $GF(2^n)$ can be reduced to mappings in $GF(2)$ thus, we first survey invertible linear transformations in $GF(2)$. The linear transformation $P \in GL(n, GF(2))$ that permutes the basis vectors in an $n$-dimensional vector space $\mid i\rangle \in (\mathcal{H}_2)^{\otimes n}$, $\mid i\rangle \mapsto P \mid i\rangle$, can be implemented using at most $3(n-1)$ `CNOT` gates. Indeed, any permutation $P$ that transforms an $n$-tuple over $GF(2)$ can be implemented as a product of at most $(n-1)$ transpositions; also, a transposition $(i, j)$ can be implemented by a quantum circuit using three `CNOT` gates,

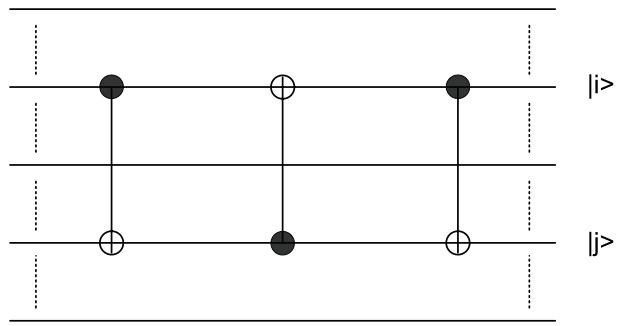



Figure 133: Any transposition $(i, j)$ can be implemented by a quantum circuit using three CNOT gates.

Figure 133; it follows that any permutation of the basis vectors in $(\mathcal{H}_2)^{\otimes n}$ requires at most $3(n-1)$ CNOT gates.

Now we show that any linear transformation $A \in GL(n, GF(2))$, invertible in $GF(2^n)$ that carries out the transformation of basis vectors $\mid i\rangle \mapsto A \mid i\rangle, \quad \mid i\rangle \in (\mathcal{H}_2)^{\otimes n}$, can be implemented using at most $(n-1)(n+3)$ CNOT gates.
Indeed, any matrix $A \in GL(n, GF(2))$ can be decomposed as a product $A = P \cdot L \cdot U$ with $P$ a permutation matrix and with $L$ and $U$ lower and, respectively, upper triangular matrices.

The lower triangular matrix $L = [L_{p,q}]_{1 \leq p,q \leq n}$ can be factored as $L = L_1 L_2 \ldots L_i \ldots L_n$ with $L_i$ obtained by replacing the $i$-th row of an identity matrix with the $i$-th row of $L$. Let the binary representation of the basis vector $\mid k\rangle$ be $k = (k_1, k_2, \ldots, k_i, \ldots, k_n)$. We multiply the vector $k$ and $L_i$

$$kL_i = (k_1 + k_i L_{i,1}, \ldots, k_{i-1} L_{i,i-1}, k_i, k_{i+1}, \ldots, k_n)$$

and notice that the $j$-th position of $k$ is inverted if both $k_i$ and $L_{i,j}$ are equal to one.

This operation requires at most $(i-1)$ CNOT gates with control qubit $i$ and target qubit $j$ whenever $L_{i,j} = 1$. Thus the implementation of $L$ requires $n(n-1)/2$ CNOT gates. A similar reasoning leads to the conclusion that implementation of $U$ requires the same number of CNOT gates. The total number of CNOT gates required for the implementation of $P$, $L$, and $U$ is

$$3(n-1) + \frac{n(n-1)}{2} + \frac{n(n-1)}{2} = (n-1)(n+3).$$

Thus, the total number of CNOT gates required for the implementation of a linear transformation $A$ is $\mathcal{O}(n^2)$.

To reduce a linear mapping in $GF(2^n)$ to a mapping in $GF(2)$ we first fix a basis $\mathcal{B}$ in $GF(2^k)$. Then we extend the homomorphism $\mathcal{B}$ and map $n$-dimensional vectors $v$ with elements in $GF(2^k)$ to binary vectors $w$ of length $kn$, as shown earlier.

**Cyclic discrete Fourier transform ($DFT$) over $GF(p^k)$.** Quantum Reed-Solomon codes discussed in the next section can be described by means of cyclic $DFT$ over $GF(p^k)$. If $n$ divides $2^k - 1$ and $\beta$ is an element of order $n$ in $GF(p^k)$ then the cyclic $DFT$ over $GF(p^k)$ is described by the matrix:

$$DFT = (\beta^{ij})_{0 \le i,j \le n-1}.$$

The condition that $n$ divides $2^k - 1$ ensures that $GF(2^k)$ has a primitive element of order $n$ and thus $DFT \in GL(n, GF(2^k))$. It is easy to see that the cyclic $DFT$ over $GF(p^k)$ can be implemented using $\mathcal{O}(n^2k^2)$ `CNOT` gates. This follows immediately from the fact that $DFT$ is a linear transformation and, according to our observation, it requires $\mathcal{O}(N^2)$ `CNOT` gates for vectors of length $N = kn$. The inverse Fourier Transform is denoted as $DFT^{-1}$.

## 5.16 Quantum Reed-Solomon Codes

Quantum Reed-Solomon (QRS) codes can be constructed as a family of CSS codes. In Section 4.15 we show that a classical $[N, K, d]$ Reed-Solomon code over $GF(2^k)$ with length $N = 2^k - 1$, distance $d$, and dimension $K = N - d + 1$, is a cyclic code with the generator polynomial

$$g(x) = (x - \beta^j)(x - \beta^{j+1}) \ldots (x - \beta^{j+d-2})$$

with $\beta$ a primitive element of $GF(2^k)$ of order N. A QRS code encodes $k(2N - K)$ qubits in states $\mid \varphi_1\rangle, \mid \varphi_2\rangle, \ldots, \mid \varphi_{k(2N-K)}\rangle$ to $kN$ qubits.

**Proposition.** *The $[N, K, d]$ Reed-Solomon code $\mathcal{C}$ is self-dual when $j = 0$ and $d > N/2 + 1$.*

*Proof:* If $g(x) = (x - \beta^0)(x - \beta^1) \ldots (x - \beta^{d-2})$ is the generator polynomial of $\mathcal{C}$ then the generator polynomial of the dual code $\mathcal{C}^\perp$ is the reciprocal polynomial of $(x^N - 1)/g(x)$

$$g^\perp(x) = (x - \beta^{-(d-1)})(x - \beta^{-d}) \ldots (x - \beta^{-(N-1)}).$$

$N = 2^k - 1$ and $\beta$ is a primitive element of $GF(2^k)$ thus, $\beta^N = 1$ and we can write

$$g^\perp(x) = (x - \beta^1)(x - \beta^2) \ldots (x - \beta^{(N-d+1)}).$$

The highest power of $\beta$ in the expression of $g^\perp(x)$, satisfies the inequality $N - d + 1 \le d - 1$; indeed, $d > N/2 + 1 \Longrightarrow N \le 2d - 2$ or $N - d + 1 \le 2d - 2 - d + 1$. It follows that $g^\perp(x)$ is a divisor of $g(x)$ thus $\mathcal{C} \subset \mathcal{C}^\perp$. □

Reed-Solomon codes can be described by means of cyclic Discrete Fourier Transform ($DFT$) over $GF(2^k)$. The <u>spectrum</u> $\bar{c}$ of a vector $c = \sum_{i=1}^n \gamma_i c_i$ with elements $\gamma_i \in GF(2^k)$ is

$$\bar{c} = c \cdot DFT \quad \Longrightarrow \quad c = \bar{c} \cdot DFT^{-1}.$$

The following proposition will allow us to show that a classical Reed-Solomon code $\mathcal{C}$ can be described as the set of $N$-tuples over $GF(2^k)$ with the first $d$ components equal to zero

$$\mathcal{C} = \{\bar{w} \cdot DFT^{-1}\} \text{ with } w = (w_1, w_2 \ldots, w_d, w_{d+1}, \ldots w_N) \in GF(2^k), w_1 = w_2 = \ldots = w_d = 0.$$

**Proposition.** *The spectrum $\bar{c}$ of a codeword $c$ of a weakly self-dual $[N, K, d]$ Reed-Solomon code over $GF(2^k)$ has its first $d - 1 = N - K$ components equal to zero, and the next $K$ components with an arbitrary value; the spectrum $\bar{c}'$ of a codeword of the dual code $c' \in \mathcal{C}^\perp$*

*has its first component as well as the last $d-2$ components with an arbitrary value while the other $K$ components are equal to zero.*

*Proof:* The transformation carried out by the cyclic $DFT$ over $GF(2^k)$ with a primitive element $\beta$ is given by the $N \times N$ matrix

$$DFT = (\beta^{ij})_{0 \le i,j \le N-1} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & (\beta^1)^1 & (\beta^1)^2 & \dots & (\beta^1)^{N-2} & (\beta^1)^{N-1} \\ 1 & (\beta^2)^1 & (\beta^2)^2 & \dots & (\beta^2)^{N-2} & (\beta^2)^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & (\beta^{K-1})^1 & (\beta^{K-1})^2 & \dots & (\beta^{K-1})^{N-2} & (\beta^{K-1})^{N-1} \\ 1 & (\beta^K)^1 & (\beta^K)^2 & \dots & (\beta^K)^{N-2} & (\beta^K)^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & (\beta^{N-1})^1 & (\beta^{N-1})^2 & \dots & (\beta^{N-1})^{N-2} & (\beta^{N-1})^{N-1} \end{pmatrix}.$$

In Section 4.15 we introduced the generator and the parity check matrix of a classical Reed-Solomon code. From that discussion it follows that the transpose of the parity check matrix of the weakly self-dual $[N, K, d]$ Reed-Solomon code is the $N \times (N-K)$ matrix

$$\mathrm{H}^T = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ (\beta^1) & (\beta^1)^2 & \dots & (\beta^1)^{N-K-1} & (\beta^1)^{N-K} \\ (\beta^2) & (\beta^2)^2 & \dots & (\beta^2)^{N-K-1} & (\beta^2)^{N-K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\beta^{N-1}) & (\beta^{N-1})^2 & \dots & (\beta^{N-1})^{N-K+1} & (\beta^{N-1})^{N-K} \end{pmatrix}$$

The cyclic $DFT$ can then be expressed as

$$DFT = \big([\mathrm{H}^T][A]\big)$$

with $A$ the $N \times K$ matrix given by

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ (\beta^1)^{N-K} & (\beta^1)^{N-K+1} & \dots & (\beta^1)^{N-2} & (\beta^1)^{N-1} \\ (\beta^2)^{N-K} & (\beta^2)^{N-K+1} & \dots & (\beta^2)^{N-2} & (\beta^2)^{N-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\beta^{N-1})^{N-K} & (\beta^{N-1})^{N-K+1} & \dots & (\beta^{N-1})^{N-2} & (\beta^{N-1})^{N-1} \end{pmatrix}.$$

If H is the parity check matrix of a linear code $\mathcal{C}$ then and $c \in \mathcal{C}$ then

$$\mathrm{H}c^T = c\mathrm{H}^T = 0.$$

It follows that if $\mathcal{C}$ is a weakly self-dual $[N, K, d]$ Reed-Solomon code and $c \in \mathcal{C}$ then

$$\bar{c} = c \cdot DFT = c \cdot [\mathrm{H}^T A].$$

The spectrum of any codeword has the first $N - K = d - 1$ components equal to zero; the last $N - d + 1$ components may have an arbitrary value, Figure 134 (a).

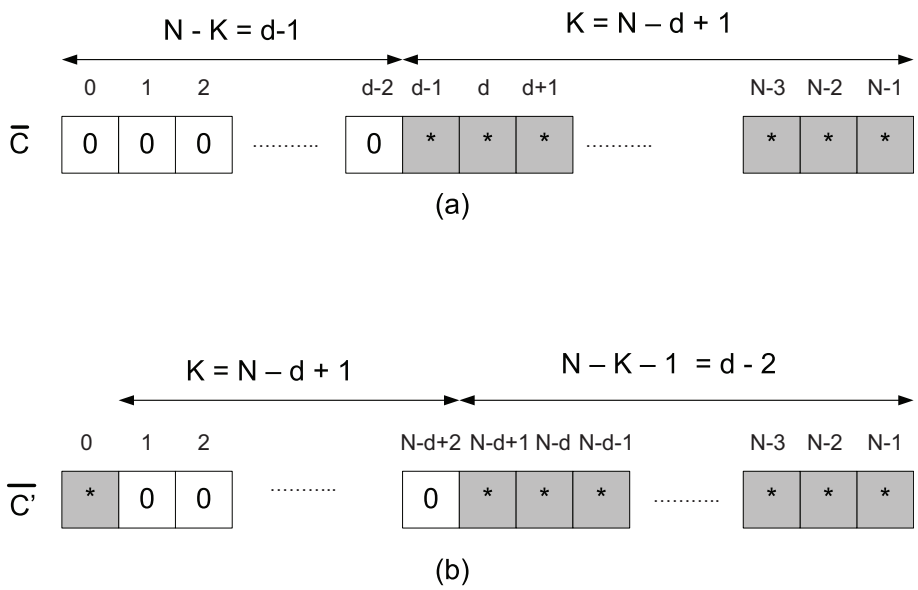



Figure 134: (a) The spectrum of codeword $c \in \mathcal{C}$ has its first $d-1$ elements equal to zero and the last $(N-d+1)$ elements may take arbitrary values. (b) The spectrum of a codeword belonging to the dual code, $c' \in \mathcal{C}^{\perp}$, allows its first component to have an arbitrary value, the next $(N-d+1)$ components are equal to zero, and the last $(d-2)$ components may have arbitrary values.

We can identify the zero components of a codeword of the dual code $\mathcal{C}^{\perp}$ knowing that the generator of the original code $\mathcal{C}$ is the parity check matrix of the dual code. The transpose of $G$, the generator matrix of the code, is the $N \times (N-K)$ matrix

$$\mathrm{G}^T = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & (\beta^1)^1 & (\beta^2)^1 & \dots & (\beta^{K-2})^1 & (\beta^{K-1})^1 \\ 1 & (\beta^1)^2 & (\beta^2)^2 & \dots & (\beta^{K-2})^2 & (\beta^{K-1})^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & (\beta^1)^{N-1} & (\beta^2)^{N-1} & \dots & (\beta^{K-2})^{N-1} & (\beta^{K-1})^{N-1} \end{pmatrix}$$

The cyclic $DFT$ can then be expressed as

$$DFT = \big([1][\mathrm{G}^T][B]\big)$$

with $B$ the $N \times (K-1)$ matrix

$$B = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ (\beta^K)^1 & (\beta^{K+1})^1 & \dots & (\beta^{N-2})^1 & (\beta^{N-1})^1 \\ (\beta^K)^2 & (\beta^{K+1})^2 & \dots & (\beta^{N-2})^2 & (\beta^{N-1})^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\beta^K)^{N-1} & (\beta^{K+1})^{N-1} & \dots & (\beta^{N-2})^{N-1} & (\beta^{N-1})^{N-1} \end{pmatrix}.$$

If G is the generator matrix of $\mathcal{C}$ and $c' \in \mathcal{C}^{\perp}$ then

$$\mathrm{G} \cdot (c')^T = c' \cdot \mathrm{G}^T = 0.$$

It follows that if $\mathcal{C}$ is a weakly self-dual $[N, K, d]$ Reed-Solomon code and $c' \in \mathcal{C}^\perp$ then

$$\bar{c'} = c' \cdot DFT = c' \cdot ([1][\mathrm{G}^T][B]).$$

The $DFT$ of any codeword of the dual code, has a first component with an arbitrary value, the next $(N - d + 1)$ components are equal to zero, and the last $(d - 2)$ components may also have an arbitrary value, Figure 134 (b). □

$\mathcal{C}$ is an $[N, K, d]$ self-dual binary code if and only if $\mathcal{C} \subset \mathcal{C}^\perp$. Let $w_j, 1 \leq j \leq 2^{N-2K}$ be a representation of the cosets $\mathcal{C}^\perp/\mathcal{C}$; then the basis states of the quantum code CSS, generated according to the rules discussed in Section 5.10 are

$$\mid s\rangle = \frac{1}{\sqrt{\mid \mathcal{C} \mid}} \sum_{c \in \mathcal{C}} \mid c + w_j\rangle.$$

Let $\mathcal{C}$ be a classical $[N, K, d]$ Reed-Solomon code such that $d > N/2 + 1$ and $j = 0$ and $\mathcal{B}$ be a self-dual basis of $GF(2^k)$ over $GF(2)$; then the quantum Reed-Solomon code $\mathcal{Q}^\mathcal{C}$ is the CSS code of length $kN$ derived from the weakly self-dual code $\mathcal{B}(\mathcal{C})$.

Now we discuss the dimension of a CSS code constructed based on the weakly self-dual $[N, K, d]$ code $\mathcal{C}$; If $\dim(\mathcal{C}) = K$ then $\dim(\mathcal{C}^\perp) = N - K$ and

$$\dim(\mathcal{Q}^\mathcal{C}) = K - (N - K) = 2K - N.$$

The duality is a symmetric relationship[39]; $\mathcal{C}^\perp$ could be considered the original code and $\mathcal{C}$ its dual. The dimension must be a positive integer, therefore

$$\dim(\mathcal{Q}^\mathcal{C}) = \begin{cases} 2K - N & \text{if} \quad K > N/2 \\ N - 2K & \text{if} \quad K < N/2. \end{cases}$$

When we assume that $K < N/2$ a CSS code $\mathcal{Q}$ based on the weakly self-dual $[N, K, d]$ code $\mathcal{C}$ is an $[N, N - 2K, d]$ code with

$$d > N/2 + 1 \implies d > (K + 1).$$

We conclude that in this case the CSS code $\mathcal{Q}$ can detect at least $K$ errors.

The transformation $\mathcal{B}$ expands each $\alpha \in GF(2^k)$ to $k$ bits. It follows that $\mathcal{Q}^\mathcal{C}$ is a $[kN, k(N - 2K), d > K + 1$ Reed-Solomon code thus, it encodes $k(N - 2K)$ qubits as $kN$ qubits, see Figure 135(a) and it can detect at least $K$ errors.

Let $\mid \psi_j\rangle$ be a codeword of the quantum Reed-Solomon code, $\mathcal{Q}^\mathcal{C}$. Then $\mid \psi_j\rangle$ can be expressed in $GF(2^k)$ as

$$\mid \psi_j\rangle = \frac{1}{\mid \mathcal{C} \mid} \sum_{c \in \mathcal{C}} \mid (c + w_j)\rangle, \quad 1 \leq j \leq 2K - N.$$

[39] A relation $\mathcal{R}$ is symmetric if $a\mathcal{R}b \implies b\mathcal{R}a$.

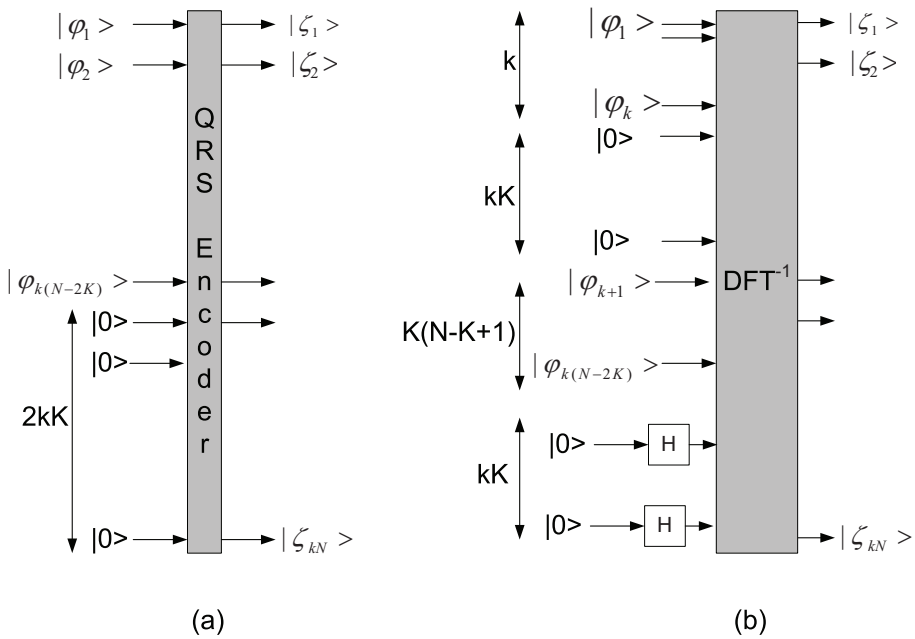



Figure 135: Encoders for Quantum Reed-Solomon codes. (a) The basic scheme for a $[kN, k(2K-N), d]$ binary QRS code; it encodes $k(N-2K)$ qubits in states $\mid \varphi_1\rangle, \mid \varphi_2\rangle, \ldots, \mid \varphi_{k(2N-K)}\rangle$ as $kN$ qubits in state, $\mid \zeta_1\rangle, \mid \zeta_2\rangle, \ldots, \mid \zeta_{kN}\rangle$ (b) A QRS encoder based on DFT.

We have $2K-N$ distinct values for $w_j$ thus, we can choose the last $K$ components of $w_j$ equal to zero; when we apply the transformation $\mathcal{B}$, then the codeword $\mid \psi_j\rangle$ can be expressed in binary as

$$\mid \psi_j\rangle = \frac{1}{\mid \mathcal{C} \mid} \sum_{c \in \mathcal{C}} \mid \mathcal{B}(c + w_j)\rangle, \quad 1 \leq j \leq 2K - N.$$

The last $kK$ components of $\mathcal{B}(w_j)$ are now zero; the spectrum of the codeword $\mid \psi_j\rangle \in \mathcal{Q}^{\mathcal{C}}$ is

$$DFT(\mid \psi_j\rangle) = \frac{1}{\mid \mathcal{C} \mid} \sum_{c \in \mathcal{C}} \mid DFT(\mathcal{B}(c)) + DFT(\mathcal{B}(w_j)\rangle = \frac{1}{\mid \mathcal{C} \mid} \sum_{c \in \mathcal{C}} \mid \bar{c} + \bar{w}_j\rangle.$$

As we have shown earlier, the spectrum of any codeword of an $[N, K, d]$ Reed-Solomon code over $GF(2^k)$ has its first $N-K$ symbols equal to zero; thus, in the binary representation, the code has its first $k(N-K)$ components equal to zero and the next $kK$ components with an arbitrary value

$$\bar{c} = (0, 0, \ldots, 0, c_{k(N-K+1)}, c_{k(N-K+2)}, \ldots, c_{kN}).$$

Now we apply a Walsh-Hadamard Transform to the last $kK$ qubits of the spectrum of $\mid \psi_i\rangle$ and leave the first $k(2K-N)$ unchanged

$$\mid \zeta\rangle = I_2^{\otimes k(N-K)} \otimes H_2^{\otimes kK} \left[DFT(\mid \psi_j\rangle\right] = I_2^{\otimes k(N-K)} \otimes H_2^{\otimes kK} \mid \bar{c} + \bar{w}\rangle.$$

The first $k(N-K)$ components of $\mid \zeta\rangle$ correspond to the DFT of $\mid w_j\rangle$; the desired result is produced by the circuit in Figure 135 (b) implementing the transformation

$$\text{Enc} = DFT^{-1} I_2^{\otimes k(N-K)} \otimes H_2^{\otimes kK}$$

applied to the input state

$$\mid \varphi\rangle = \mid \varphi_1\rangle \otimes \ldots \mid \varphi_k\rangle \otimes \mid 0\rangle \ldots \mid \varphi_{k+1}\rangle \otimes \ldots \mid \varphi_{N-2K}\rangle \otimes \mid 0\rangle \otimes \ldots \mid 0\rangle.$$

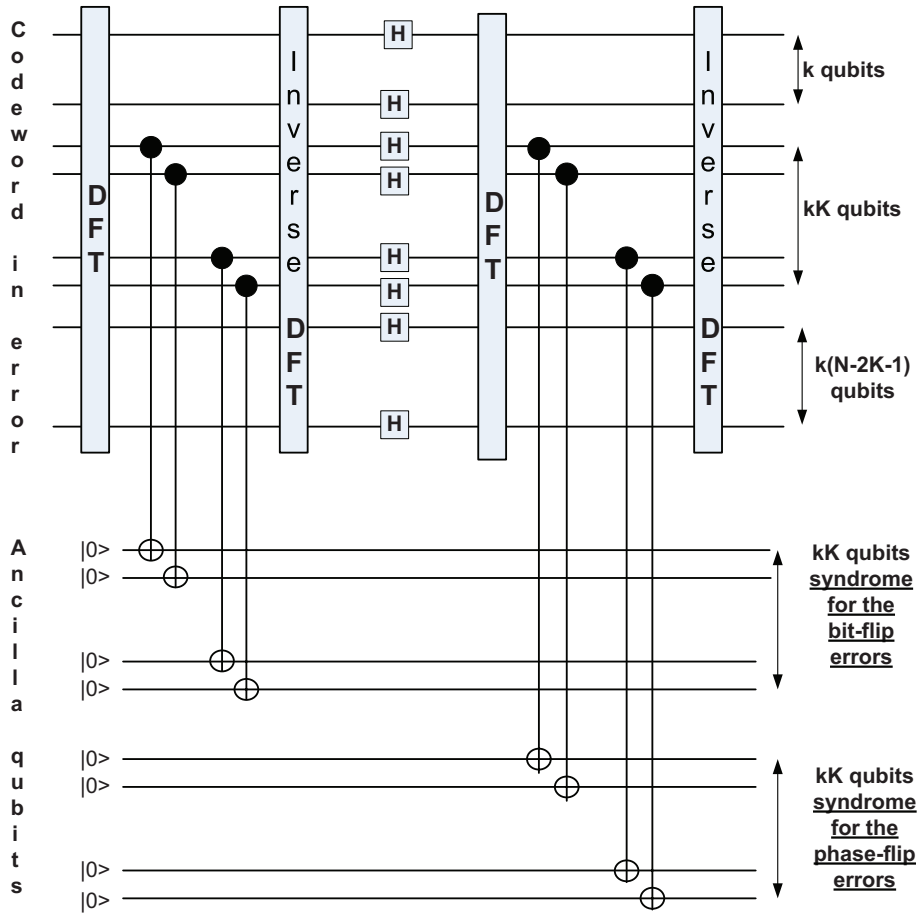



Figure 136: Quantum circuit to compute the syndrome for bit-flip and phase-flip errors for an $[N, K, d]$ quantum Reed-Solomon code.

Decoding quantum Reed-Solomon codes follows the general scheme for CSS codes discussed in Section 5.10 and illustrated in Figure 128; we compute first the syndrome for the bit-flip errors and then the syndrome for the phase-flip errors, as shown in Figure 136. Each symbol in $GF(2^k)$ is represented by $k$ qubits and encoding is done by applying the inverse cyclic DFT to a system of $kN$ qubits, the last $kK$ of them in state $\mid 0\rangle$; after computing the bit-flip syndrome we have to restore the original state, thus we apply a DFT, and then we rotate the base by applying a Walsh-Hadamard Transform, $H^{\otimes(kN)}$. The dimension of the original Reed-Solomon code $\mathcal{C}$ is $K$, therefore the bit-flip and the phase-flip syndromes require $kK$ qubits each.

**Example.** *The* $[7, 3, 5]$ *Reed-Solomon code* $\mathcal{C}$. This code discussed is Section 4.15 with $N = 7$, $K = 3$, $d = 5$, and the generator polynomial

$$g(x) = (x - \beta^0)(x - \beta^1)(x - \beta^2)(x - \beta^3)$$

has a dual, $\mathcal{C}^\perp$, a $[7, 4, 4]$ linear code with generator matrix $g^\perp(x)$

$$g^\perp(x) = \frac{x^6 - 1}{g(x)} = (x - \beta^{-4})(x - \beta^{-5})(x - \beta^{-6}) = (x - \beta^3)(x - \beta^2)(x - \beta^1).$$

The corresponding quantum code $\mathcal{Q}^\mathcal{C}$ is a $[21, 3, d]$ code; indeed, $k \times N = 7 \times 3 = 21$, $k(N - 2K) = 3(7 - 6) = 3$ and $d \geq N - K = 7 - 3 = 4$.

Next we discuss concatenated quantum codes.

## 5.17 Concatenated Quantum Codes

Classical concatenated codes discussed in Section 4.18 have two simultaneous properties:

(i) the probability of error decreases exponentially at all the rates below channel capacity, and,

(ii) the decoding complexity increases only algebraically with the number of errors corrected by the code and implicitly with the length of a codeword.

Classical concatenated codes have a multilevel structure. The data is encoded first using an $[n, k, d]$ code; then each bit in a block is again encoded using this time an $[n_1, 1, d_1]$ code and the process continues for say $l$ levels. The result is an $[n_1 n_2 \ldots n_{l-1}, k, d_1 d_2 \ldots d_{l-1}]$ code. To decode the concatenated code we first compute the error syndrome for the first level, the $[n_{l-1}, 1, d_{l-1}]$ code, for all the blocks of $n_{l-1}$ bits in parallel, then compute the error syndrome for the second level, the $[n_{l-2}, 1, d_{l-2}]$ code for all the blocks of $n_{l-2}$ bits in parallel, and continue for all the $l$ levels. Thus, we compute the error syndrome for the entire code in a number of steps equal to the sum of the number of steps to compute the syndrome at each level.

Concatenated quantum codes were introduced by Knill and Laflamme in 1996 [245] and their performance and behavior was investigated by several groups including [349]. We consider two codes:

- An outer code. The $M$-qubit code $\mathcal{C}^{out} = \{E^{out}, D^{out}\}$ with $E^{out}$ the encoding procedure and $D^{out}$ the decoding procedure.
- An inner code. The $N$-qubit code $\mathcal{C}^{in} = \{E^{in}, D^{in}\}$ with $E^{in}$ the encoding procedure and $D^{in}$ the decoding procedure.

The qubit carrying the information is first encoded with the procedure $E^{out}$ to $M$ qubits, then each of the $M$ qubits is encoded again to a block of $N$ qubits, this time using the procedure $E^{in}$. After passing through a noisy channel the $MN$ qubits are first decoded block by block using the procedure $D^{in}$, and then the decoding procedure $D^{out}$ is applied. The density matrix of the system evolves as follows:

$$\rho_i^1 \mapsto \rho_\alpha^{(M)} \mapsto \rho_\beta^{(MN)} \mapsto \rho_\gamma^{(MN)} \mapsto \rho_\delta^{(M)} \mapsto \rho_f^1$$

with $\rho_i^{(1)}$ the density matrix of the original qubit, $\rho_\alpha^{(M)}$ the density matrix after the outer encoding procedure, $E^{out}$, $\rho_\beta^{(MN)}$ the density matrix after the inner encoding procedure, $E^{in}$, $\rho_\gamma^{(MN)}$ the density matrix before the inner decoding procedure,$D^{in}$, $\rho_\delta^{(M)}$ the density matrix before the outer decoding procedure $D^{out}$, and $\rho_f^1$ the density matrix of the qubit after the error correction, Figure 137.

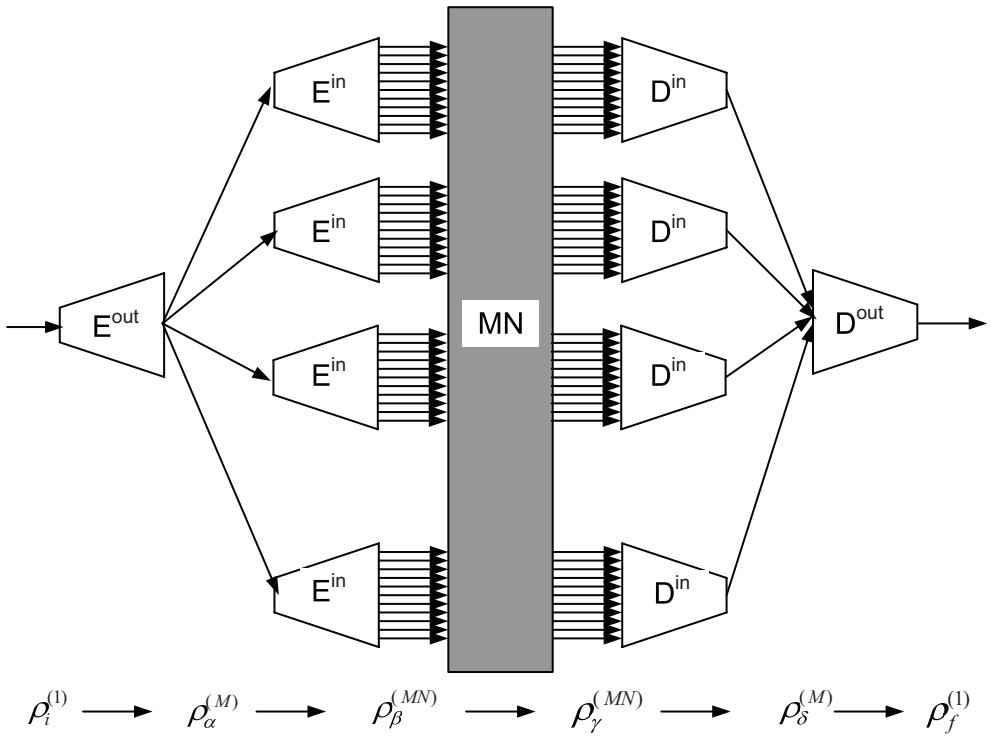



Figure 137: The encoding and decoding of a concatenated code, $\mathcal{C} = \{\mathcal{C}^{out}, \mathcal{C}^{in}\}$. The qubit carrying the information is first encoded with the procedure $E^{out}$ to $M$ qubits, then each of the $M$ qubits is encoded again to a block of $N$ qubits, this time using the procedure $E^{in}$. After passing through a noisy channel the $MN$ qubits are first decoded block by block using the procedure $D^{in}$, and then the decoding procedure $D^{out}$ is applied. The noise operator $\mathcal{N}$ may represent a noisy communication channel, or noisy quantum circuits.

## 5.18 Quantum Convolutional and Quantum Tail-Biting Codes

Quantum convolutional codes (QCC) were introduced by Ollivier and Tillich [315]. More recently, Forney, Grassl, and Guha [156] showed how to extend the stabilizer formalism and to construct convolutional and tail biting coded from classical self-dual codes with rates $R = 1/n$ over $GF(4)$. The authors show that quantum convolutional codes compare favorably with quantum block codes; they have a better code rate and a lower decoding complexity thus, a better performance.

This section builds up on the concepts related to classical convolutional and tail-biting codes, therefore a review of Section 4.16 is recommended. The $D$-transform is used in the literature of quantum convolutional and quantum tail-biting codes [156], [315]; we introduce the D-transform and use it to discuss convolutional encoders based on linear time-invariant systems. Then we discuss a finite state representation of convolutional codes and review codes over $GF(4)$.

**D-transform.** We first review a few concepts related to generating functions and the $Z$-transform. A generating function is a formal power series whose coefficients encode information about a sequence $a_n$ indexed by natural numbers

$$G(a_n; x) = a_n x^n.$$

For example,

$$G(n^2; x) = \sum_{n=0}^{\infty} n^2 x^2 = \frac{x(x+1)}{(1-x)^3}.$$

Informally, a generating function is a "clothline on which we hang up a sequence of numbers to display."

The Z-transform converts a discrete time-domain signal, a sequence of real numbers $a_n$, to a complex frequency-domain representation A(z)

$$A(z) = \mathcal{Z}(a_n) = \sum_{n=0}^{\infty} a_n z^{-n}.$$

The $Z$-transform allows us to compute the response of linear circuits. A linear circuit is characterized by a transfer function. When the circuit has one input $a_k$ and one output $c_k$ and the transfer function, also called an *impulse response*, is $h_k$ the output sequence is

$$c_k = \sum_{i>0} h_i a_{k-i}.$$

Let $A(z)$, $H(z)$, and $C(z)$ be the $Z$-transforms of $a_k, h_k$ and $c_k$ respectively; then

$$C(z) = H(z) \cdot A(z).$$

The convolution giving the response of the linear circuit in the time domain is transformed to a multiplication in the $Z$-domain.

Recall that a convolutional code is characterized by a triplet $[n, k, M]$ with $n$ the length of the output block $c$, $k$ the length of the input block $a$, and $M$ the memory of the code. A more general encoding rule for a convolutional code is

$$c_k = \sum_{i \geq 0} G_i a_{k-i},$$

where $G_i$ are $k \times n$ matrices; for $k = 1$ this expression coincides with the encoding rule given in Section 4.16, $c_k = \sum_{i \geq 0} g_i a_{k-i}$ with $g_i$ the components of the generating vector $g$.

The analogy with linear circuits is clear. Now, instead of the indeterminate $1/z$ used in case of the $Z$-transform we shall use the indeterminate $D$. If $a(D), G(D)$, and $c(D)$ are the $D$-transforms of the input, the generating matrix/vector, and the output, respectively, then the encoding rule can be written as a product of $D$-transforms

$$c(D) = a(D) \cdot G(D).$$

The $D$-transform allows us to describe a convolutional code in a more compact format. We shall see that $D$ can be viewed as a "delay" operator. For example, the expression $c_k = a_k + a_{k-1} + a_{k-2}$ encoding the fact that $c_k$, the output at time $k$, is a function of the input at times $k$, $k-1$ and $k-2$ is represented by the polynomial $(1 + D + D^2)$; when $c_k = a_k + a_{k-2}$ the polynomial is $(1 + D^2)$.

Consider a rate $R = 1/2$ convolutional code with generating vector $g = (11 \mid 10 \mid 11)$. In this case each input bit is encoded as two bits. The first bit is generated function of the first bit of each pair of the generating vector $(11 \mid 10 \mid 11)$, namely 111; the second output bit is generated depending on the second bit of each pair, namely 101. The $D$-transforms of the two sequences are

$$(111) \mapsto (1 + D + D^2) \quad \text{and} \quad (101) \mapsto (1 + D^2).$$

Thus $G(D) = [(1 + D + D^2), (1 + D^2)]$.

**Linear time-invariant (LTI) systems over finite fields.** The encoders for convolutional codes are linear-time invariant systems over a finite field $GF(q)$ and are characterized by an *impulse response* given by the $D$-transform $G(D)$.

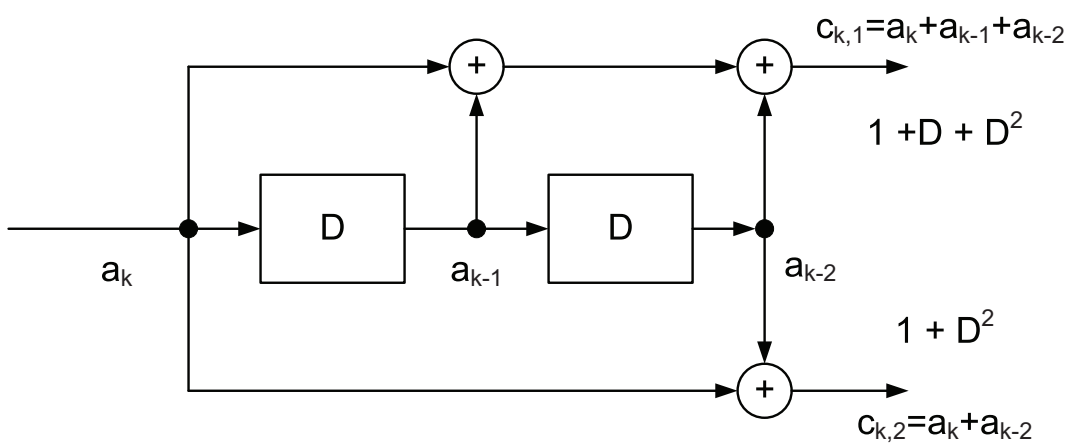



Figure 138: An encoder for a rate $R = 1/2$ block convolutional code. The encoder generates two output bits for each input bit. The LTI system has two delay/memory elements and three adders. The generating vector is $g = (11.10.11)$ thus the upper output bit stream is generated according to the sequence is 111 (the first bit of each pair of bits in $g$) thus, $c_{k,1} = (a_k + a_{k-1} + a_{k-2})$ and the corresponding polynomial is $(1 + D + D^2)$; the lower output bit stream is generated according to the sequence 101 (the second bit of each pair of bits in $g$) thus, $c_{k,2} = (a_k + a_{k-2})$ and the corresponding polynomial is $(1 + D^2)$.

LTI systems are implemented using adders, multipliers and delay/memory elements over $GF(q)$; they are linear systems because the output is a linear function of the input, the adders implement the addition in $GF(q)$, the multipliers implement the multiplication in $GF(q)$, and the delay/memory is also a linear operation. The systems are time-invariant because the relation between input and output is time-invariant, the output of a convolutional code with generating vector $g = (g_1, g_2, \ldots g_s)$ at time $k$ is $c_k = \sum_{i \geq 0} g_i a_{k-i} \quad \forall k \in \mathbb{Z}$.

Figure 138 shows an encoder for the rate $R = 1/2$ block convolutional code in the previous example. The two output sequences are: $(a_k + a_{k-1} + a_{k-2})$ and $(a_k + a_{k-2})$ thus the encoder is characterized by the *impulse response* $G(D) = [(1 + D + D^2), (1 + D^2)]$.

It is easy to see that this encoder generates a code capable to correct 2 errors. Indeed, the distance of the code is $d = 5$. To show that, we observe that the minimum weight code is generated by the input sequence $\ldots, 0, 0, 1, 0, 0, \ldots$; in this case the sequence of the output bits is 111 for the upper bit stream and 101 for the lower bit stream. Thus the minimum weight codeword is 5.

The encoder for a rate $R = 1/n$ convolutional code is characterized by $n$ polynomials in $D$, one for each of the $n$ output bits. The degree of each polynomial of a code with a generating vector $g = (g_0, g_1, \ldots, g_s)$ can be at most $s$.

Consider now an LTI system with $\mu$ delay/memory elements over $GF(q)$ and characterize the state of the system by the contents of the $\mu$ memory elements. Then the system can by in one of $\mu^q$ states. It is easy to see that the impulse response $G(D)$ must be a rational function. Indeed, the *autonomous* behavior corresponding to zero input, must be periodic after time zero if the impulse response at time zero is $g_0(D) = 1$. Thus $G(D)$ must be periodic, i.e., a rational function, $G(D) = G_1(D)/G_2(D)$.

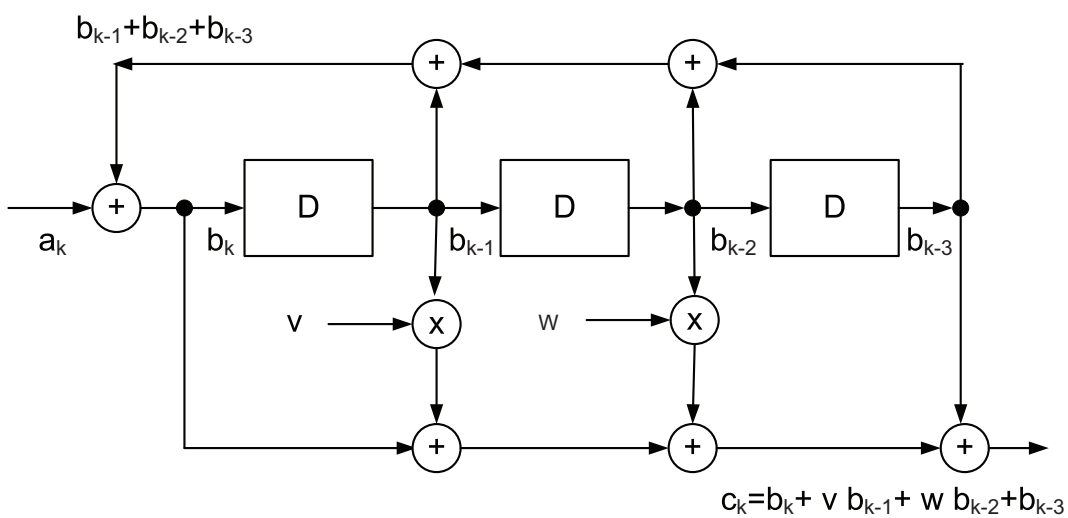



Figure 139: An encoder with a rational impulse response over $GF(2)$. The impulse response is infinite and the realization of the LTI system involves feedback. There are three memory elements (thus, the system has $3^2$ states), six adders and two multipliers with $v, w \in GF(2)$.

Figure 139 shows an encoder with a rational impulse response over $GF(2)$. In this case the output and the feedback are

$$G_1(D) = 1 + vD + wD^2 + D^3 \quad \text{and} \quad G_2(D) = 1 - D - D^2 - D^{-3}.$$

The impulse response is

$$G(D) = \frac{G_1(D)}{G_2(D)} = \frac{1 + vD + wD^2 + D^3}{1 - D - D^2 - D^{-3}}.$$

**Finite-state representation of convolutional codes.** The encoder of a convolutional code over $GF(q)$ is a finite-state machine. A state is labelled by the symbols stored by the memory elements, thus a system with $M$ memory elements has $M^q$ states. The transitions between states are labelled by the output streams. For example, the encoder for a rate $R = 1/2$ block convolutional code over $GF(2)$ in Figure 138 has $2^2$ states labelled $00, 01, 10, 11$, as shown in Figure 140. The trellis diagram in Figure 140 (b) conveys the same information as the traditional state transition diagram in 140 (a) and, in addition, it allows us to follow the path of an encoded sequence.

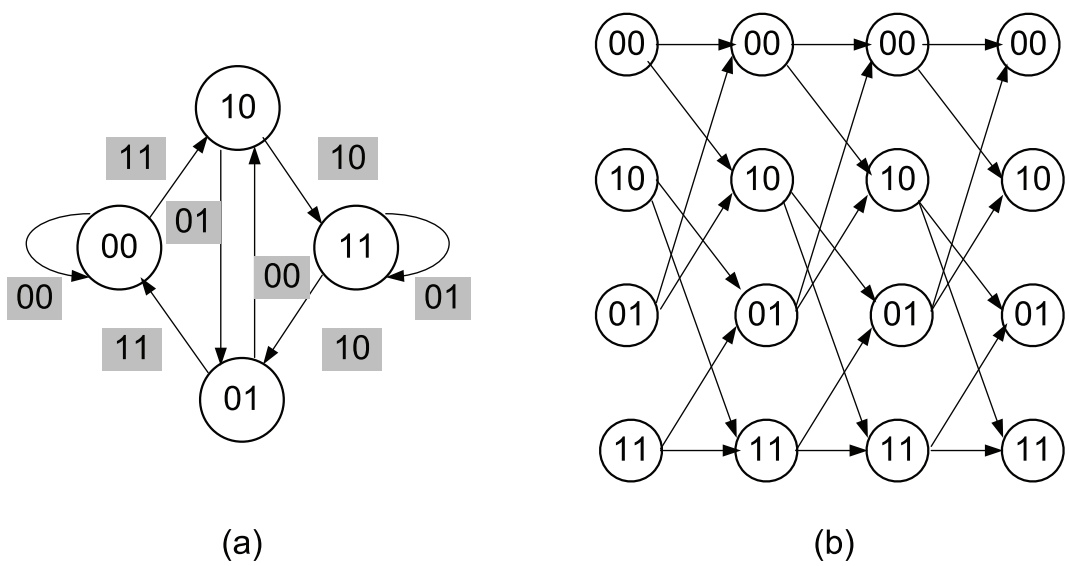



Figure 140: The states and transitions for the rate $R = 1/2$ block convolutional code over $GF(2)$ in Figure 138. A state is characterized by the contents of the two memory elements and a transition is labelled by the two output symbols (a) State transition diagram. (b) The trellis diagram of the encoder; the labels on the transition are omitted.

**Linear codes over $GF(4)$.** In this section we prepare the ground to extend the stabilizer formalism and to construct convolutional and tail biting coded from classical self-dual codes with rates $R = 1/n$ over $GF(4)$.

Recall that $\mathcal{P}$, the set of Pauli operators is not an Abelian group, but the set $\Pi$ of equivalence classes of Pauli matrices, also called the *projective Pauli group* is an Abelian group. The multiplication tables of $\mathcal{P}$ and $\Pi$ are

| $\times$ | $\mathbb{I}$ | $X$ | $Y$ | $Z$ |
|---|---|---|---|---|
| $\mathbb{I}$ | $\mathbb{I}$ | $X$ | $Y$ | $Z$ |
| $X$ | $X$ | $\mathbb{I}$ | $iZ$ | $-iY$ |
| $Y$ | $Y$ | $-iZ$ | $\mathbb{I}$ | $iX$ |
| $Z$ | $Z$ | $iY$ | $-iX$ | $\mathbb{I}$ |

and

| $\times$ | $[\mathbb{I}]$ | $[X]$ | $[Y]$ | $[Z]$ |
|---|---|---|---|---|
| $[\mathbb{I}]$ | $[\mathbb{I}]$ | $[X]$ | $[Y]$ | $[Z]$ |
| $[X]$ | $[X]$ | $[\mathbb{I}]$ | $[Z]$ | $[Y]$ |
| $[Y]$ | $[Y]$ | $[Z]$ | $[\mathbb{I}]$ | $[X]$ |
| $[Z]$ | $[Z]$ | $[Y]$ | $[X]$ | $[\mathbb{I}]$ |

.

We now consider the additive group with four elements, $GF(4) = \{0, 1, \omega, \bar{\omega}\}$. The addition and multiplication tables of $GF(4)$ are, respectively,

| $+$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
|---|---|---|---|---|
| $0$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
| $1$ | $1$ | $0$ | $\bar{\omega}$ | $\omega$ |
| $\omega$ | $\omega$ | $\bar{\omega}$ | $0$ | $1$ |
| $\bar{\omega}$ | $\bar{\omega}$ | $\omega$ | $1$ | $0$ |

and

| $\times$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
| $\omega$ | $0$ | $\omega$ | $\bar{\omega}$ | $1$ |
| $\bar{\omega}$ | $0$ | $\bar{\omega}$ | $1$ | $\omega$ |

.

Group isomorphisms allow us to discover and express properties of the elements of a group relative to another group. Recall that the elements of two isomorphic groups are in an one-to-one correspondence and respect the group operations. More precisely, two groups $G = (S, \perp)$ and $G' = (S', \top)$ are isomorphic if there is a one-to-one mapping $f : G \mapsto G'$ such that

$$\forall x_i, x_j \in G \quad f(x_i \perp x_j) = f(x_i) \top f(x_j).$$

The projective Pauli group, $\Pi$, is isomorphic with the additive group $GF(4)$, $L : \Pi \mapsto GF(4)$; this implies that

$$L([A]) + L([B]) = L([A \times B]), \qquad \forall([A], [B]) \in [\Pi].$$

The correspondence between the respective elements of $\Pi$ and $GF(4)$ is summarized in the following table

| $\Pi$ | $GF(4)$ |
|---|---|
| $[\mathbb{I}]$ | $0$ |
| $[X]$ | $\omega$ |
| $[Y]$ | $1$ |
| $[Z]$ | $\bar{\omega}$ |

.

Let us now consider tensor products of $n$ Pauli operators

$$A = A_1 \otimes A_2 \ldots \otimes A_i \ldots \otimes A_n \quad \text{and} \quad B = B_1 \otimes B_2 \ldots \otimes B_i \ldots \otimes B_n$$

with $A_i, B_i \in \{\mathbb{I}, X, Y, Z\}$.

The mapping $L$ of such tensor products induces a vector space $\mathcal{V} = \{a, b, \ldots\}$ over $GF(4)$ where

$$a = L(A), b = L(B), \quad a = (a_1, a_2, \ldots, a_n), \; b = (b_1, b_2, \ldots, b_n), \quad \text{and} \quad a_i = L(A_i), \; b_j = L(B_j).$$

It is easy to show that

$$L([A]) + L([B]) = L([A \times B])$$

for any $[A], [B] \in \mathcal{P}$. For example, consider $n = 1$, $A_1 = X$, and $B_1 = Y$; we have to show that

$$L([X]) + L([Y]) = \omega + 1 \quad \text{and} \quad L([X] \times [Y]) = \bar{\omega}.$$

From the correspondence table it follows that

$$L([X]) = \omega, \quad L([Y]) = 1, \quad \Longrightarrow \quad L([X]) + L([Y]) = \omega + 1 = \bar{\omega}.$$

The last equality follows from the addition table in $GF(4)$; then

$$[X] \times [Y] = [Z], \quad L([Z]) = \bar{\omega} \quad \Longrightarrow \quad L([X] \times [Y]) = L([Z]) = \bar{\omega}.$$

Now we introduce the Hermitian inner product and the trace inner product in $\mathcal{V}$ and use them to study the commutativity/anticommutativity properties of elements of the Pauli group. The *Hermitian inner product* $\langle a, b\rangle$ and the *trace inner product* $\mathrm{tr}\langle a, b\rangle\rangle$ of two elements $a, b \in GF(4)$ are defined as

| $\langle , \rangle$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
| $\omega$ | $0$ | $\bar{\omega}$ | $1$ | $\omega$ |
| $\bar{\omega}$ | $0$ | $\omega$ | $\bar{\omega}$ | $1$ |

and

| $\mathrm{tr}\langle , \rangle$ | $0$ | $1$ | $\omega$ | $\bar{\omega}$ |
|---|---|---|---|---|
| $0$ | $0$ | $0$ | $0$ | $0$ |
| $1$ | $0$ | $0$ | $1$ | $1$ |
| $\omega$ | $0$ | $1$ | $0$ | $1$ |
| $\bar{\omega}$ | $0$ | $1$ | $1$ | $0$ |

Note that $\langle a, b\rangle = a^\dagger b$ with $a^\dagger$ the adjoint of $a$. These definitions can be extended to $a, b \in \mathcal{V}$.

The commutativity properties of matrices in the Pauli group $\mathcal{P}$ can now be established based on the *Hermitian inner product* and the *trace inner product* of two elements $a, b \in \mathcal{V}$.

Thus, $\forall A, B \in \Pi$ we have

$$A \cdot B = (-1)^{\mathrm{tr}\langle L(A), L(B)\rangle} B \cdot A.$$

We conclude that *two matrices* $A, B \in \Pi$ *commute when* $\mathrm{tr}\langle L(A), L(B)\rangle = 0$ *and anticommute when* $\mathrm{tr}\langle L(A), L(B)\rangle = 1$.

The Hermitian inner product in $\mathcal{V}$ is zero when the trace inner product is zero; indeed, if $(a, b) \in \mathcal{V}$ and $\alpha \in GF(4)$ then

$$\mathrm{tr}\langle \alpha \times a, b\rangle = \mathrm{tr}[\alpha \times \langle a, b\rangle] = \alpha \times \mathrm{tr}[\langle a, b\rangle].$$

It follows that two matrices $A, B \in \Pi$ commute if and only if the Hermitian inner product of their images in $\mathcal{V}$ is equal to zero.

To prove these properties we have to consider all pairs of elements $A, B \in G$ and show that the property holds; for example, consider the case $n = 1$, $A_1 = X$ and $B_1 = Y$

$$A_1 \cdot B_1 = X \cdot Y = iZ \quad \text{and} \quad B_1 \cdot A_1 = Y \cdot X = -iZ.$$

From the correspondence table we have

$$L(X) = \omega, \quad L(Y) = 1.$$

From the definition of the trace inner product $\mathrm{tr}\langle \omega, 1\rangle = 1$. We have thus verified that $X$ and $Y$ anticommute by showing that the trace inner product $\mathrm{tr}\langle L(X), L(Y)\rangle = 1$. These results allow us to extend the stabilizer formalism and define quantum convolutional and quantum tail-biting codes.

**Stabilizer codes from codes over $GF(4)$.** Now it is a good time to review the basic concepts regarding stabilizer codes discussed in Section 5.12. Consider an $[n, k]$ code in $(\mathcal{H}_2)^{\otimes n}$ stabilized by the set $S = \{M_1, M_2, \ldots, M_{n-k}\}$; the $(n-k)$ generators $M_i$ of the stabilizer $\mathcal{S}$ are tensor products of $n$ Pauli matrices. Each $M_i$ has two eigenvalues, $\pm 1$, and has two corresponding eigenspaces of dimension $2^{n-1}$. There are $(n-k)$ generators thus, there are $2^{n-k}$ orthogonal eigenspaces, each of dimension $2^k$. The quantum code is the $2^k$-dimensional eigenspace such that all $(n-k)$ generators have eigenvalues equal to $+1$.

Table 30 summarizes the correspondence between stabilizer codes in the Hilbert space $(\mathcal{H}_2)^{\otimes n}$ and linear codes in $\mathcal{V}$ induced by the mapping $L$ and its inverse $L^{-1}$ introduced in [156].

Error correction for a stabilizer code is done as follows. We measure $(n-k)$ generators $M_j$ simultaneously and map the eigenvalues of the error operator $E$ applied to the current state $\mid \psi\rangle$ to a binary $(n-k)$ tuple $s = \{s_1, s_2, \ldots, s_{n-k}\}$ called the syndrome. Bit $s_i$ of the syndrome depends only on the eigenvalue of $E$, $(+1) \mapsto 0$, $(-1) \mapsto 1$.

We can exploit the mapping discussed in this section to construct an $[n, k, d]$ stabilizer code from classical codes over $GF(4)$ when $(n-k)$ is even. Given a self-dual $[n, (n-k)/2, d]$ linear code $\mathcal{C}$ and its $[n, (n+k)/2, d]$ dual, $\mathcal{C}^\perp$, the resulting $[n, k, d]$ stabilizer code will be able to correct $\lfloor (d-1)/2 \rfloor$ errors. For example, the $[5, 2, 4]$ doubly-extended Reed-Solomon code generated by:

$$\begin{matrix} 0 & \bar{\omega} & \omega & \omega & \bar{\omega} \\ \bar{\omega} & 0 & \bar{\omega} & \omega & \omega \end{matrix}$$

could be used to construct a $[5, 1, 3]$ single-error correcting stabilizer code.

**Convolutional stabilizer codes.** A convolutional quantum code $\mathcal{Q}$ with minimum distance $d$ can be constructed from a classical convolutional $[n, (n-k)/2, d]$ linear code $\mathcal{C}$ and its dual $\mathcal{C}^\perp$ an $[n, (n+k)/2, d]$. The rates of $\mathcal{C}$, $\mathcal{C}^\perp$ and $\mathcal{Q}$ are respectively (see Section 4.5)

$$R_{\mathcal{C}} = \frac{n-k}{2n}, \qquad R_{\mathcal{C}^\perp} = \frac{n+k}{2n}, \qquad \text{and} \qquad R_{\mathcal{Q}} = \frac{k}{n}.$$

Consider for example $\mathcal{C}$, a rate $R = 1/3$ convolutional code over $GF(4)$ with the $D$-transform

$$m(D) = (1 + D, 1 + \omega D, 1 + \bar{\omega} D)$$

A generator sequence is $m_j = (\ldots \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid 000 \mid \ldots)$ and the generators are

$$\begin{matrix} \ldots \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid 000 \mid 000 \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid 000 \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 000 \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid \ldots \end{matrix}.$$

Table 30: The correspondence between stabilizer codes in the Hilbert space $\mathcal{H}_2^n$ and linear codes in $\mathcal{V}$ induced by mapping $L$.

| | | |
|---|---|---|
| $(\mathcal{H}_2)^{\otimes n}$ | $\overset{\text{L}}{\longrightarrow}$ | $\mathcal{V}$ |
| $M_i \in \mathcal{P}, 1 \leq i \leq (n-k)$<br>$M_i$ tensor product of $n$ Pauli operators<br>generator of the stabilizer group $\mathcal{S}$ | | $m_i = L(M_i),\ 1 \leq i \leq (n-k)$<br>$m_i$ vector with elements from $GF(4)$<br>generator in $\mathcal{V}$ |
| $M_i \cdot M_j = M_j \cdot M_i$<br>($M_i$ and $M_j$ commute) | | $\text{tr}\langle m_i, m_j\rangle = 0$<br>(Hermitian inner product $\langle m_i, m_j\rangle = 0$ ) |
| $S = \{M_1, M_2, \ldots, M_{n-k}\}$<br>$S \subset \Pi$ is an Abelian subgroup | | $\Sigma = \{m_1, m_2, \ldots, m_{n-k}\}$<br>$\Sigma \subset \mathcal{V}$ generates a classical code |
| $[n, k]$ stabilizer code<br>subspace of $\mathcal{H}_2^n$ stabilized by $\mathcal{S}$ | | $L(S)$ set of linear combinations $\omega m_i$ and $\bar{\omega} m_i$<br>$[n, (n-k)/2]$ block linear code over $GF(4)$<br>$\langle m_i, m_j\rangle = 0 \implies L(S)$ is self-dual<br>under the Hermitian inner product in $\mathcal{V}$ |
| $N(S)$<br>normalizer group of stabilizer $\mathcal{S}$ | $L^{-1}$<br>$\longleftarrow$ | $L(S)^{\perp}$ classical code, the dual of $L(S)$<br>$[n, (n+k)/2]$ block linear code over $GF(4)$ |
| error pattern $E = \sigma_a^{\otimes n}, a = \{I, x, y, z\}$<br>syndrome $s = \{s_1, s_2, \ldots, s_{n-k}\}$ | | $e = L(E)$<br>syndrome $\gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_{n-k}\},\ \gamma_j = L(s_j)$<br>$\gamma_j = \text{tr}\langle e, m_j\rangle$ |

$\mathcal{C}$ is self dual, the generators are orthogonal $m_j \times m_j = 0$, according to the multiplication table in $GF(4)$. Then $\mathcal{C}^{\perp}$ is a rate $R = 2/3$ convolutional code generated by

$$m'(D) = (\bar{\omega}, \omega, 1).$$

It follows that $m'_j = (\ldots \mid 000 \mid 111 \mid \bar{\omega}\omega 1 \mid 000 \mid \ldots)$. The generators of $\mathcal{C}^{\perp}$ are

$$\begin{array}{l} \ldots \mid 000 \mid 111 \mid \bar{\omega}\omega 1 \mid 000 \mid 000 \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 111 \mid \bar{\omega}\omega 1 \mid 000 \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 000 \mid 111 \mid \bar{\omega}\omega 1 \mid 000 \mid \ldots \\ \ldots \mid 000 \mid 000 \mid 000 \mid 000 \mid 111 \mid \bar{\omega}\omega 1 \mid \ldots \end{array}.$$

A generator $M_j$ of the stabilizer $\mathcal{S}$ can be determined using the mapping between $GF(4)$ and $\Pi$

$$M = F^{-1}[\ldots \mid 000 \mid 111 \mid 1\omega\bar{\omega} \mid 000 \mid \ldots] = (\ldots \mid \mathbb{III} \mid XXX \mid XZY \mid \mathbb{III} \mid \ldots].$$

All generators of $\mathcal{S}$ can be obtained by a cyclic shift of $M$ to the right with an integral number of blocks

$$\begin{array}{l} \ldots \mid \mathbb{III} \mid XXX \mid XZY \mid \mathbb{III} \mid \mathbb{III} \mid \mathbb{III} \mid \ldots \\ \ldots \mid \mathbb{III} \mid \mathbb{III} \mid XXX \mid XZY \mid \mathbb{III} \mid \mathbb{III} \mid \ldots \\ \ldots \mid \mathbb{III} \mid \mathbb{III} \mid \mathbb{III} \mid XXX \mid XZY \mid \mathbb{III} \mid \ldots \\ \ldots \mid \mathbb{III} \mid \mathbb{III} \mid \mathbb{III} \mid \mathbb{III} \mid XXX \mid XZY \mid \ldots \end{array}.$$

Another quantum convolutional code is described in [315]. The generators of the stabilizer group are

$$\begin{array}{lcl} M_0 & = & X \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \ldots \\ M_1 & = & Z \otimes X \otimes X \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \ldots \\ M_2 & = & \mathbb{I} \otimes Z \otimes X \otimes X \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \ldots \\ M_3 & = & \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes X \otimes X \otimes Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \ldots \\ M_4 & = & \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes X \otimes X \otimes Z \otimes \mathbb{I} \otimes \ldots \\ \vdots & & \\ M_{4i+j} & = & \mathbb{I}^{\otimes 5i} \otimes M_j, \quad 0 < i, \ 1 \leq j \leq 4 \\ \vdots & & \\ M_\infty & = & \ldots \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes Z \otimes X. \end{array}$$

The code so defined has a particular structure: all generators, other than $M_0$ and $M_\infty$, can be grouped together to sets of four; each set acts on a set of seven consecutive qubits and each set has a fixed ovelap of two qubits with the set imediately before and after it.

**Block codes and tail-biting codes.** Two methods to make a block code out of a convolutional code, termination and tail-biting, are discussed in [156].

When using termination, we truncate the set of generators $\Sigma$ of $\mathcal{C}$ to create a block code $\mathcal{B}$. For example if we start with the $R = 2/3$ code $\mathcal{C}^\perp$ from the previous example we can construct a $[9, 5, 3]$ terminated block code $\mathcal{B}^\perp$ using the generator matrix

$$G_t = \begin{pmatrix} \mid \bar{\omega}\omega 1 & \mid 000 & \mid 000 \mid \\ \mid 111 & \mid 1\omega\bar{\omega} & \mid 000 \mid \\ \mid 000 & \mid \bar{\omega}\omega 1 & \mid 000 \mid \\ \mid 000 & \mid 111 & \mid 1\omega\bar{\omega} \mid \\ \mid 000 & \mid 000 & \mid \bar{\omega}\omega 1 \mid \end{pmatrix}.$$

The $[9, 4]$ orthogonal block code $\mathcal{B}$ is generated by the truncated generators of $\mathcal{C}$ which are not self orthogonal; $\mathcal{B}$ is not self-orthogonal. The generator matrix is

$$G'_t = \begin{pmatrix} \mid 1\omega\bar{\omega} & \mid 000 & \mid 000 \mid \\ \mid 111 & \mid 1\omega\bar{\omega} & \mid 000 \mid \\ \mid 000 & \mid 111 & \mid 1\omega\bar{\omega} \mid \\ \mid 000 & \mid 000 & \mid 111 \mid \end{pmatrix}.$$

We can also construct a block code $\mathcal{B}$ from a concatenated code $\mathcal{C}$ by tail-biting by limiting the generators to those whose "starting time" lies in a given interval. We continue with the example from [155] and use the rate $R = 2/3$ code $\mathcal{C}^{\perp}$. The generator matrix in this case is

$$G_{tb} = \begin{pmatrix} |\,\bar{\omega}\omega 1 & |\,000 & |\,000\,| \\ |\,111 & |\,1\omega\bar{\omega} & |\,000\,| \\ |\,000 & |\,\bar{\omega}\omega 1 & |\,000\,| \\ |\,000 & |\,111 & |\,1\omega\bar{\omega}\,| \\ |\,000 & |\,000 & |\,\bar{\omega}\omega 1\,| \\ |\,1\omega\bar{\omega} & |\,000 & |\,111\,| \end{pmatrix}.$$

## 5.19 Correction of Time-Correlated Quantum Errors

The quantum threshold theorem mentioned in Section 1.14 is based on the assumption that the quantum noise is Markovian and establishes the foundation for fault-tolerant computations in environments where spatial and temporal correlations decay exponentially. The "classical" theory of quantum error correcting codes is based on the independent qubit decoherence (IQD) model which makes several assumptions:

- Errors are uncorrelated in space. An error affecting qubit $i$ of an $n$-qubit register does not affect other qubits of the register. An error affecting one qubit is equally likely to be due to an interaction with the environment described by $X, Y$, or $Z$ operators.
- Errors are uncorrelated in time. An error affecting qubit $i$ at time $t$ and corrected at time $t + \Delta$ will have no further effect either on qubit $i$ or on other qubits of the register.
- The error rate is constant. Moreover, whenever the probability of a qubit to be in error, $\epsilon$, is very small, $\epsilon \leq 10^{-4}$, then the probability of $t$ or more errors is very small indeed, $\mathcal{O}(\epsilon^t)$, and we ignore such a case.

Different physical implementations of quantum devices reveal that the interactions of the qubits with the environment are more complex and force us to consider spatially- as well as, time-correlated errors. For example, if the qubits of an $n$-qubit register are confined to a `3D` structure, an error affecting one qubit will propagate to the qubits in a volume centered around the qubit in error.

Two major questions pertinent to error correction of correlated errors need to be addressed: first, can the quantum threshold theorem be extended for non-Markovian noise, and second, what are the practical means to deal with correlated errors. In this section we only address the second question and restrict our discussion on how to deal with time-correlated errors.

Correlated quantum errors could cause two or more errors to occur during an error correction cycle. There are two obvious approaches to deal with this problem:

1. Design a code capable to correct $(e_u + e_c)$ errors, where $e_u$ is the number of uncorrelated errors and $e_c$ the expected number of time-correlated errors, or
2. Use the classical information regarding past errors and quantum error correcting codes capable to correct a single error.

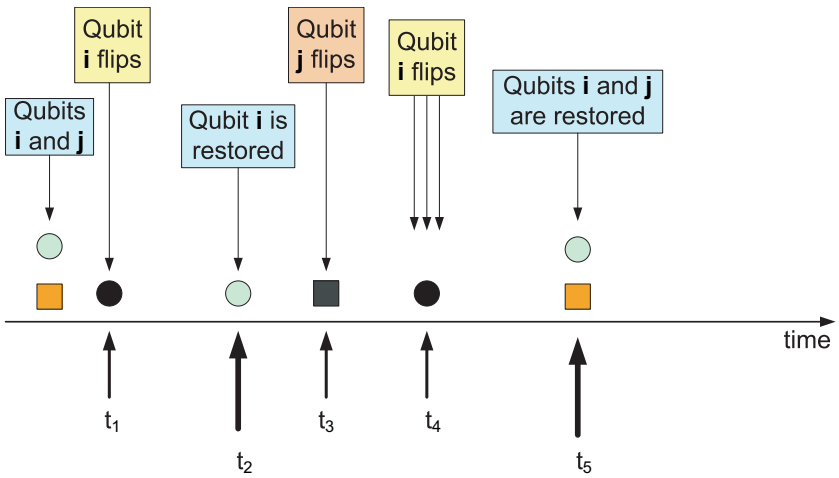



Figure 141: Time-correlated quantum errors. We follow the evolution in time of two qubits, $i$ and $j$ for two error correction cycles. The first error correction cycle starts at time $t_2$ and the second at time $t_5$. At time $t_1$ qubit $i$ is affected by decoherence and flips; at time $t_2$ it is flipped back to its original state during the first error correction step; at time $t_3$ qubit $j$ is affected by decoherence and it is flipped; at time $t_4$ the correlation effect discussed in this section affects qubit $i$ and flips it to an error state; the second error correction occurs at time $t_5$ and qubits $i$ and $j$ are corrected.

Spatially-correlated errors and means to deal with the spatial noise are analyzed in recent papers [7, 101, 242]. Time-correlated cause errors occur after an error correction cycle and could affect the qubit restored to its original state after an error, or a different qubit. Figure 141 illustrates the evolution in time of two qubits, $i$ and $j$ for two error correction cycles and the effect of time-correlation. We expect that the correlations in the quantum system decay in time, and the latest error will influence the system most. A recent paper on the decoherence in a spin-boson model assumes that the qubits are perfect, thus the only errors are due to de-phasing and considers a linear coupling to an ohmic bath [311]; then the Hamiltonian of the model can be written as

$$H = \frac{v_b}{2} \int_{-\infty}^{+\infty} dx[\partial_x \phi(x)]^2 + [\Pi(x)]^2 + \frac{\lambda}{2} \sum_n \partial_x \phi(n) \sigma_n^z$$

with $\phi$ and $\Pi = \partial_x \theta$ canonic conjugate variables and $\sigma_n^z$ the Pauli operators acting in the Hilbert space of the qubits. $v_b$ is the velocity of the bosonic excitations and the units are such that $\hbar = 1$. The exact time evolution between gates of a qubit, can be expressed as the product of two vertex operators of the free bosonic theory

$$U_n(t, 0) = \exp[i\sqrt{\frac{\pi}{2}}\lambda(\theta(n, t) - \theta(n, 0))\sigma_n^z]$$

with $\lambda$ the qubits bosonic coupling strength; this model is then applied to the three qubit Steane's code to calculate the probability of having errors in quantum error correction cycles starting at times $t_1$ and $t_2$. The conclusion of [311] is that the probability of errors consists

of two terms; the first is the uncorrelated probability and the second is the contribution due to correlation between errors in different cycles ($\Delta$ is the error correcting cycle)

$$P \approx (\frac{\epsilon}{2})^2 + \frac{\lambda^4 \Delta^4}{8(t_1 - t_2)^4}.$$

This implies that the correlations in the quantum system decay in time, and the latest error will influence the system most.

An important question for the error model discussed in [311] is the independence of phase- and bit-flip errors, namely, can a phase-flip error be correlated to a posterior bit-flip error? The answer will depend on the particular environment one tries to model. For solid-state qubits, these errors are caused by different bath modes which may not efficiently communicate with each other (such as for phonon-induced tunnelling and bias fluctuations in quantum dot charge qubits, which cause bit- and phase-flip errors, respectively). In this case, errors of a different nature are expected to be very weakly correlated in time and space while correlated errors are mostly of the same kind.

An algorithm based on the stabilizer formalism for perfect as well as non-perfect quantum codes is discussed in [279]. The quantum code has a codeword consisting of $n$ qubits and uses $k$ ancilla qubits for syndrome measurement; the algorithm makes the standard assumptions for quantum error correction:

- Quantum gates are perfect and operate much faster than the characteristic response of the environment;
- The states of the computer can be prepared with no errors.

It is also assumed that: (i) There is no spatial-correlation of errors, a qubit in error does not influence its neighbors; (ii) In each error correcting cycle, in addition to a new error $\mathbf{E}_a$ that occurs with a constant probability $\varepsilon_a$, a time-correlated error $\mathbf{E}_b$ may occur with probability $\varepsilon_b(t)$. As correlations decay in time, the qubit affected by error during the previous error correction cycle, has the highest probability to relapse.

QEC requires *non-demolition measurements of the error syndrome* in order to preserve the state of the physical qubits. In other words, a measurement of the probe (the ancilla qubits) should not influence the free motion of the signal system. The syndrome has to identify precisely the qubit(s) in error and the type of error(s). Thus, the qubits of the syndrome are either in the state with density matrix $\mid 0\rangle\langle 0 \mid$ or in the state with density matrix $\mid 1\rangle\langle 1 \mid$, which represent classical information. A quantum non-demolition measurement is used to construct the error syndrome $\Sigma_{current}$; after examining the syndrome $\Sigma_{current}$ an error correcting algorithm should be able to decide if:

1. No error has occurred; no action should be taken;
2. One "new" error, $\mathbf{E}_a$, has occurred; then we apply the corresponding Pauli transformation to the qubit in error;
3. Two or more errors have occurred. There are two distinct possibilities: (a) We have a "new" error as well as an "old" one, the time-correlated error; (b) There are two or more "new" errors. A quantum error correcting code capable of correcting a single error will fail in both cases.

It is rather hard to distinguish the last two possibilities. For perfect codes, the syndrome $S_{ab}$ corresponding to two errors, $\mathbf{E}_a$ and $\mathbf{E}_b$, is always identical to the syndrome $S_c$ for some single error $\mathbf{E}_c$. Thus, for perfect quantum codes the stabilizer formalism does not allow us to distinguish two errors from a single one; for a non-perfect code it is sometimes possible to distinguish the two syndromes and then using the knowledge regarding the time-correlated error it may be possible to correct both the "old" and the "new" error.

As we may recall from classical error correcting codes, burst error correcting codes are designed to deal with disturbances which affect a group of consecutive bits. Similar effects are noticeable in quantum computing and quantum communication when the correlation length of the source of decoherence is smaller than the separation in space of the qubits of a quantum register. In such cases, the independent qubit decoherence (IQD) model is no longer valid, the source of decoherence affects an entire group of qubits, e.g., those within a volume of radius $r$, in case of a 3D-structure. One method to deal with spatially-correlated quantum errors is to consider burst-error correcting codes.

The quantum error correcting codes discussed in Sections 5.1 to 5.19 are subspaces of a Hilbert space and are collectively known as subspace codes; a code $\mathcal{C}$ is a subspace of the Hilbert space $\mathcal{H}$ defined with the aid of the projector $\mathbf{\Pi}_{\mathcal{C}}$; the correction of error $\mathbf{E}$ produces the desired results only if $\mathbf{\Pi}_{\mathcal{C}}\mathbf{E}\mathbf{\Pi}_{\mathcal{C}} = \alpha_{\mathcal{C}}\mathbf{\Pi}_{\mathcal{C}}$ with $\alpha_{\mathcal{C}}$ a scalar. The detection of the error consists of the following steps:

1. Encode the information as a state $\mid \varphi\rangle \in \mathcal{C}$.
2. Allow the error to occur.
3. Carry out a measurement of the resulting state $\mathbf{E} \mid \varphi\rangle$
   - if the outcome is $\mathbf{\Pi}_{\mathcal{C}}\mathbf{E} \mid \varphi\rangle$ then the state is in $\mathcal{C}$ - accept;
   - if the outcome is $(\mathbf{I} - \mathbf{\Pi}_{\mathcal{C}})\mathbf{E} \mid \varphi\rangle$ then the state is in the orthogonal complement - reject.

We now switch our focus to subsystem codes; when using them, we limit the quantum information restoration process of traditional subspace codes and encode the information in a noiseless subsystem; the subsystem codes are said to be based on a passive quantum error correction strategy, while subspace codes are based on an active QEC strategy.

## 5.20 Quantum Error Correcting Codes as Subsystems

The search for systematic ways to bypass decoherence was initiated as a study of the effects of pure de-phasing; decoherence-free subspaces (DFS) capture the behavior of systems decoupled from the environment and with completely unitary evolution; for example, two qubits possessing identical interactions with the environment which do not decohere are discussed in [318]. Noiseless quantum codes were introduced in [467] and noiseless encoding in quantum dot arrays was reported in [468].

Emmanuel Knill, Raymond Laflamme, and Lorenza Viola introduced subsystem codes in a paper on the theory of quantum error correction for general noise [250]; a few years later

David Kribs, Raymond Laflamme, and David Poulin developed Operator Quantum Error Correcting (OQEC) codes [260] and Poulin proposed a stabilizer formalism for OQEC codes [338]. David Bacon introduced a class of subsystem codes referred to as Bacon-Shor codes [17, 18]. In this section and the next ones we follow the notations in the original papers to describe the evolution of ideas in the field of subsystem codes and we drop the bold notations for operators.

The basic idea of subsystem codes is to use algebraic methods to classify quantum errors and, implicitly, to describe quantum error correcting codes as tensor products of subspaces, or more accurately, as subsystems where the useful information is confined to one subsystem [250]. When the system consists of a number of qubits then each qubit is a subsystem; the approach of subsystem QEC is more general, it describes physical systems that cannot be canonically decomposed into qubits; for example, when the system consists of a number of photon modes each photon mode is a subsystem.

**Subspace versus subsystem quantum error correction.** A Hilbert space $\mathcal{H}$ can be decomposed as the direct sum of two subspaces $\mathcal{H} = \mathcal{H}^C \bigoplus \mathcal{H}^D$ or as the tensor product of two subspaces $\mathcal{H} = \mathcal{H}^C \otimes \mathcal{H}^D$; in case of a direct sum, $\dim(\mathcal{H}) = \dim(\mathcal{H}^C) + \dim(\mathcal{H}^D)$ while, $\dim(\mathcal{H}) = \dim(\mathcal{H}^C) \times \dim(\mathcal{H}^D)$ for the tensor product. The direct sum and the tensor product obey a distributive law thus, in the general case, we can *decompose $\mathcal{H}$ as a direct sum of multiple tensor products of Hilbert spaces.* It follows immediately that a decomposition of the Hilbert space $\mathcal{H}$ as a subspace $\mathcal{H}^E$ and a perpendicular subspace $(\mathcal{H}^E)^\perp$ is possible, with $(\mathcal{H}^E)^\perp$ expressed as a tensor product of two other subspaces $\mathcal{H}^C$ and $\mathcal{H}^D$

$$\mathcal{H} = (\mathcal{H}^E)^\perp \bigoplus \mathcal{H}^E \quad \text{or} \quad \mathcal{H} = (\mathcal{H}^C \otimes \mathcal{H}^D) \bigoplus \mathcal{H}^E.$$

The <u>subsystem encoding</u> for this decomposition requires the preparation of the quantum state

$$\rho = (\rho^C \otimes \rho^D) \oplus 0^E$$

and encoding the information into subsystem $\mathcal{C}$ described in the Hilbert space $\mathcal{H}^C$ for a fixed encoding into the subsystem $\mathcal{D}$ described in the Hilbert space $\mathcal{H}^D$; in this expression $\rho^C$ is the density matrix of the encoded quantum information, $\rho^D$ is the density matrix of the arbitrary quantum information encoded into the subsystem $\mathcal{D}$, and $0^E$ is the all zero matrix on the subspace $\mathcal{H}^E$.

$\mathcal{D}$ is referred to as a <u>gauge subsystem</u> and $E$ as a <u>logical subsystem</u>. Regardless of the nontrivial transformation applied to gauge subsystem $\mathcal{D}$, the information encoded into subsystem $\mathcal{C}$ is not affected; *the quantum error correction is restricted to the restoration of the information encoded into subsystem $\mathcal{C}$.*

The statement that subsystem QEC is the most general expression for quantum error correction is justified: when $\mathcal{H}^D = 0$ we recognize the case of subspace codes; the subsystem codes correspond to the case $\mathcal{H}^D \neq 0$, when we only need to correct errors modulo the subsystem structure. It follows that *the properties of subsystem codes are not fundamentally different than those of the subspace codes* we have analyzed extensively. As expected, the subsystem codes are more efficient than their subspace counterparts because: (i) they require the computation of fewer syndrome bits to identify the error(s); and (ii) the quantum circuits to implement the error correction are simpler.

**Noiseless subsystems.** Consider a system $R$ described in an $N$-dimensional Hilbert space $\mathcal{H}$ with an environment $B$ and let $B_i$ be the linearly independent environment operators; then the interaction Hamiltonian of $R$ and the environment $B$ can be expressed as

$$J = \sum_i J_i \otimes B_i.$$

When $\mathrm{tr}(J_i) = 0$ for all $i$, then we are able to retain the internal interaction of the system $R$ with the environment $B$ and remove the internal evolution of $B$. Define the noise strength parameter, $\lambda$, to be the maximum eigenvalue of $\sqrt{JJ^\dagger}$.

Let $\mathcal{J}_1$ be the set of operators $\mathcal{J}_1 = \{\mathbb{1}, J_1, J_2, \ldots\}$ with $\mathbb{1}$ the identity operator. A rigorous definition of interaction algebra is given in Section 5.22; for now we only mention that the interaction algebra $\mathcal{J}$ consists of linear combinations of products of operators in $\mathcal{J}_1$; for example, $\mathcal{J}_d = (\mathcal{J}_1)^d$ is the linear span of products of $d$ or fewer operators from $\mathcal{J}_1$. When the system $R$ consists of $n$ qubits then the dimension of the Hilbert space is $N = 2^n$ and each operator $J_i$ involves only a Pauli operator acting on a single qubit.

A c-code is a code intended only for transmission of classical information in some basis; an e-error correcting code is a c-code which permits correction of all errors described by operators in $\mathcal{J}_e$. To correct all errors in $\mathcal{J}_e$, a code $\mathcal{C}$ detects the operators in $\mathcal{J}_e^\dagger \mathcal{J}_e$; indeed, the c-code $\mathcal{C}$ with an orthonormal basis $(\mid c_1\rangle, \mid c_2\rangle, \ldots)$ is able to detect error $E$ if $\langle c_i \mid E \mid c_j\rangle = \alpha_i \delta_{i,j}$. Thus, when we use the definition of distance it follows immediately that a c-code with the minimum distance $d = 2e + 1$ is an e-error correcting code. The error amplitude of the information protected in an e-error correcting code is the amplitude of the part of the state orthogonal to the intended state; it is given as a function of time $t$ and is at most $(\lambda t)^{e+1}/(e+1)!$ with $\lambda$ the noise strength parameter [250]. The noise strength parameter is given as $\lambda = \mid \mathcal{J}_e \mid$ with $\mid \mathcal{J}_e \mid$, the maximum eigenvalue of $\sqrt{\mathcal{J}_e^\dagger \mathcal{J}_e}$. The minimum c-distance $d$ implies the existence of an orthonormal basis $(\mid c_1\rangle, \mid c_2\rangle, \ldots, \mid c_k\rangle)$ such that every operator in the set $E_l$ can be detected. It follows that $\langle c_i \mid E_l \mid c_j\rangle = \alpha_{i,l}\delta_{i,j}$; in this expression the operators in the set $\{E_1 = \mathbb{1}, E_2, \ldots, E_D\}$ form a basis for $\mathcal{J}_{d-1}$ and the dimension of this basis is $D$. Two results are proved in [250]:

(i) There exits codes of dimension at least $\lceil \frac{N}{D} \rceil$ of $R$ with minimum c-distance equal to d.

(ii) There exits codes of dimension at least $\lceil \frac{N}{D} \rceil \frac{1}{D+1}$ of $R$ with minimum distance equal to d.

Subsystem quantum error correcting codes use a quantum operation before the errors occur instead of the classical approach of encoding first and then restoring the state affected by error; this quantum operation, described by a family $\mathcal{A} = \{A_i\}$ of linear operators which transform the state, $\rho \mapsto \sum_i A_i \rho A_i^\dagger$, ensures the preservation of the information in a subsystem after the errors in $\mathcal{J}_e$ occur. The result of the state transformation of the system $R$ due to the operator $A_i \in \mathcal{A}$ followed by the error $E \in \mathcal{J}_e$ is then expressed as the product of the two operators.

This treatment is possible because subsystems of $R$ appear as factors of a tensor product of subspaces of the Hilbert space $\mathcal{H}$. Let $\mathrm{Mat}(\mathcal{H})$ be the set of all linear operators from $\mathcal{H}$ to itself and $Z(\mathcal{A})$ be the commutant of $\mathcal{A}$, the space of all operators commuting with those in $\mathcal{A}$. Then according to a theorem from representation theory the system $R$ can be described as a direct sum of tensor products of subsystems [250]

$$R \sim \bigoplus_i \mathcal{C}_i \otimes \mathcal{Z}_i \quad \text{with} \quad \mathcal{A} = \bigoplus_i \text{Mat}(\mathcal{C}_i) \otimes I^{Z_i} \quad \text{and} \quad Z(\mathcal{A}) = \bigoplus_i \mathbb{1}^{C_i} \otimes \text{Mat}(\mathcal{Z}_i)$$

with $\mathcal{Z}_i(\mathcal{C}_i)$ the state space of the subsystem $Z_i(C_i)$. The states of the subsystem $Z_i$ in $\mathcal{Z}_i$ are immune to noise; the interaction operators act only on the co-factor $\mathcal{C}_i$ thus, $Z_i$ is a noiseless subsystem and information in it is immune to noise.

Next we discuss examples of subsystem codes.

## 5.21 Bacon-Shor Code

To develop an intuition for the subsystem quantum error correction formalism we discuss first a self-correcting two-dimensional structures of $n^2$ qubits [17] and then a generalization of this idea known as Bacon-Shor code [18]. This code leads to simple and efficient methods for fault-tolerant error correction which does not require entangled ancilla states and can be implemented with nearest neighbor two-qubit measurements [9].

Recall from Section 5.12 that the stabilizer groups $\mathcal{S}$ is a subgroup of the Pauli group $\mathcal{P}_n$ that does not contain $-1$. The code space $C$ is the span of the vectors fixed by $\mathcal{S}$; if $\mathcal{S} = \{S_1, S_2, \ldots, S_s\}$ with $s < n$ then $S_i \mid \varphi\rangle = \mid \varphi\rangle,\ i \in (1, s)$ and $\dim(C) = 2^{n-s}$.

If $X_i$ and $Z_i$ are the operators corresponding to the Pauli operators $X$ and $Z$, respectively, acting on qubit $i$ we define $\{\bar{X}_i, \bar{Z}_i\},\ 1 \leq i \leq n$, a set of operators in $\mathcal{P}_n$ isomorphic with $X_i$ and $Z_i$ thus, satisfying the same commutation relations as $X_i$ and $Z_i$. Several choices for $\{\bar{X}_i, \bar{Z}_i\},\ 1 \leq i \leq n$ are possible as we can think about them as possible degrees of freedom.

**Subsystem error correction for a two-dimensional structure of $n^2$ qubits.** We consider a two-dimensional lattice of $n \times n$ qubits and define a set of products of Pauli operators acting on the $n^2$ qubits. Then we partition the set into three subsets of operators; this process induces a partition of the Hilbert space $(\mathcal{H}_2)^{\otimes n^2}$ into subspaces. Lastly, we discuss the construction of a subsystem code and the error correction procedure for the code.

The Pauli operators $P(a, b)$ acting on the $n^2$ qubits are defined with the aid of two binary configuration matrices $a = [a_{i,j}]$ and $b = [b_{i,j}]$ with $(a_{i,j}, b_{i,j}) = \{0, 1\}$ , $1 \leq (i, j) \leq n$ as

$$P(a, b) = \prod_{i=1}^{n} \prod_{j=1}^{n} \begin{cases} X_{i,j} & \text{if} \quad a_{i,j} = 1 \quad \text{and} \quad b_{i,j} = 0 \\ Z_{i,j} & \text{if} \quad a_{i,j} = 0 \quad \text{and} \quad b_{i,j} = 1 \\ -iY_{i,j} & \text{if} \quad a_{i,j} = 1 \quad \text{and} \quad b_{i,j} = 1 \end{cases} .$$

We also define

$$e_j(a) = \oplus_{i=1}^{n} a_{i,j} \quad \text{and} \quad f_j(b) = \oplus_{j=1}^{n} b_{i,j}$$

with $\oplus$ the $XOR$ operation. The set $P(a, b)$ is partitioned in three subsets $\mathcal{T}, \mathcal{S} \subset \mathcal{T}$, and $\mathcal{L}$ with the patterns of operators in each set depicted graphically in Figure 142:

- $\mathcal{T}$ - operators in $P(a, b)$ with an even number of $X_{i,j}$ operators in each column and an even number of $Z_{i,j}$ operators in each row. If $e_j = f_i = 0,\ 1 \leq (i, j) \leq n$ then $P(a, b) \in \mathcal{T}$. The operators in $\mathcal{T}$ form a non-Abelian group under multiplication.

- $\mathcal{S}$ - the stabilizers, operators in $P(a,b)$ with an even number of rows consisting entirely of $X$ operators and an even number of columns consisting entirely of $Z$ operators. If $s_i^X$ and $s_i^Z$ are the eigenvalues of the operators $S_i^X \in \mathcal{S}$ and $S_i^Z \in \mathcal{S}$, respectively, call $s^X$ and $s^Z$ the corresponding binary vectors and recognizee that the Hilbert space $(\mathcal{H}_2)^{\otimes n^2}$ can be represented as a direct sum of subspaces labelled by the binary vectors $s^X$ and $s^Z$

$$(\mathcal{H}_2)^{\otimes n^2} = \bigoplus_{s^X,s^Z} \mathcal{H}_{s^X,s^Z}.$$

  The operators in $\mathcal{S}$ form a multiplicative Abelian subgroup of $\mathcal{T}$; every operator in $\mathcal{S}$ commutes with all other operators in $\mathcal{S}$ and also with all operators in $\mathcal{T}$.

- $\mathcal{L}$ - operators in $P(a,b)$ with an odd number of rows consisting only of $X_{i,j}$ operators and an odd number of columns consisting only of $Z_{i,j}$ operators. The operators in $\mathcal{L}$ form a non-Abelian group under multiplication and commute with the operators in $\mathcal{S}$.

The operators in the sets $\mathcal{T}$ and $\mathcal{L}$ must be represented on the subspaces $\mathcal{H}_{s^X,s^Z}$ as subsystem actions; $\mathcal{T}$ and $\mathcal{L}$ induce a partitioning of the Hilbert spaces into direct sums of tensor products of subspaces $\mathcal{H}_{s^X,s^Z}$ as follows:

$$(\mathcal{H}_2)^{\otimes n^2} = \bigoplus_{s^X,s^Z} \mathcal{H}^{\mathcal{T}}_{s^X,s^Z} \otimes \mathcal{H}^{\mathcal{L}}_{s^X,s^Z}.$$

The dimensions of the two spaces are $2^{(n-1)^2}$ and 2, respectively [17]; $\mathcal{H}^{\mathcal{T}}_{s^X,s^Z}$ is the gauge subsystem mentioned in Section 5.20. Operators $T \in \mathcal{T}$ and $L \in \mathcal{L}$ act only on the first and, respectively, the second subspace of the two tensor products

$$T = \bigoplus T_{s^X,s^Z} \otimes \mathbb{1}_2 \quad \text{and} \quad L = \bigoplus \mathbb{1}_{2^{(n-1)^2}} \otimes L_{s^X,s^Z}.$$

Under this decomposition an operator $P(a,b)$ is block diagonal and can be expressed as

$$P(a,b) = \bigoplus_{s^X,s^X} E_{s^X,s^Z} \otimes \mathbb{1}_2$$

with $E_{s^X,s^Z}$ an error operator depending on the error $(a,b)$ and the subspaces labelled $s^X$ and $s^Z$. The information encoded in the subspaces $\mathcal{H}^{\mathcal{L}}_{s^X,s^Z}$ identified by $s^X = s^Z = 11\ldots11$, or, equivalently $e_j = f_i = 0$ is not affected by these errors, these subspaces are noiseless.

The subsystem code encodes one qubit of information as $n^2$ qubits in the Hilbert space $\mathcal{H}^{\mathcal{L}}_{s^X,s^Z}$ with $s^X = s^Z = 11\ldots11$. To restore the system to $e_j(a) = f_i(a)$ we carry out simultaneously $2(n-1)$ measurements of the eigenvectors of $S_i^X$ and $S_i^Z$. Measuring $S_i^X$ and $S_i^Z$ is equivalent to computing

$$\bigoplus_{j=1}^{n}(b_{i,j} \oplus b_{i+1,j}) = f_i(b) \oplus f_{i+1}(b) \quad \text{and} \quad \bigoplus_{i=1}^{n}(a_{i,j} \oplus a_{i,j+1}) = e_j(b) \oplus e_{j+1}(b).$$

Indeed, if $f_i(b)$ is treated as an $n$-bit codeword of a code with two logical codewords, $f_i(b) = 0,\ 1 \le i \le n$ and $f_i(b) = 1,\ 1 \le i \le n$, then the measurement of the $(n-1)$

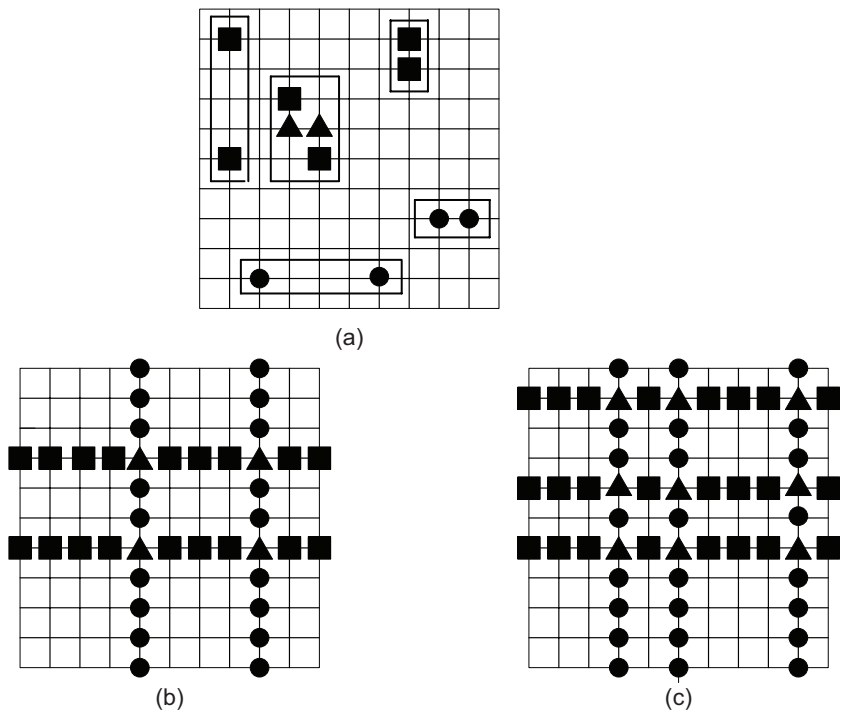



Figure 142: The operators in $\mathcal{T}, \mathcal{S}$, and $\mathcal{L}$. The qubits are located at the vertices of the $(11\times 11)$ $2D$-lattice; a square on a vertex represents an $X$ operator acting on the corresponding qubit, a circle denotes a $Z$ operator, a triangle an $Y$ operator, and the absence of any shape represents the identity operator, $\mathbb{1}$. (a) $\mathcal{T}$ - the subset of $P(a,b)$ operators with an even number of $X$ operators in each column and an even number of $Z$ operators in each row; the rectangles mark such sets. (b) $\mathcal{S}$ - the subset of $P(a,b)$ operators with an even number of rows consisting entirely of $X$ operators and an even number of columns consisting entirely of $Z$ operators. (c) $\mathcal{L}$ - the subset of $P(a,b)$ operators with an odd number of rows consisting entirely of $X$ operators and an odd number of columns consisting entirely of $Z$ operators.

eigenvalues of the stabilizer $S_i^X$ is equivalent to computing the syndrome and then using this syndrome to correct the errors and restore $f_i(b) = 1,\ 1 \leq i \leq n$ to either 0 or 1 logical codewords. A similar reasoning can be carried out for $e_j(a)$.

Next we discuss a stabilizer formalism for classical linear codes and review the basic ideas of Shor codes analyzed in Sections 5.5, 5.6, and 5.7, then we introduce the generalized Shor codes, and the Bacon-Shor code.

**A stabilizer formalism for a binary linear code used to correct bit-flip or phase-flip errors.** Let $\mathcal{C}$ be an $[n,k,d]$ classical linear code with a $k \times n$ generator matrix $\mathrm{G} = [g_{i,j}]$ and an $(n-k) \times n$ parity check matrix $\mathrm{H} = [h_{i,j}]$. The $k$ rows of G are linearly independent vectors in $GF(2^n)$ and $(n-k)$ rows of H are also linearly independent.

Call $\mathrm{G}^c = [g^c_{i,j}]$ a $(n-k) \times n$ matrix whose rows are linearly independent of each other and also of the rows of G; together the rows of G and the rows of $\mathrm{G}^c$ form a basis in $GF(2^n)$. Similarly, $\mathrm{H}^c = [h^c_{i,j}]$ is a $k \times n$ matrix whose rows are linearly independent of each other and also of the rows of H; together the rows of H and the rows of $\mathrm{H}^c$ form a basis in $GF(2^n)$.

When we use $\mathcal{C}$ to correct bit-flip quantum errors we define four sets of operators: stabilizers, $S_i$, logical $X$ operators $\bar{X}_i$, logical $Z$ operators $\bar{Z}_i$, and error operators, $E_i$.

*The stabilizers are obtained from the rows of the parity check matrix;* the stabilizer generator $S_i$ is a tensor product of $n$ terms; each term is either the identity, $\mathbb{1}$, or the $Z$ operator thus, the generators of the stabilizer group commute with each other. The $Z$ operators in this product are selected by the elements $h_{i,j}$ of the parity check matrix

$$S_i = \bigotimes_{j=1}^{n} Z^{h_{i,j}}.$$

To compute the syndrome we measure the eigenvalues of the $(n-k)$ stabilizer generators $S_i$; indeed, the measurements are equivalent to computing the syndrome using the parity check matrix H.

The $k$ generators of *logical* $X$ operators, $\bar{X}_i$, for the $k$ encoded qubits are selected by the rows of the generator matrix; each one is a tensor product of $n$ terms each being an $X$ operator selected by the elements of the generator matrix G

$$\bar{X}_i = \bigotimes_{j=1}^{n} X^{g_{i,j}}.$$

It is easy to see that $\bar{X}_i$ commute with $S_j$, $\bar{X}_i S_j = S_j \bar{X}_i$, and they are not in the stabilizer set as they consist entirely of $X$ operators. The $k$ generators of the *logical* $Z$ operator, $\bar{Z}_i$, for the $k$ encoded qubits are selected by the rows of the matrix $\mathrm{H}^c$; each one is a tensor product of $n$ terms each being an $X$ operator selected by the elements of the matrix $\mathrm{H}^c = [h^c_{i,j}]$

$$\bar{Z}_i = \bigotimes_{j=1}^{n} Z^{h^c_{i,j}}.$$

The *logical* $Y$ operators are $\bar{Y}_i = i\bar{X}_i\bar{Z}_i$. Lastly, the $(n-k)$ generators of error operators are selected by the rows of the matrix $\mathrm{G}^c = [g^c_{i,j}]$

$$E_i = \bigotimes_{j=1}^{n} X^{g^c_{i,j}}.$$

The products of $E_i$ operators anticommute with at least one $S_i$ thus, we can detect all error that can be expressed as a product of one or more $E_i$ with $\bar{X}_i$ logical operators

To correct phase-flip errors we interchange the $Z$ operators with $X$ operators in $S_i$ and $\bar{Z}_i$ and the $X$ operators with $Z$ operators in $\bar{X}_i$. The code for phase-flip error correction operates in the dual basis $(|+\rangle, |-\rangle)$ with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ instead of operating in the computational basis $(|0\rangle, |1\rangle)$.

**Generalized Shor code.** After reviewing concatenated codes we recognize that the basic idea of the Shore code discussed in Section 5.7 is to concatenate an inner code, $\mathcal{C}_1 = [n_1, k_1, d_1]$, to correct bit-flip, or $X$-errors, with an outer code, $\mathcal{C}_2 = [n_2, k_2, d_2]$, to correct phase-flip, or $Z$-errors. The generator matrices of $\mathcal{C}_1$ and $\mathcal{C}_2$ are $\mathrm{G}_1 = [g^1_{i,j}],\ 1 \le i \le k_1, 1 \le j \le n_1$ and $\mathrm{G}_2 = [g^2_{i,j}],\ 1 \le i \le k_2, 1 \le j \le n_2$, respectively; the parity check matrices of $\mathcal{C}_1$ and $\mathcal{C}_2$ are

$\mathrm{H}_1 = [h^1_{i,j}],\ 1 \le i \le n_1 - k_1, 1 \le j \le n_1$ and $\mathrm{H}_2 = [h^2_{i,j}],\ 1 \le i \le n_2 - k_2, 1 \le j \le n_2$, respectively.

The generalized Shor code denoted as $[n_1 n_2, k_1 k_2, \min(d_1, d_2)]$ is defined as follows [18]: arrange the $n_1 n_2$ physical qubits in a two-dimensional array with $n_2$ rows and $n_1$ columns; each row of $n_1$ qubits encodes $k_1$ logical qubits; each column of $n_2$ qubits encodes $k_2$ qubits; use $\mathcal{C}_1$ to correct bit-flip errors and $\mathcal{C}_2$ to correct phase-flip errors.

Phase errors on the qubits of a row of the array are phase errors on the qubits already encoded with $\mathcal{C}_1$ thus we can use $\mathcal{C}_2$ to correct phase-flip errors. Indeed, a tensor product of the identity, $\mathbb{1}$, and the $Z$ Pauli operators is either a stabilizer or a product of stabilizers; acting on a row of this array will result in a phase-flip error on some of the encoded qubits times an element of the stabilizer. We conclude that the $[n_1 n_2, k_1 k_2, \min(d_1, d_2)]$ code has distance $d_1$ for bit-flip errors, distance $d_2$ for phase-flip errors, and distance $\min(d_1, d_2)$ for bit- and phase-flip errors; thus, the code can correct $e = \lfloor (\min(d_1, d_2) - 1)/2 \rfloor$ errors. Error correction for a generalized Shor code requires the measurement of $[(n_1 - k_1)n_2 + (n_2 - k_2)]$ eigenvalues of the stabilizer generators.

**The subsystem code.** The procedure follows closely the one discussed earlier in this section for the two-dimensional lattice of $n^2$ qubits. We use the rows of the parity check matrix $\mathrm{H}_1$ of $\mathcal{C}_1$ to construct the a stabilizer code on $n_1$ qubits and the rows of the parity matrix $\mathrm{H}_2$ of $\mathcal{C}_2$ to construct the a stabilizer code on $n_2$ qubits. The the first stabilizer group consists of $n_1 - k_1$ stabilizers, $\mathcal{S}_1 = \{S_1, S_2, \ldots S_{n_1 - k_1}\}$ with $S_i = \otimes_{j=1}^{n_1} Z^{h^1_{i,j}}$; the second stabilizer group consists of $n_2 - k_2$ stabilizers, $\mathcal{S}_2 = \{T_1, T_2, \ldots T_{n_2 - k_2}\}$ with $T_i = \otimes_{j=1}^{n_2} Z^{h^2_{i,j}}$. The stabilizers in $\mathcal{S}_1$ will be used on the columns and the ones in $\mathcal{S}_2$ on the rows to construct the non-Abelian group $\mathcal{T}$. The group $\mathcal{T}$ consists of $\mathcal{T}_1$, operators from $\mathcal{S}_1$ acting on the columns of the array, and $\mathcal{T}_2$, operators from $\mathcal{S}_2$ acting on the rows of the array.

Next we construct an Abelian subgroup $\mathcal{S} \subset \mathcal{T}$ of operators from $\mathcal{S}_1$ and from $\mathcal{S}_2$ which commute with all the elements of $\mathcal{T}$. We take an operator from $\mathcal{S}_1$ and a codeword $v = (v_1, v_2, \ldots v_n) \in \mathcal{C}_2$ and construct an operator acting on $n$ qubits with $S_1^{v_i}$ acting on each column $j,\ 1 \le j \le n$. This operator is in $\mathcal{T}$ and commutes with every operator in $\mathcal{T}$; indeed, in a particular row the operators in $\mathcal{T}$ are constructed of $Z$ operators and in a particular column they are made of operators in $\mathcal{S}_1$. Similarly, we take an operator from $\mathcal{S}_2$ and a codeword $w = (w_1, w_2, \ldots w_n) \in \mathcal{C}_2$ and construct an operator on $n$ qubits with $S_2^{w_j}$ acting on each row $i,\ 1 \le i \le n$. This operator is in $\mathcal{T}$ and commutes with every operator in $\mathcal{T}$. We conclude that $\mathcal{S}$ is an Abelian invariant subgroup of $\mathcal{T}$ as all operators in $\mathcal{T}$ commute with all operators from $\mathcal{S}_1$ and $\mathcal{S}_2$.

$\mathcal{S}$ is the stabilizer group of the stabilizer code and the operators in $\mathcal{T}$ and $\mathcal{L}$ are in the normalizer of this code, all their elements commute with the operators in $\mathcal{S}$. The subspaces of the Hilbert space $(\mathcal{H}_2)^{\otimes n_1 n_2}$ are labelled with the eigenvalues of the $(n_1 - k_1)k_2$ stabilizer generators which are tensor products of $\mathbb{1}$ and $Z$ Pauli operators and $k_1(n_2 - k_2)$ stabilizer generators which are tensor products of $\mathbb{1}$ and $X$ Pauli operators.

Thus we have: (i) a stabilizer $\mathcal{S}$ with $(n_1 - k_1)k_2 + k_1(n_2 - k_2)$ generators; (ii) $\mathcal{T}/\mathcal{S}$ generators which act as a Pauli group on $(n_1 - k_1)(n_2 - k_2)$ qubits, and (iii) encoded Pauli operators $\mathcal{L}$ which act on $k_1 k_2$ independent qubits which carry the information [18]. The total number of Pauli operators in these sets is: $[(n_1 - k_1)k_2 + k_1(n_2 - k_2)] + [(n_1 - k_1)(n_2 - k_2)] + [k_1 k_2] = n_1 n_2$.

The Hilbert space can then be decomposed into tensor products of subspaces $\mathcal{H}_s^{\mathcal{T}/\mathcal{S}}$ of

dimension $2^{(n_1-k_1)(n_2-k_2)}$ and $\mathcal{H}_s^{\mathcal{L}}$ of dimension $2^{k_1k_2}$ with $s$ an $[(n_1-k_1)k_2+k_1(n_2-k_2)]$-tuple of eigenvalues equal to $(\pm 1)$ of the stabilizer generators

$$(\mathcal{H}_2)^{\otimes n_1 n_2} = \bigoplus_{s\in\{0,1\}^{(n_1-k_1)k_2+k_1(n_2-k_2)}} \mathcal{H}_s^{\mathcal{T}/\mathcal{S}} \otimes \mathcal{H}_s^{\mathcal{L}}.$$

The quantum information is encoded into the subsystem described in $\mathcal{H}_s^{\mathcal{L}}$ a subspace were $s_i = +1$. The measurement of the generators of the stabilizer group $\mathcal{S}$ allows us to correct errors that can be corrected with $\mathcal{C}_1$ and $\mathcal{C}_2$ [18].

The concatenation for the generalized Shor codes led to an asymmetry between bit- and phase-flips and the Bacon-Shor code removes this asymmetry and produces a substantial saving in the number of measurements of the stabilizers for error correction; this number is reduced from $[(n_1-k_1)n_2+(n_2-k_2)]$, for a generalized Shor code, to $[(n_1-k_1)k_2+k_1(n_2-k_2)]$ for a subsystem code constructed from two classical codes [18]. Indeed, $k_1 \lll n_1$ and $k_2 \lll n_2$ and this reduction can be exploited successfully by fault-tolerant systems [9].

The operator quantum error formalism discussed next unifies the subspace codes based on the standard model for quantum error correction, the decoherence-free subspaces model [318], and the noiseless subsystem model discussed in Section 5.20.

## 5.22 Operator Quantum Error Correction

In this section we discuss a generalized formalism for quantum error correction, the Operator Quantum Error Correction (OQEC) [260, 261]. The roots of the operator quantum correction formalism can be traced back to Werner Heisenberg and Pascual Jordan and to concepts from modern functional analysis applied to quantum mechanics. OQEC is based on the concept of equivalence between states of composite systems expressed as tensor products; two states $\rho^A \otimes \rho^B$ and $\rho^A \otimes (\rho')^B$ are said to be equivalent if they carry the same information even though $\rho^B$ and $(\rho')^B$ are different.

**Basic concepts.** A Banach algebra, $A$, is an associative algebra over $\mathbb{R}$ or $\mathbb{C}$, the fields of real and complex numbers, respectively, when the multiplication obeys the rule

$$\|ab\| \leq \|a\|\|b\| \quad \forall a,b \in \mathbb{R} \;\text{ or }\; \forall a,b \in \mathbb{C}.$$

A $C^*$-algebra is a Banach algebra over $\mathbb{C}$ together with an involution, a map $* : A \mapsto A$ with several properties [110]:

$$\begin{array}{lll}
1. & (a+b)^* = a^* + b^*, & \forall a,b \in A \\
 & (ab)^* = b^* a^*, & \forall a,b \in A \\
2. & (\lambda a)^* = \bar{\lambda} a^*, & \forall \lambda \in \mathbb{C} \;\text{ and }\; \forall a \in A \\
3. & (a^*)^* = a, & \forall a \in A \\
4. & \|aa^*\| = \|a\|\|a^*\|, & \forall a \in A \\
 & \|a^*a\| = \|a^*\|\|a\|, & \forall a \in A.
\end{array}$$

Condition (4) is called the $C^*$-identity and can also be expressed as $\|aa^*\| = \|a\|^2$.

If $A$ and $B$ are two $C^*$-algebras the mapping $F : A \mapsto B$ is a $C^*$-homomorphism when the following two conditions are satisfied:

$$\begin{array}{ll} 1. & F(ab) = F(a)F(b), \quad \forall a, b \in A \\ 2. & F(a^*) = F(a)^*, \qquad \forall a \in A. \end{array}$$

The $C^*$-identity implies that any homomorphism between $C^*$-algebras is bounded, its norm is smaller or equal to 1; a bijective $C^*$-homomorphism is a $C^*$-isomorphism.

**Example.** *A finite-dimensional $C^*$-algebra, also called a $\dagger$-algebra.* Let $\mathcal{M}_m$ be the set of $m \times m$ matrices over $\mathbb{C}$ with the involution defined as the conjugate transpose. When the matrices in $\mathcal{M}_n$ are viewed as operators over $\mathbb{C}^m$ with the operator norm $\|.\|$ we have a finite-dimensional $C^*$-algebra.

The finite direct sum of two algebraic structures is the Cartesian product of their underlaying sets. For example, given two Abelian groups $(A, \top)$ and $(B, \perp)$ then $A \bigoplus B$ is

$$(a_1, b_1) \circ (a_2, b_2) = (a_1 \top a_2), (b_1 \perp b_2)$$

with $\circ$ denoting the group operation. It is easy to show that a finite-dimensional $C^*$-algebra is isomorphic with a finite direct sum.

Given a Hilbert space $\mathcal{H}$ the algebra of bounded operators $\mathcal{B}(\mathcal{H})$ is a $C^*$-algebra with $a^*$ the adjoint (dual) of the operator $a : \mathcal{H} \mapsto \mathcal{H}$.

A quantum channel, also called a quantum operator, $\mathcal{E}$, is a completely positive and trace-preserving linear map $\mathcal{E} : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ with an operator sum (Kraus) representation

$$\mathcal{E}(\sigma) = \sum_a E_a \sigma E_a^\dagger, \quad \forall \sigma \in \mathcal{B}(\mathcal{H}).$$

$\mathcal{E}$ is unital if

$$\mathcal{E}(\mathbb{1}) = \sum_a E_a E_a^\dagger = \mathbb{1}.$$

The algebra $\mathcal{A}$ associated with $\mathcal{E} = \{E_a\}$ is a $\dagger$-algebra called the interaction algebra; $\mathcal{A}$ is the algebra generated by $E_a$ and $E_a^\dagger$. An orthonormal basis exists and the matrix representation of the operators in $\mathcal{A}$ with respect to this basis is of the form

$$\mathcal{A} \approx \bigoplus_J \left(\mathcal{M}_{m_J} \otimes \mathbb{1}_{n_J}\right).$$

This decomposition is unique up to a unitary equivalence.

The noise commutant associated with the channel $\mathcal{E}$ is defined as

$$\mathcal{A}' = \left\{\sigma \in \mathcal{B}(\mathcal{H}) : E\sigma = \sigma E, \quad \forall E \in \{E_a E_a^\dagger\}\right\}.$$

The noise commutant is unitarily equivalent with the direct sum

$$\mathcal{A}' \approx \bigoplus_J \left(\mathbb{1}_{m_J} \otimes \mathcal{M}_{n_J}\right).$$

We see that elements of $\mathcal{A}'$ are not affected by error of $\mathcal{A}$ when $\mathcal{E}$ is unital. It was proven that when $\mathcal{E}$ is unital then the noise commutant coincides with the fixed point set[40] of $\mathcal{E}$

$$\mathcal{A}' = \mathrm{Fix}\,(\mathcal{E}) = \{\sigma \in \mathcal{B}(\mathcal{H}) : \mathcal{E} = \sum_a E_a \sigma E_a^\dagger = \sigma\}.$$

The $\mathcal{A}'$ may be used to produce noiseless subsystems for unital $\mathcal{E}$.

The interaction algebra $\mathcal{A}$, due to its structure, induces a decomposition of the original Hilbert space

$$\mathcal{H} = \bigoplus_J \mathcal{H}_J^A \otimes \mathcal{H}_J^B,$$

where the noisy subsystems $\mathcal{H}_J^A$ and $\mathcal{H}_J^B$ have dimensions $m_J$ and $n_J$, respectively. We only discuss the case when the information is encoded in a single noiseless sector; then the decomposition is

$$\mathcal{H} = \left(\mathcal{H}^A \otimes \mathcal{H}^B\right) \bigoplus \mathcal{K}$$

with $\dim(\mathcal{H}^A) = m$, $\dim(\mathcal{H}^B) = n$, and $\dim(\mathcal{K}) = \dim(\mathcal{H}) - mn$. $\sigma^A$ and $\sigma^B$ are the operators in $\mathcal{B}(\mathcal{H}^A)$ and $\mathcal{B}(\mathcal{H}^B)$, respectively. When $\sigma = \mathbb{1}^A$ is the identity element of $\mathcal{B}(\mathcal{H}^A)$ the restriction of the noise commutant $\mathcal{A}'$ to $\mathcal{H}^A \otimes \mathcal{H}^B$ consists of the operators $\sigma = \mathbb{1}^A \otimes \sigma^B$.

**Generalized noiseless systems.** We define a set of *matrix units* in $\mathcal{A}$ for the decomposition $\mathcal{H} = \left(\mathcal{H}^A \otimes \mathcal{H}^B\right) \bigoplus \mathcal{K}$ as

$$P_{kl} \equiv \mid \alpha_k\rangle\langle\alpha_l \mid \otimes \mathbb{1} \quad 1 \le k, l \le m$$

with $\mid \alpha_k\rangle$ and $\mid \beta_k\rangle$ basis vectors for $\mathcal{H}^A$ and $\mathcal{H}^B$, respectively, $\mathcal{H}^A = \mathrm{span}\{\mid \alpha_k\rangle\}_{k=1}^m$ and $\mathcal{H}^B = \mathrm{span}\{\mid \beta_k\rangle\}_{k=1}^n$; with this choice of the basis vectors the matrix representation of the subalgebra of $\mathcal{A}'$ is $\mathbb{1}^A \otimes \mathcal{B}(\mathcal{H}^B)$. The following identities for a family of matrix elements are established in [260, 261]:

$$\begin{array}{lcl} P_{kl} & = & P_{kk} P_{kl} P_{ll}, \quad 1 \le k, l \le m \\ P_{kl}^\dagger & = & P_{lk}, \quad 1 \le k, l \le m \\ P_{kl} P_{l'k'} & = & \begin{cases} P_{kk'} & \text{if} \quad l = l' \\ 0 & \text{if} \quad l \ne l' \end{cases} \end{array}$$

The projection $P_{\mathcal{U}}$ and the superoperator $\mathcal{P}_{\mathcal{U}}$ are defined as $P_{\mathcal{U}} = P_{11} + P_{22} + \ldots + P_{mm}$ and as the action $\mathcal{P}_{\mathcal{U}}(\cdot) = P_{\mathcal{U}}(\cdot)P_{\mathcal{U}}$, respectively.

For the fixed decomposition of $\mathcal{H}$ when the information is encoded in a single noiseless sector, we define the operators

$$\mathcal{U} = \{\sigma \in \mathcal{B}(\mathcal{H}) : \sigma = \sigma^A \otimes \sigma^B \ \text{ for some } \ \sigma^A, \sigma^B\}.$$

Given a quantum operation $\mathcal{E}$ on $\mathcal{B}(\mathcal{H})$ a system is said to be noiseless for the decomposition $\mathcal{H} = \left(\mathcal{H}^A \otimes \mathcal{H}^B\right) \bigoplus \mathcal{K}$ if any one of the following three conditions are equivalent and define the properties of the noiseless subsystem $B$:

[40] $x$ is a fix point of $f(x)$ if $f(x) = x$.

$$
\begin{array}{lll}
1. & \forall \sigma^A, \forall \sigma^B & \exists \tau^A : \mathcal{E}(\sigma^A \otimes \sigma^B) = \tau^A \otimes \sigma_B \\
2. & \forall \sigma^A & \exists \tau^A : \mathcal{E}(\mathbb{1}^A \otimes \sigma^B) = \tau^A \otimes \sigma_B \\
3. & \forall \sigma \in \mathcal{U} & : (\mathrm{tr}_A \circ \mathcal{P}_{\mathcal{U}} \circ \mathcal{E})(\sigma) = \mathrm{tr}_A(\sigma).
\end{array}
$$

Given a map $\mathcal{E} = \{E_a\}$ and a semigroup $\mathcal{U}$ of $\mathcal{B}(\mathcal{H})$, the three equivalent necessary and sufficient conditions for existence of a noiseless subsystem are [260]:

1. The $B$-sector of $\mathcal{U}$ encodes a noiseless subsystem for $\mathcal{E}$ or a decoherence-free subspace when $m = 1$;

2. The subspace $P_{\mathcal{U}}\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^B$ is invariant for the operation $E_a$ and the restrictions $E_a \mid_{P_{\mathcal{U}}\mathcal{H}}$ belong to the algebra $\mathcal{B}(\mathcal{H}^A) \otimes \mathbb{1}^B$;

3. Two condition hold for any choice of the matrix units $\{P_{kl} : 1 \leq k, l \leq m\}$ for $\mathcal{B}(\mathcal{H}^A) \otimes \mathbb{1}^B$

$$
\begin{array}{ll}
1. & P_{kk} E_a P_{ll} = \lambda_{akl} P_{kl}, \quad \forall a, k, l \ \text{ and for some set of scalars } (\lambda_{akl}) \\
2. & E_a P_{\mathcal{U}} = P_{\mathcal{U}} E_a P_{\mathcal{U}}, \quad \forall a.
\end{array}
$$

The codes described by Operator Quantum Error Correction consist of operators semigroups and algebras as we can see from the following definition: a <u>correctable B-sector of $\mathcal{U}$ for a channel $\mathcal{E}$ and a recovery $\mathcal{R}$</u> is one when the $\mathcal{H}^B$ sector $\mathcal{U}$ encodes a noiseless subsystem for the map $\mathcal{R} \circ \mathcal{E}$ given the decomposition $\mathcal{H} = \left(\mathcal{H}^A \otimes \mathcal{H}^B\right) \bigoplus \mathcal{K}$; formally we express this condition as

$$
\left(\mathrm{tr}_A \circ \mathcal{P}_{\mathcal{U}} \circ \mathcal{R} \circ \mathcal{E}\right)(\sigma) = \mathrm{tr}_A(\sigma), \ \forall \sigma \in \mathcal{U}.
$$

The following proposition is proved in [261]:

**Proposition.** *Let $\mathcal{E} = \{E_a\}$ be a quantum operation on $\mathcal{B}(\mathcal{H})$ and let the set of operators*

$$
\mathcal{U} = \{\sigma \in \mathcal{B}(\mathcal{H}) : \sigma = \sigma^A \otimes \sigma^B \ \ \textit{for some} \ \ \sigma^A, \sigma^B\}
$$

*be a semigroup. Then the B-sector of $\mathcal{U}$ is correctable for $\mathcal{E}$ if and only if a quantum recovery operation $\mathcal{R} \in \mathcal{B}(\mathcal{H})$ exists and*

$$
(\mathcal{R} \circ \mathcal{E})(\sigma) = \sigma, \ \forall \sigma \in \mathcal{U}.
$$

We conclude this section with an example.

**Example.** *Two noiseless subsystems for a non-unital quantum channel [261].* Let $\mathcal{E} = \{E_0, E_1\}$ be a channel on $\mathbb{C}^2 \otimes \mathbb{C}^2$ with $E_0$ and $E_1$ given by

$$
E_0 = \sqrt{1-2p}\,(\mid 00\rangle\langle 00 \mid + \mid 11\rangle\langle 11 \mid) + \mid 01\rangle\langle 01 \mid + \mid 10\rangle\langle 10 \mid
$$

and

$$
E_1 = \sqrt{p}\,(\mid 00\rangle\langle 00 \mid + \mid 10\rangle\langle 00 \mid + \mid 01\rangle\langle 11 \mid + \mid 11\rangle\langle 11 \mid)
$$

with $0 < p < 1$. Recall that a channel $\mathcal{E} = \{E_a\}$ is unital if $\mathcal{E}(\mathbb{1}) = \sum_a E_a E_a^\dagger = \mathbb{1}$; the channel in this example is non-unital as $E_0 E_0^\dagger + E_1 E_1^\dagger \neq \mathbb{1}$.

(a) The first subsystem decomposition is $\mathcal{H}^{A_1} = \mathbb{C}$ and $\mathcal{H}^{B_1} = \text{span}\{|\ 01\rangle, |\ 10\rangle\}$; if subsystem $B_1$ is a noise-free subsystem then the states immune to the effects of the channel $\mathcal{E}$ are the logical zero and the logical one, $|\ 0\rangle_L = |\ 01\rangle$ and $|\ 1\rangle_L = |\ 10\rangle$. For this decomposition the condition 3.1 above for the existence of a noise-free subspace is satisfied for $Q = |\ 01\rangle\langle 01\ | + |\ 10\rangle\langle 10\ |$. Indeed,

$$E_0 Q = Q = Q E_0 = Q E_0 Q \qquad \text{and} \qquad E_1 Q = 0 = Q E_1 Q.$$

Alternatively, we can show that there exists $\sigma \in \mathbb{B}(\mathcal{H}^{B_1}) = \mathbb{B}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{B_1})$ such that $\mathcal{E}(\sigma) = \sigma$; let $a, b, c, d \in \mathbb{C}$ and define

$$\sigma = a\ |\ 01\rangle\langle 01\ | + b\ |\ 01\rangle\langle 10\ | + c\ |\ 10\rangle\langle 01\ | + d\ |\ 10\rangle\langle 10\ |\ .$$

We see immediately that the condition $\mathcal{E}(\sigma) = \sigma$ is satisfied. Lastly, we observe that $E_1 \sigma = 0$ and $\sigma E_1 \neq 0$ thus, the noiseless subsystem $B_1$ is not supported by a noise commutant for $\mathcal{E}$.

(b) The second subsystem decomposition is supported by a noise commutant for $\mathcal{E}$. In this case: $\mathcal{H}^{A_2} = \text{span}\{|\ \alpha_1\rangle, |\ \alpha_2\rangle\}$ and $\mathcal{H}^{B_2} = \text{span}\{|\ \beta_1\rangle, |\ \beta_2\rangle\}$. If $B_2$ is a noise-free subsystem then the states immune to the effects of the channel $\mathcal{E}$ are the logical zero and the logical one, $|\ 0\rangle_L = |\ \beta_1\rangle$ and $|\ 1\rangle_L = |\ \beta_2\rangle$. Indeed, $B_2$ is a noise-free subsystem as condition 3.2 above is satisfied; it is left as an exercise, see Section 5.28, to prove the following equalities

$$P_{11} E_0 P_{11} = \alpha P_{11}, \quad P_{11} E_1 P_{11} = \beta P_{11}, \quad P_{11} E_0 P_{22} = 0 P_{12}, \quad P_{11} E_1 P_{22} = 0 P_{22},$$

and

$$P_{22} E_0 P_{22} = P_{22}, \quad P_{22} E_1 P_{22} = 0 P_{22}, \quad P_{22} E_0 P_{11} = 0 P_{21}, \quad P_{22} E_1 P_{22} = 0 P_{21}.$$

Then we see that

$$\mathcal{E}(\mathbb{1}_2 \otimes \sigma) = \begin{pmatrix} 1-p & p \\ p & 1+p \end{pmatrix} \otimes \sigma, \ \forall \sigma \in \mathbb{B}(\mathcal{H}^{B_2}).$$

We conclude the presentation of the subsystem codes with a discussion of stabilizers for OQEC in the next section, a topic analyzed in [338].

## 5.23 Stabilizers for Operator Quantum Error Correction

Operator Quantum Error Correction is based on a fixed partition of the Hilbert space $\mathcal{H}$, $\mathcal{H} = A \otimes B \oplus C^\perp$, with $C$ the code space. When the information is encoded in subsystem $A$ then the density matrix of the resulting state is $\rho^A \otimes \rho^B \oplus 0^{C^\perp}$ with $\rho^A \in \mathcal{B}(A)$ the density matrix of the logical state carrying useful information and $\rho^B$ is the density matrix of an arbitrary state.

We consider a system consisting of $n$ qubits in $\mathcal{H}_{2^n}$. We choose a set of $2n$ operators $\{X'_j, Z'_j\}_{j=1,n} \subset \mathcal{P}_n$ [338]; the operators $\{X'_j, Z'_j\}_{j=1,n}$ act on $n$ virtual qubits in the code space C, the same way the Pauli operators $X_i$ and $Z_i$ act on a single physical qubit.

The stabilizer group of the code, $\mathcal{S}$, is a Abelian subgroup of $\mathcal{P}_n$ that does not contain $-1$; the generators $S_j$ in $\mathcal{S} = \{S_1, S_2, \ldots, S_s\},\ s \leq n$ are independent commuting elements of $\mathcal{P}_n$ thus, they can be diagonalized simultaneously; we can choose $S_j = Z'_j$. The code space $C$ has dimension $2^{n-s}$ and is the span of the vectors $\mid \varphi\rangle$ invariant to the transformation carried out by any stabilizer, $S_j \mid \varphi\rangle = \mid \varphi\rangle,\ 1 \leq j \leq s$; the projector onto the code space, $P$, satisfies the relation $S_j P = P,\ 1 \leq j \leq s$.

Every pair of elements of the Pauli group $\mathcal{P}_n$ either commutes or anticommutes. $N(\mathcal{S})$ is the subgroup of $\mathcal{P}_n$ that commutes with every element of $\mathcal{S}$ and is called the normalizer group; the normalizer maps the code subspace to itself. In our case the normalizer is

$$N(\mathcal{S}) = \{i, Z'_1, Z'_2, \ldots, Z'_n, X'_1, X'_2, \ldots, X'_n\}.$$

We define an equivalence relation "$\sim$" induced by a set of gauge transformations; two states with the density matrices $\rho$ and $\rho'$ are said to be equivalent, $\rho \sim \rho'$, if and only if there exists a transformation $g$ such that $\rho = g\rho' g^\dagger$; the gauge transformations $g$ belong to the gauge group $\mathcal{G}$.

The equivalence relation requires the gauge group $\mathcal{G}$ to be a subgroup of the normalizer group, $\mathcal{G} \subset N(\mathcal{S})$ in order to keep the states $\mid \varphi\rangle \in C$ in the code subspace $C$. The operators in $\mathcal{S}$, as well as the complex number $i$ leave the states $\mid \varphi\rangle \in C$ invariant under conjugation and should be in $\mathcal{G}$. The gauge group can then be generated by the stabilizer generators, $S_j,\ 1 \leq j \leq s$, the complex number $i$, and an arbitrary subset of $X'_{i_j}$ and $Z'_{i_j}$ with $j \leq s$

$\mathcal{G} = \{S_1, S_2, \ldots S_s, X'_{i_1}, \ldots, X'_{i_\alpha}, Z'_{i_1}, \ldots, Z'_{i_\beta}\}$

Thus, $\mathcal{G}$ is a normal subgroup of $N(\mathcal{S})$. Then $\mathcal{L} = \{L_j\} = N(\mathcal{S})/\mathcal{G}$ is the quotient group[41] of $N(\mathcal{S})$ and $\mathcal{G} \times \mathcal{L} \approx \mathcal{N}(\mathcal{S})$. In [469] it is shown that $[\mathcal{G}, \mathcal{L}] = 0$ and the operators $X'_{i_j}$ and $X'_{i_j}$ must appear in pairs thus,

$\mathcal{G} = \{S_1, S_2, \ldots S_s, X'_{s+1}, Z'_{s+1}, \ldots, X'_{s+r}, Z'_{s+r}\}$ with $S_j = Z'_j,\ r \leq n-s,\ 1 \leq j \leq s$

and

$\mathcal{L} \approx \{X'_{s+r+1}, Z'_{s+r+1}, \ldots, X'_n, Z'_n\}$.

The gauge group $\mathcal{G}$ and the quotient group $\mathcal{L}$ induce a subsystem structure and partition the code space $C$ into subspaces $C = A \otimes B$ with $\dim(A) = 2^k$ and $\dim(B) = r$. The $n$ virtual qubits are partitioned in: $s$ - stabilizer qubits; $r$ - gauge qubits of subsystem $B$ which do not encode any useful information and are only used to absorb the transformation carried out by operators in $\mathcal{G}$; and $k$ - logical qubits, with $n = r + s + k$. The actions of the operators in the two groups $g \in \mathcal{G}$ and $L \in \mathcal{L}$ restricted to the code subspace are limited to one subsystem

$$gP = \mathbb{1}^A_{2^k} \otimes g^B,\ g^B \in \mathcal{B}(B) \qquad \text{and} \qquad LP = L^A \otimes \mathbb{1}^B_{2^{rk}},\ L^A \in \mathcal{B}(A).$$

The group $\mathcal{L}_\mathcal{B}$ of Pauli operators acting on the $r$ virtual qubits of subsystem $B$ is generated by the operators $Z'_{s+j}$ and $X'_{s+j}$ denoted as $g^z_j$ and $g^x_j$ with $1 \leq j \leq r$. The $k$ virtual qubits of subsystem $A$ are acted upon by the operators in $\mathcal{L}$, a group generated by the operators $Z'_{s+r+j}$ and $X'_{s+r+j}$ denoted as $\bar{Z}^j$ and $\bar{X}^j$ with $1 \leq j \leq k$. The gauge group $\mathcal{G}$ leaves the encode information invariant under conjugation $\mathcal{G} = \mathcal{L}_\mathcal{B} \otimes \mathcal{S} \otimes (i)$. Thus, *the standard stabilizer formalism corresponds to an Abelian gauge group, while the Operator Quantum Error Correction corresponds to a non-Abelian $\mathcal{G}$.*

[41] A quotient group is a group formed by identifying the elements of a larger group based on an equivalence relation.

Our discussion of the error correction for OQEC codes starts with an analysis of a set of correctable errors and continues with the error detection and error recovery procedures. The projector $P$ onto the code space $C$ has the property that $PC = \mathcal{H} = A \otimes B$; if $g_{ab}^B \in \mathcal{B}(B)$ is an arbitrary operator then the set of correctable errors $\{E_a\}$ must satisfy the condition

$$PE_a^\dagger E_b P = \mathbb{1}^A \otimes g_{ab}^B \ , \forall a, b.$$

When this condition is satisfied then a recovery map $\mathcal{R} : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ exists; the recovery $\mathcal{R}$ reverses the action of the transformation $\mathcal{E} : \mathcal{B}(H) \mapsto \mathcal{B}(\mathcal{H})$ on the subsystem $A$. Then $\mathcal{R} \circ \mathcal{E}(\rho^A \otimes \rho^B) = \rho^A \otimes {\rho'}^B$ with $\mathcal{E}(\rho) = \sum_a E_a \rho E_a^\dagger$.

The error detection requires the measurement of the eigenvalues of the stabilizer generators $S_j \in \mathcal{S},\ 1 \leq j \leq s$. The error syndromes can be $\pm 1$: an eigenvalue equal to $+1$ indicates that the corresponding eigenvector is in the code subspace $C$; detectable errors anticommute with at least one of the stabilizer generators, they correspond to an eigenvalue of $-1$ and *do not have any effect on the useful information.*

The operator $E_a E_b$ (the conjugation is irrelevant and it is omitted) is an element of either:

- $\mathcal{P}_n - N(\mathcal{S}) \quad \longrightarrow$ the error can be corrected; there exists $S \in \mathcal{S}$ such that $\{E_a E_b, S\} = 0$ and $PE_aE_bP = PE_aE_bSP = -PSE_aE_bP = -PE_aE_bP = 0$, or
- $N(\mathcal{S}) - \mathcal{G} \longrightarrow$ the error cannot be corrected as $PE_aE_bP = L_{ab}^A \otimes g_{ab}^B$ and $L_{ab}^A \neq \mathbb{1}^A$, or
- $\mathcal{G} \quad \longrightarrow$ the error can be corrected; $PE_aE_bP = \mathbb{1}^A \otimes g_{ab}^B$.

We conclude that an error is correctable if and only if $E_aE_b \neq N(\mathcal{S}) - \mathcal{G},\ \forall (a, b)$.

The error recovery is based on the observation that equivalent errors have the same syndrome; then the measurement of the syndrome identifies the coset of $\{E_a\}/\mathcal{G}$ of the error and we recover the information by applying any element of the coset.

We conclude this section with the observation that a stabilizer OQEC code is characterized by $[n, k, r, d]$ with: $n$, the number of physical qubits; $k$, the number of logical qubits; $r$, the number of gauge qubits, and $d$, the distance of the code. In [338] Poulin illustrates the possibility of simplifying existing $[n, k, 0, d]$ codes by identifying gauge symmetries in their stabilizers.

After the discussion of subsystems codes in Sections 5.20 to 5.23 we discuss briefly an approach for correction of systematic errors proposed by Grover.

## 5.24 Correction of Systematic Errors Based on Fixed-Point Quantum Search

In this section we overview a different approach to QEC based on Grover's fixed-point quantum search [191] discussed in Section 1.24. This strategy can be applied to the correction of systematic errors and of slow varying random errors, when the parameters of the quantum circuit do not change in time. An error is systematic if the process is repetitive and perfectly reversible.

In Section 1.24 we concentrated on unitary transformations $\mathbf{U}$ from a known initial state $\mid \gamma\rangle$ to a well-defined final state $\mid \tau\rangle$ and considered the effect of an error operator, $\mathbf{E}$. Now,

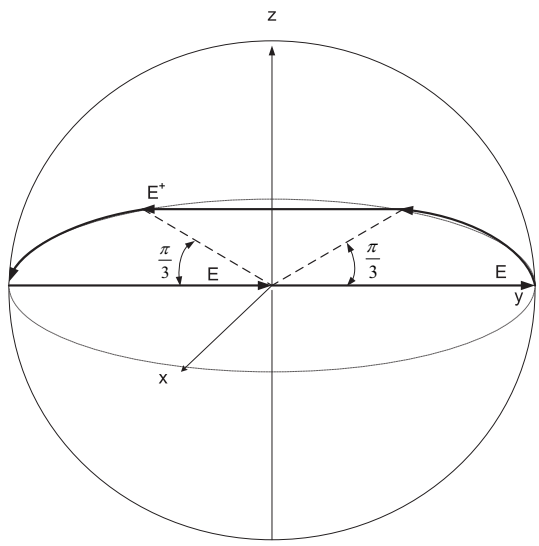



Figure 143: A Bloch sphere representation of error correction based on fixed-point search. The current state is a vector along $z$. We show the effect of the transformation $\mathbf{E}$, followed by a rotation of $\frac{\pi}{3}$ radians, $\mathbf{r}_\tau(\pi/3)$, followed by $\mathbf{E}^\dagger$, followed by another rotation of $\frac{\pi}{3}$ radians, $\mathbf{R}_\gamma(\pi/3)$, and finally followed by $\mathbf{E}$.

we do not have either a source or a target state, the input can be any arbitrary state; instead of the transformation $\mathbf{U}$ we apply $\bar{\mathbf{U}} = \mathbf{UE}$. The reversibility implies that instead of $\mathbf{U}^\dagger$ we apply $\bar{\mathbf{U}}^\dagger$. The transformation has a probability of error equal to $\epsilon$, in other words $\bar{\mathbf{U}}$ applied to an arbitrary initial state $\mid \gamma\rangle$ reaches the corresponding desired state, $\mid \tau\rangle$, with probability $(1-\epsilon)$. In this formulation it becomes clear that the fixed-point quantum search can be used to increase the probability of success from $(1-\epsilon)$ to $(1-\epsilon^3)$ when, instead of $\bar{\mathbf{U}}$, we apply the composite operation $\bar{\mathbf{U}}\mathbf{R}_\gamma\bar{\mathbf{U}}^\dagger\mathbf{R}_\tau\bar{\mathbf{U}}$. In this case $\mathbf{R}_\tau$, and respectively, $\mathbf{R}_\gamma$ rotate the state with an angle $\theta = \frac{\pi}{3}$.

The error operator $\mathbf{E}$ associated with the linear transformation $\mathbf{U}$ is restricted. The probability of error, the amount of over-rotation of the initial state vector, is small, yet unknown; it is a function of the angle $\theta$ and the axis of rotation $\bar{n}$, $\epsilon = \epsilon(\mid \theta \mid, \bar{n})$. The error operator is

$$\mathbf{E} = e^{i\epsilon\bar{n}\bar{\sigma}}$$

with $\bar{\sigma} = (X, Y, Z)$ the Pauli matrices. It is easy to observe that $\mathbf{E} = \mathbf{E}^\dagger\bar{\mathbf{E}}$ because $\mathbf{EE}^\dagger = \mathbf{I}$.

Note that $\mid \tau\rangle = \mathbf{E} \mid \gamma\rangle$ thus $\langle\tau \mid = \langle\gamma \mid \mathbf{E}^\dagger$. Then the over-rotation when we apply $\mathbf{U}$ to the arbitrary initial state is

$$\langle\tau \mid \mathbf{E} \mid \gamma\rangle = \langle\gamma \mid \mathbf{E}^\dagger\bar{\mathbf{E}} \mid \gamma\rangle = \langle\gamma \mid \mathbf{E}^\dagger\mathbf{EE} \mid \gamma\rangle = \langle\gamma \mid \mathbf{E} \mid \gamma\rangle.$$

Similarly,

$$\begin{aligned}\langle\tau \mid \bar{\mathbf{U}}\mathbf{R}_\gamma(\pi/3)\bar{\mathbf{U}}^\dagger\mathbf{R}_\tau(\pi/3)\bar{\mathbf{U}} \mid \gamma\rangle &= \langle\gamma \mid \mathbf{U}^\dagger\mathbf{UER}_\gamma(\pi/3)E^\dagger\mathbf{UR}_\tau(\pi/3)\mathbf{UE} \mid \gamma\rangle \\ &= \langle\gamma \mid \mathbf{ER}_\gamma(\pi/3)\mathbf{E}^\dagger\mathbf{R}_\tau(\pi/3)\mathbf{E} \mid \gamma\rangle.\end{aligned}$$

Reichardt and Grover [356] show that if the error is small then by concatenating the basic sequence

$$\bar{\mathbf{U}}\mathbf{R}_\tau(\pi/3)\bar{\mathbf{U}}^\dagger\mathbf{R}_\gamma(\pi/3)\bar{\mathbf{U}}$$

we obtain an even smaller error

$$\begin{aligned}
&\langle\tau \mid \mathbf{U} \mid \gamma\rangle = \langle\gamma \mid \mathbf{E} \mid \gamma\rangle \;\rightarrow\; 1 \\
&\quad\Longrightarrow\quad \langle\tau \mid \bar{\mathbf{U}}\mathbf{R}_\gamma(\pi/3)\bar{\mathbf{U}}^\dagger\mathbf{R}_\tau(\pi/3) \mid \gamma\rangle = \langle\gamma \mid \mathbf{E}\mathbf{R}_\gamma(\pi/3)\mathbf{E}^\dagger\mathbf{R}_\tau(\pi/3)\mathbf{E} \mid \gamma\rangle \;\rightarrow\; 1.
\end{aligned}$$

Figure 143 shows the effect of the transformation $\bar{\mathbf{U}}\mathbf{R}_\tau(\pi/3)\bar{\mathbf{U}}^\dagger\mathbf{R}_\gamma(\pi/3)\bar{\mathbf{U}}$. We assume that the current state $\mid \gamma\rangle$ is along the $\mid z\rangle$ axis of the Bloch sphere and show the effect of the composite transformation on the error.

Finally, we discuss an application of QEC to reliable quantum gates. Reliability of quantum systems is a vast subject and the discussion in the next section only illustrates the role played by quantum error correcting codes in the effort to build reliable quantum systems.

## 5.25 Reliable Quantum Gates and Quantum Error Correction

A reliable quantum computer can only be built with reliable quantum circuits. To design reliable quantum circuits we have to consider "logical quantum gates" performing reliably the same function as the "physical quantum gates" we are accustomed to, e.g., `CNOT` and `Toffoli` gates; the logical gates should be built using quantum error correcting codes.

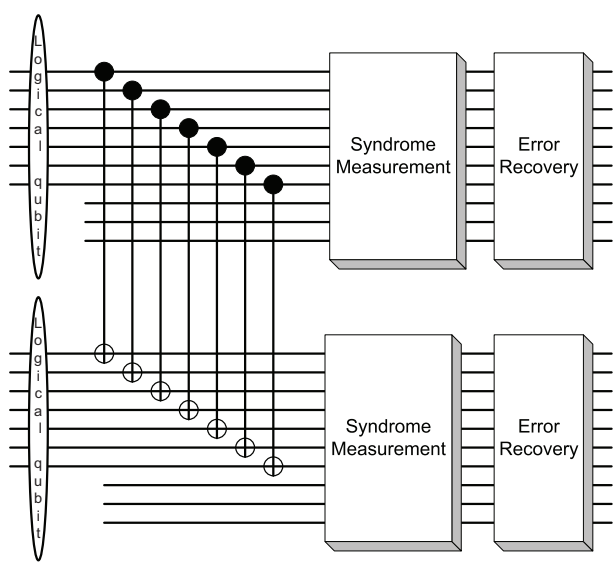



Figure 144: A reliable `CNOT` gate based on the seven-qubit quantum error correcting code of Steane. Each one of the seven pairs of physical qubits traverses a `CNOT` gate; for each logical qubit there are three ancilla qubits for syndrome calculation. The fourteen physical qubits are subject to syndrome measurement and error recovery.

Figure 144 shows a reliable CNOT gate based on the 7-qubit code of Steane. This circuit is able to correct one qubit in error after the application of the gate. This circuit is similar to the classical CNOT, it has two logical qubits and requires 7 physical qubits to encode each logical qubit, as well as 3 ancilla qubits for syndrome calculation for each logical qubit.

If an error affecting a register of quantum bits is described by a unitary transformation $\mathbf{E}$ and if $\mathbf{G}_{cnot}$ is the transformation carried out by the CNOT gate then:

$$\mathbf{G}_{cnot}\mathbf{E} = \mathbf{E}^{\dagger}\mathbf{G}_{cnot}.$$

This equation shows that it makes no difference if an error occurs before, or after the application of a CNOT gate, and it is true for other types of quantum gates as well. This important property enables us to build reliable quantum gates consisting of a possibly large number of error-prone physical gates.

If $p$ is the probability of an error affecting a single qubit what is the probability $P_e$ that the circuit in Figure 144 is affected by two or more errors? The answer to this question is critical for our ability to build a reliable CNOT gate based on a quantum error correcting code, e.g., the 7-qubit Steane's code capable to correct a single bit-flip or phase-flip error. We shall show that $P_e = kp^2$ with $k$ a constant dependent on the quantum error correcting code.

For Steane's code $k = k_s \approx 10^4$. If for a given $\epsilon$ we wish to achieve a probability of error $P_e \lll \epsilon$ then $p$, the probability of a single qubit error should satisfy the following condition:

$$p = \sqrt{\frac{P_e}{k}} \lll \sqrt{\frac{\epsilon}{k}}.$$

For example, when $k_s = 10^4$ (Steane's code) and when we require that $\epsilon \approx 10^{-10}$ then $p \lll 10^{-3}$, e.g., $p = 10^{-5}$.

Let us now discuss the procedure to evaluate $k_s$ when we assume that the probability of more than two errors in the circuit in Figure 144 is very low. This seems to be a reasonable assumption, indeed if $p = 10^{-5}$ then $p^3 = 10^{-15}$. We evaluate several possible scenarios and compute the probability $P_i = k_i p^2$ of an error caused by each scenario [288] and then compute

$$k_s = \sum_i k_i \quad \text{and} \quad P_e = k_s p^2.$$

There are several possibilities:

1. We have two independent errors, one affecting each logical qubit of the current gate. Each one of the 7 physical qubits of a logical qubit is produced by a logical gate of a previous stage. For simplicity we assume that the previous gate is also a reliable CNOT based on Steane's code, thus there are 10 possible points of failure (7 qubits plus 3 ancilla qubits). Thus, $P_1 = (k_1' p)^2 = k_1 p^2$ with $k_1 = {k_1'}^2$ and $k_1' \approx 7 \times 10 = 70$.

2. A single error enters the logical gate and an additional error occurs during syndrome calculations and error recovery. For each of the seven qubits there are $2k_1$ possible pairs of failure points and $P_2 = (k_2 p) \times p = k_2 p^2$ with $k_2 \approx 7 \times (2k_1') = 7 \times 140$.

3. Two failures occur during traversal of CNOT gates; then there are $10 \times 10$ combinations for possible points of failure $P_3 = k_3 p^2$ with $k_3 \approx 10 \times 10$.

4. Two failures occur during processing, one during the traversal of `CNOT` gates and one during the syndrome measurement; then $P_4 = k_4 p^2$ with $k_4 \approx 10 \times 10$.

5. Two failures occur during syndrome measurement; $P_5 = k_5 p^2$ with $k_5 = k_1' \times k_1' \approx 70 \times 70$.

6. One failure occurs during syndrome measurement and another failure occurs during the recovery stage and then $P_6 = k_6 p^2$ with $k_6 = k_1' \times 7 \approx 70 \times 7$.

7. Two failures occur during the recovery stage and then $P_7 = k_7 p^2$ with $k_7 \approx 7 \times 7$.

We conclude that $k_s = \sum_{i=1}^{7} k_i \approx 10^4$.

Quantum error correction is intimately related to the physical processes which cause decoherence. It is critical to have some understanding of the time scale for the decoherence phenomena for different qubit candidates. If $\tau_{gate}$ is the time required for a single gate operation and $\tau_{decoherence}$ is the decoherence time of a qubit, then $n_{gates}$, the number of gates that can traversed by a register before it is affected by decoherence is given by

$$n_{gates} = \frac{\tau_{decoherence}}{\tau_{gate}}.$$

Table 31 presents sample values of the time required for a single gate operation $\tau_{gate}$, the decoherence time of a qubit, $\tau_{decoherence}$, and $n_{gates}$, the number of gates that can be traversed before a register of qubits is affected by decoherence, for several qubit implementations [127]. We notice a fair range of values for the number of quantum gate operations that can be performed before decoherence affects the state, from about $10^3$ for quantum dots to $10^{13}$ for a trapped Indium ion.

Table 31: The time required for a single gate operation, $\tau_{gate}$, the decoherence time of a qubit, $\tau_{decoherence}$, and the number of gates that can be traversed before a register of qubits is affected by decoherence, $n_{gates}$. The qubit implementations are listed in decreasing order of decoherence time.

| Qubit implementation | $\tau_{decoherence}$ (sec) | $\tau_{gate}$ (sec) | $n_{gates}$ |
|---|---|---|---|
| Nuclear spin | $10^4$ | $10^{-3}$ | $10^7$ |
| Trapped Indium ion | $10^{-1}$ | $10^{-14}$ | $10^{13}$ |
| Quantum dots/charge | $10^{-9}$ | $10^{-12}$ | $10^3$ |
| Quantum dots/spin | $10^{-6}$ | $10^{-9}$ | $10^3$ |
| Optical cavity | $10^{-5}$ | $10^{-14}$ | $10^9$ |

*The process of error correction introduces new errors at the same time it corrects others.* The discussion in this Section shows that the error correction may, or may not, reduce the error rate during a quantum computation. The error correction will help if and only if the gate error rate is low enough, otherwise it will produce more harm than benefits. Moreover, fault tolerance may introduce additional overhead and limit the ability of a quantum computer to carry out long computations.

There is another aspect of quantum error correction we have to pay attention to, *the complexity of decoding.* We can always use a code that corrects more errors and presumably

a quantum computer based on such codes will be able to withstand decoherence for longer periods of time and carry out longer computations. However, for most error correcting codes the complexity of decoding increases rapidly with the number of errors corrected by the code. The time to do error correction for such codes may in fact be so large, that the performance of the quantum computer may decline, when the number of errors corrected by the code increases above a certain threshold.

Thus, it is important to find quantum error correcting codes where the time to measure the error syndrome and to carry out the error recovery and the complexity of the quantum circuits to carry out these operations increase slowly with the error correcting capability of the code. This problem was discussed in Section 4.18 when we showed that there exist polynomial time algorithms $\mathcal{E}' : \{0,1\}^K \mapsto \{0,1\}^N$ for encoding and $\mathcal{D}' : \{0,1\}^N \mapsto \{0,1\}^K$ for decoding such that: Prob(decoding error) $< 1/N$.

## 5.26 History Notes

In 1995 Peter Shor described a 9-qubit single error-correcting code [381]. Before this breakthrough it was not clear that quantum error correction is even possible.

A 7-qubit quantum error-correcting code was found in 1995 by Steane [402] and Calderbank and Shor [81]; this was the first code of an entire class of quantum error correcting codes, the so called CSS (Calderbank, Shor, Steane) codes. In 1996 Bennett *et al.* [40] and Laflamme *et al.* [264] discovered an even more efficient 5-qubit code.

The stabilizer formalism which provides a general theory of quantum error correcting codes was developed soon thereafter by Gottesman and others [82, 174, 343].

Emmanuel Knill, Raymond Laflamme, and Lorenza Viola introduced subsystem codes in a paper on the theory of quantum error correction for general noise [250]; David Kribs, Raymond Laflamme, and David Poulin developed Operator Quantum Error Correcting (OQEC) codes [260] and Poulin proposed a stabilizer formalism for OQEC codes [338]. David Bacon introduced a class of subsystem codes referred to as Bacon-Shor codes [17, 18].

The algorithmic approach to error correction is intimately tied to fault-tolerance. In 1996, Shor made a major contribution to fault-tolerant quantum computation on encoded states. In [383] Shor discusses the implementation of quantum gates to evolve the logical state of a quantum circuit.

Shor emphasizes the use of ancilla entangled states that are partially verified and the repetition of syndrome measurement. He also introduced a discrete universal set of quantum operations. The same year, DiVincenzo and Shor [129] generalized the fault-tolerant protocol for syndrome measurement for any stabilizer code. Daniel Gottesman also proposed fault-tolerant universal methods for stabilizer codes [174, 175].

## 5.27 Summary and Further Readings

The properties of quantum information make error correction very challenging. Quantum error correction requires a multi-qubit measurement that does not disturb the quantum information in the encoded state but retrieves information about the error. The measurement of the syndrome should reveal as much as possible about the error, but should not provide

any information about the value that is stored in the logical qubit; otherwise the measurement would destroy any quantum superposition of this logical qubit with other qubits in the quantum computer.

The measurement of the syndrome determines if a qubit has been corrupted and the type of error. The effect of the noise is either a bit-flip, or a phase-flip, or both and can be represented by transformations carried out by the Pauli operators X, Z, and Y. A puzzling question that comes to mind is how can we consider only a few distinct possibilities when the noise is arbitrary? The answer is that we perform a projective measurement of the syndrome; though the effect of the noise is arbitrary, it can be expressed as a superposition of basis operations given by the $X, Y$, and $Z$ Pauli operators and the identity, when the qubit is not affected by error. The measurement of the syndrome "forces" the qubit to "choose" a specific "Pauli error" identified by the syndrome; to correct the error we apply to the corrupted qubit the same Pauli operator.

A quantum code is denoted by $\{n, k, d_1, d_2\}$ with $n$ the number of qubits used to store or transmit $k$ bits of information and allow correction of up to $\lfloor (d_1 - 1)/2 \rfloor$ bit-flip errors and simultaneously up to $\lfloor (d_2 - 1)/2 \rfloor$ phase-flip errors. The code $\mathcal{C}$ is a subspace of dimension $2^k$ of a Hilbert space $H_{2^n}$ and it is constructed by specifying a basis of it.

The following proposition [81] relates quantum and classical error-correcting codes: *If $\mathcal{C}_1$ and $\mathcal{C}_2$ are both $[n, k, d]$ codes with $\{0\} \subseteq \mathcal{C}_2^{\perp} \subseteq \mathcal{C}_1 \subseteq \mathcal{F}_2^n$, then the quantum code $\mathcal{Q}_{\mathcal{C}_1,\mathcal{C}_1}$ is a $t$-error correcting code, with $t = \lfloor (d-1)/2 \rfloor$.*

The stabilizer formalism presented in [174] is very useful in quantum error correction. An $[n, k]$ stabilizer code can be used to encode the state of a $k$-qubit system and correct a set $\Sigma$ of error patterns. There are $(n - k)$ independent generators, $\mathbf{M}_i, 1 \leq i \leq (n - k)$ of the stabilizer group $S$, which commute with each other.

The generators, $\mathbf{M}_i, 1 \leq i \leq (n - k)$, are $n$-tuples of Pauli operators (tensor products on $n$ Pauli operators), and their eigenvectors span two orthogonal eigenspaces of dimension $2^{n-1}$, corresponding to the two eigenvalues $\pm 1$, respectively, of the generators. The stabilizer group $S$ generated by $\mathbf{M}_i,\ 1 \leq i \leq (n - k)$ has a set of $2^{n-k}$ orthogonal eigenspaces, each of dimension $2^k$. Each eigenspace corresponds to one of the $2^{n-k}$ combinations of values $\pm 1$, the eigenvalues of generators.

The *code subspace*, $\mathcal{C}$, is defined as the subspace of $\mathcal{H}_{2^n}$ stabilized by $S$; the code subspace is a $2^k$ subspace for which the $(n - k)$ generators have eigenvalues equal to $+1$.

The basic idea of subsystem codes is to use algebraic methods to classify quantum errors and, implicitly, to describe quantum error correcting codes as tensor products of subspaces, or more accurately, as subsystems where the useful information is confined to one subsystem [250].

A Hilbert space $\mathcal{H}$ can be decomposed as the direct sum of two subspaces $\mathcal{H} = \mathcal{H}^C \bigoplus \mathcal{H}^D$, or as the tensor product of two subspaces $\mathcal{H} = \mathcal{H}^C \otimes \mathcal{H}^D$; in case of a direct sum $\dim(\mathcal{H}) = \dim(\mathcal{H}^C) + \dim(\mathcal{H}^D)$ while $\dim(\mathcal{H}) = \dim(\mathcal{H}^C) \times \dim(\mathcal{H}^D)$ for the tensor product. A decomposition of the Hilbert space $\mathcal{H}$ as a subspace $\mathcal{H}^E$ and a perpendicular subspace $(\mathcal{H}^E)^{\perp}$ is possible, with $(\mathcal{H}^E)^{\perp}$ expressed as a tensor product of two other subspaces $\mathcal{H}^C$ and $\mathcal{H}^D$

$$\mathcal{H} = (\mathcal{H}^E)^{\perp} \bigoplus \mathcal{H}^E \quad \text{or} \quad \mathcal{H} = (\mathcal{H}^C \otimes \mathcal{H}^D) \bigoplus \mathcal{H}^E.$$

The subsystem encoding for this decomposition requires the preparation of the quantum state

$$\rho = (\rho^C \otimes \rho^D) \oplus 0^E$$

and encoding the information into subsystem $\mathcal{C}$ described in the Hilbert space $\mathcal{H}^C$ for a fixed encoding into the subsystem $\mathcal{D}$ described in the Hilbert space $\mathcal{H}^D$; in this expression $\rho^C$ is the density matrix of the encoded quantum information, $\rho^D$ is the density matrix of the arbitrary quantum information encoded into the subsystem $\mathcal{D}$, and $0^E$ is the all zero matrix on the subspace $\mathcal{H}^E$. $\mathcal{D}$ is called a gauge subsystem and $E$ a logical subsystem. Regardless of the nontrivial transformation applied to gauge subsystem $\mathcal{D}$, the information encoded into subsystem $\mathcal{C}$ is not affected; *the quantum error correction is restricted to the restoration of the information encoded into subsystem $\mathcal{C}$.*

A different approach to quantum error correction is taken by Lov Grover who in 2005 introduced fixed-point quantum search [190, 191] and then later discussed the application of fixed-point quantum search to error correction [356]. Grover's approach can be applied to correct small, systematic errors.

The most important readings on the subject of quantum error correcting codes are the seminal papers of Shor [381], Steane [403], Calderbank and Shor [81], Schumacher [369], Knill, Laflamme and Zurek [244], Knill and Laflamme [245]. A comprehensive presentation of the stabilizer formalism can be found in Gottesman's thesis [174] and in his later papers [177]. David Forney, Grassl, and Guha have contributed to concatenated quantum codes [156]. The relevant references for OQEC are due to Emmanuel Knill, Raymond Laflamme, and Lorenza Viola [250] David Kribs, David Poulin, [260], [338] and David Bacon [17, 18]. Fixed-point quantum search is discussed by Grover in [191] and Grover and Reichardt in [356].

## 5.28 Exercises and Problems

**Problem 1.** Let $\mathcal{C}$ be a linear $[n, k, d]$ code, capable of correcting $\lfloor \frac{d-1}{2} \rfloor$ errors and $\mathcal{C}^\perp$ its dual. Assume that $\mathcal{C}^\perp \subset \mathcal{C}$. Let

$$\mid s_v\rangle = 2^{-\frac{n-k}{2}} \sum_{w \in \mathcal{C}^\perp} \mid v \oplus w\rangle$$

Prove that given $(\mid v\rangle, \mid u\langle) \in \mathcal{C}$ the two vectors are identical, $\mid s_v\rangle = \mid s_u\rangle$, if and only if $\mid v\rangle = \mid u\rangle$.

**Problem 2.** Show that the generators of a group provide a more compact description of the group than the enumeration of all the elements of the group.

**Problem 3.** Show that in the case of the nine qubit error correcting code any error operator $\mathbf{E}^i_{x|y|z}$, $1 \leq i \leq 9$ anticommutes with at least one of the eight operators $\mathbf{M}_j$, $1 \leq j \leq 8$, thus any single qubit error can be corrected (see Section 5.12).

**Problem 4.** Show that all elements of the Pauli group $\mathcal{P}$ are: unitary, either Hermitian or anti-Hermitian, and either commute or anticommute.

**Problem 5.** Show that $S$ is a normal subgroup of $N(S)$.

**Problem 6.** Show that $\mid NS \mid = 4 \cdot 2^{n+k}$. *Hint:* The Pauli group $\mathcal{P}_n$ has $4^{n+1}$ elements including an overall phase of $\pm 1$ and $\pm i$. The factor of 4 accounts for the overall phase factor.

**Problem 7.** Show that it makes no difference if a qubit error occurs before or after the application of a `CNOT` gate. If $\mathbf{A}$ is a unitary transformation and $\mathbf{G}_{CNOT}$ is the transformation carried out by the `CNOT` gate show that:

$$\mathbf{G}_{CNOT}\mathbf{A} = \mathbf{A}^{\dagger}\mathbf{G}_{CNOT}.$$

**Problem 8.** Show that $S$ is a stabilizer of a non-trivial Hilbert subspace $V_{2^n} \subset \mathcal{H}_{2^n}$ only if: $S = \{\mathbf{S}_1, \mathbf{S}_2, \ldots\}$ is an Abelian group:

$$\mathbf{S}_i\mathbf{S}_j = \mathbf{S}_j\mathbf{S}_i, \quad \forall \mathbf{S}_i, \mathbf{S}_j \in S \quad i \neq j.$$

Note: In Section 5.12 we showed that this condition is necessary.

**Problem 9.** Show that error operators for Steane's seven-qubit code anticommute with the generators listed in Table 23 from Section 5.12.

**Problem 10.** Establish the correspondences between the errors and the syndromes for Steane's seven-qubit code, see Table 24 from Section 5.12.

**Problem 11.** Show that error operators for the perfect quantum code discussed in Section 5.13 anticommute with the generators listed in Table 28. .

**Problem 12.** Establish the correspondences between the errors and the syndromes for the perfect code discussed in Section 5.13; see Table 27.

**Problem 13.** Construct a table showing the error operators which anti-commute with each one of the four generators for the quantum restoration circuit in Section 5.14.

**Problem 14.** Establish the correspondences between the errors and the syndromes for the quantum restoration circuit in Section 5.14, see Table 29.

**Problem 15.** Show that the following equalities given in the discussion of the example in Section 5.22 hold:

$$P_{11}E_0P_{11} = \alpha P_{11}, \quad P_{11}E_1P_{11} = \beta P_{11}, \quad P_{11}E_0P_{22} = 0P_{12}, \quad P_{11}E_1P_{22} = 0P_{12},$$

and

$$P_{22}E_0P_{22} = P_{22}, \quad P_{22}E_1P_{22} = 0P_{22}, \quad P_{22}E_0P_{11} = 0P_{21}, \quad P_{22}E_1P_{22} = 0P_{22}.$$

*Hint:* Show first that:

$$\alpha_1 \otimes \beta_1 = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}}, \quad \alpha_1 \otimes \beta_2 = \frac{\mid 00\rangle - \mid 11\rangle}{\sqrt{2}}$$

and

$$\alpha_2 \otimes \beta_1 = \frac{\mid 10\rangle + \mid 01\rangle}{\sqrt{2}}, \quad \alpha_2 \otimes \beta_2 = \frac{\mid 10\rangle - \mid 01\rangle}{\sqrt{2}}.$$

Compute the matrix units for this decomposition using the expression

$$P_{kl} = \mid \alpha_k \rangle \langle \alpha_l \mid \otimes \mathbb{1}^{B_2} = \mid \alpha_k \rangle \langle \alpha_l \mid \otimes \left( \mid \beta_1 \rangle \langle \beta_1 + \mid \beta_2 \rangle \langle \beta_2 \mid \right), \ 1 \leq k, l, \leq 2.$$

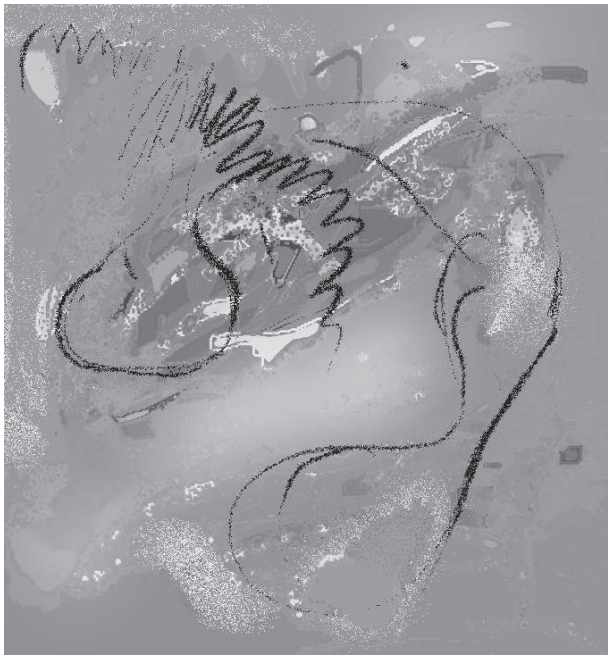
Then,

$$P_{11} = \mid 00 \rangle \langle 00 \mid + \mid 11 \rangle \langle 11 \mid, \quad P_{12} = \mid 00 \rangle \langle 10 \mid + \mid 11 \rangle \langle 01 \mid$$

and

$$P_{21} = \mid 10 \rangle \langle 00 \mid + \mid 01 \rangle \langle 11 \mid, \quad P_{22} = \mid 10 \rangle \langle 10 \mid + \mid 01 \rangle \langle 01 \mid .$$

‘Reality is merely an illusion, albeit a very persistent one.” Albert Einstein.



# 6 Physical Realization of Quantum Information Processing Systems

Throughout the previous chapters we limited ourselves to an abstract view of a quantum computer as a device suitable to conduct a special class of thoughts or gedanken experiments in computing and communication. It is now the time to address the question of the physical implementation of quantum computing and quantum information theory concepts discussed earlier. We wish to see if indeed it is feasible to build quantum computers and, if so, what type of technological advances and scientific breakthroughs should occur before the powerful concepts presented earlier could have a practical impact.

The reason we deferred this topic to the very end of our journey is two fold. First, several physical implementations are possible and at this time it is very difficult to assess in full the relative merits and potential of each of these approaches. Significant progress is reported in each of these areas, but a real breakthrough has not been achieved in any of them. Second, the physical phenomena involved in each of these areas are very complex and non-trivial to explain in detail.

The basic problems facing a physical implementation of a quantum computer are related to finding a system which is sufficiently insulated from unwanted interactions with the

environment but, at the same time, can be controlled through some required interactions implementing logical quantum gates, and can be measured.

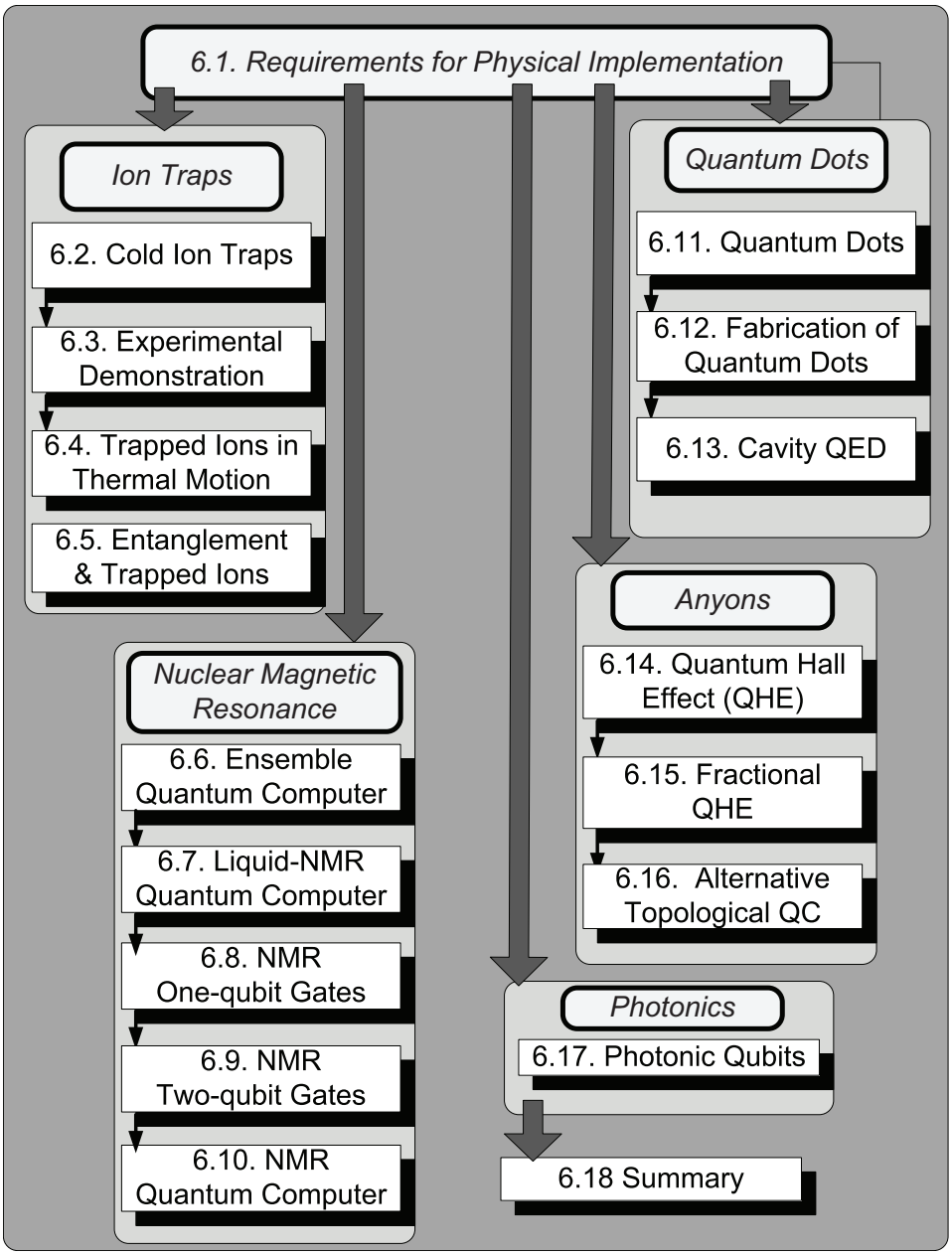



Figure 145: Chapter 6 at a glance.

So far, there have been several serious proposals for quantum computing and quantum communication physical implementations and they fall into several basic categories [474]: (i) nuclear magnetic resonance systems, (ii) ion traps, (iii) neutral atoms, (iv) optical systems, (v) mesoscopic superconducting systems with tunnel junctions, (vi) solid state systems (quantum dots, impurities). Among other unique qubit implementations we mention incompressible higher order fractional quantum Hall states with non-Abelian braiding statistics which have

been suggested as a good candidate for topological quantum computing.

Figure 145 summarizes the organization of this chapter which starts with an analysis of the requirements for a physical implementation of a quantum information processing systems. From the multitude of proposals for the physical implementation of quantum information processing devices we have selected only a subset. First, we discuss the cold ion traps, then the liquid NMR; we continue with quantum dots, and with an in depth discussion of the quantum Hall effect and conclude with a summary presentation of photonics.

## 6.1 Requirements for Physical Implementations of Quantum Computers

Quantum computing and communication devices must satisfy a set of generic requirements; these requirements guide the implementation efforts for quantum information processing systems. DiVincenzo formulated several criteria for realization of quantum computers in 2000; a decade later, in 2010, these criteria were reformulated and generalized for systems with low level of decoherence.

**DiVincenzo's criteria.** In his seminal paper on the physical implementation of quantum computers [131] DiVincenzo formulated five criteria related to quantum computation plus two additional ones required for transfer of information:

(1) *The physical system has well characterized qubits and is scalable.* The embodiment of a qubit is a quantum two-level system, such as the two spin states of a spin-1/2 particle, or the ground and some excited state of an atom, or the vertical and horizontal polarization of a single photon. The qubits can be in superposition states and/or entangled states and these unique quantum mechanical characteristics are associated with the enormous computational power of quantum computing devices. These characteristics can be effective only if the qubits are well isolated from the environment; the process of decoherence due to the possible interaction with the environment must be very slow.

The physical parameters of a well characterized qubit must be accurately known and include: the internal Hamiltonian (it determines the energy eigenstates of the qubit); the presence of other states of the physical qubit besides the two characteristic (basis) states and the couplings among them; the interactions with other qubits; and the couplings to external fields that might be used to manipulate the qubit. In general, the qubit has other (higher) energy levels than the two basis states and the probability of the transitions to such states from the characteristic states must be very low and under the control of the physical system.

Some of the proposed qubit implementations are based in atomic physics such as, pairs of energy levels of ions confined in a linear electromagnetic trap [97], Zeeman-degenerate ground states used in the NMR approach [105], atomic energy levels of neutral atoms [76, 219]. Other qubits are implemented as the position of atoms in a trap or lattice, the presence or absence of a photon in an optical cavity, or the vibrational quanta of trapped electrons, ions or atoms.

Some of the physical implementations based on solid state, such as impurities in solids and the quantum dots take advantage of the fact that they have well characterized discrete energy levels that can be controlled through manufacturing processes. The qubits in solid state systems include spin states or charge states of quantum dots, quantized states of superconductive devices, Cooper-pair charge, or the flux.

(2) *The physical system must be able to have the state of the qubits initialized to a known low-entropy state, such as* $\mid 000 \ \ldots\rangle$, as in classical computation. The initial state of the qubits can be prepared by "cooling" which can be either natural cooling (when the state of interest is the ground state of the qubit Hamiltonian), or by projecting the system into the state of interest through the appropriate measurement. The natural cooling method is used in the case of electron spin resonance techniques: the electrons are placed in a strong magnetic field and their spins are allowed to align with it while interacting with a heat bath (natural thermalization). The state projection method is associated with laser cooling techniques which are applied to cooling ion states to near the ground state in a trap [97] and are usually accompanied by fluorescence detection used to measure the state of these ions (state projection). The cooling times by state projection could be much shorter than by natural relaxation. The state initialization time becomes an important factor if this time is relatively long compared with the gate operation time and/or repeated initialization is required by say, quantum error correction.

(3) *The decoherence times of the physical system implementing the qubit must be much longer than the quantum gate operation time.* The decoherence of a quantum system is due to thermodynamically irreversible interactions with the environment; it represents the principal mechanism for the transition from quantum to classical behavior. The evolution of a quantum system in contact with its environment is characterized by various decoherence times; each decoherence time is related to a different degree of freedom of the system.

The decoherence times relevant for a quantum computer are associated with the degrees of freedom that characterize the physical qubits (see Table 31 in Chapter 5); they also depend on the specifics of the qubits' couplings to these degrees of freedom. For example, the decoherence time of the spin of an impurity in a perfect semiconductor depends on its location, whether it is in the bulk of the solid or near the surface of the device, and how close it is to the structures used to manipulate its quantum state [131]. The decoherence times must be long enough to allow the quantum system to evolve undisturbed until the computation is complete. No physical system can be completely free of decoherence, but decoherence times of the order of $(10^4 - 10^5) \times$ (*gate operation time*) are considered acceptable for fault-tolerant quantum computation [42], in which case the error probabilities are lower than a critical threshold. The reality is that quantum systems with long decoherence times of the order mentioned above are relatively hard to find.

(4) *A physical system as the embodiment of a quantum computer must have available a "universal" set of quantum logic gates.* A quantum computer executes a sequence of unitary transformations $U_1, \ U_2, \ U_3, \ \ldots \ , U_n$, as specified by a quantum algorithm, with each transformation acting on one, two, or, at most three qubits. All these unitary operations are implemented by a "universal" set of quantum gates; a convenient set of universal quantum gates contains one two-qubit gate plus a set of one-qubit gates.

The physical implementation of a quantum algorithm requires (1) identifying the Hamiltonians which generate the unitary transformations specified by the algorithm, given that

$$U_1 \ = \ e^{it\mathbf{H}_1/\hbar}, \quad U_2 \ = \ e^{it\mathbf{H}_2/\hbar}, \quad U_3 \ = \ e^{it\mathbf{H}_3/\hbar}, \quad \ldots \quad U_n \ = \ e^{it\mathbf{H}_n/\hbar};$$

[42] *Gate operation time*, the time for the execution of an individual quantum gate represents the *clock time* of a quantum computer.

(2) designing the physical apparatus so that the interaction Hamiltonians $\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \ldots, \mathbf{H}_n$ are successively turned on at precise moments of time and for predetermined lengths of time. In reality, only some types of Hamiltonians can be turned on and off as requested by a quantum computation. Most of the physical implementations proposed so far consider only two-body interactions, i.e., two-qubit unitary transformations (gates) of the type `CNOT`. As we have seen in Section 1.17, three-qubit unitary transformations can be expressed in terms of sequences of one- and two-qubit interactions [16]. In case two-qubit interactions are not easily achievable, the `CNOT` gate can be constructed using the appropriate sequence of one-qubit interactions.

*One-qubit gates* act on either the phase or the excitation of a single qubit. A phase gate performs a unitary transformation of the type

$$U_z^{(1)}(\varphi) = e^{-i\varphi\sigma_z}.$$

The state vector executes a rotation with angle $\varphi$ about the $z$-axis in a Bloch sphere representation; the characteristic states become

$$\begin{array}{ll} \mid 0\rangle & \rightarrow e^{i\varphi/2} \quad \mid 0\rangle, \\ \mid 1\rangle & \rightarrow e^{-i\varphi/2} \quad \mid 1\rangle. \end{array}$$

An excitation gate performs the unitary transformation

$$U_x^{(1)}(\theta) = e^{-i\theta\sigma_x}.$$

The state vector executes a rotation with angle $\theta$ about the $x$-axis in a Bloch sphere representation; the characteristic states become

$$\begin{array}{rl} \mid 0\rangle \rightarrow & \cos\theta \mid 0\rangle - \ i \ \sin\theta \mid 1\rangle, \\ \mid 1\rangle \rightarrow & -i \ \sin\theta \mid 0\rangle + \ \cos\theta \mid 1\rangle. \end{array}$$

*Two-qubit gates* carry out controlled transformations of the second qubit state conditioned by the state of the first one. The controlled transformations can be of two types:

(1) `Controlled-NOT` (`CNOT`):

$$U_{\texttt{CNOT}}^{(2)} = \mid 0\rangle\langle 0 \mid \otimes I + \mid 1\rangle\langle 1 \mid \otimes \sigma_x.$$

The gate changes the states of the two qubits in the well-known manner

$$\begin{array}{ll} \mid 0,0\rangle \rightarrow \mid 0,0\rangle, & \mid 1,0\rangle \rightarrow \mid 1,1\rangle, \\ \mid 0,1\rangle \rightarrow \mid 0,1\rangle, & \mid 1,1\rangle \rightarrow \mid 1,0\rangle. \end{array}$$

(2) `Controlled-Phase` (`C-Phase`):

$$U_\pi^{(2)} = \mid 0\rangle\langle 0 \mid \otimes I - \mid 1\rangle\langle 1 \mid \otimes \sigma_x.$$

The states of the two qubits are transformed as follows

$$\begin{array}{ll} \mid 0,0\rangle \rightarrow \mid 0,0\rangle, & \mid 1,0\rangle \rightarrow \quad \mid 1,0\rangle, \\ \mid 0,1\rangle \rightarrow \mid 0,1\rangle, & \mid 1,1\rangle \rightarrow \ - \mid 1,1\rangle. \end{array}$$

Here are two examples of universal sets of gates:

$$S_1 : \{U^{(2)}_{\texttt{CNOT}}, U^{(1)}_x(\pi/4), U^{(1)}_z(\varphi),\ \varphi \in [0, 2\pi]\}$$

$$S_2 : \{U^{(2)}_{\pi}, U^{(1)}_z(\pi/4), U^{(1)}_x(\varphi),\ \varphi \in [0, 2\pi]\}$$

The time profile of a two-qubit interaction is usually described as a square pulse. In reality, the time profile is determined by requirements of each specific implementation of the interaction Hamiltonian. For example, when the interaction Hamiltonian could also couple the qubit state to other energy levels (states) of the quantum system (other than the levels involved in the embodiment of the qubit), then the desired transformation can be obtained by turning on and off the interaction slowly enough to make it compatible with an adiabatic approximation; the front and tail of the pulse are far from vertical. The adiabatic condition is satisfied if the duration of the pulse front is long enough.

The time duration of the interaction pulse is controlled by the maximum size attained by the matrix elements of $\mathbf{H}(t)$ since the quantum gate action is determined only by the time integral $\int \mathbf{H}(t)dt$. The maximum size of the matrix elements can be determined by fundamental constraints, such as the requirement that the linear approximation should be applicable to the system, and by practical constraints, such as the laser power that can be delivered to a particular ion. In the end, the clock time of the quantum computer is determined by the time interval needed to ensure that two consecutive pulses have negligible overlap.

The time profile of the interaction Hamiltonian $\mathbf{H}(t)$ is physically controlled by "classical means", such as the intensity of a laser beam, the value of the gate voltage, the current intensity in a wire, which are also quantum mechanical in nature at the atomic level. The fact that these control devices behave classically means that there should be no entanglement between their quantum states and the states of any physical implementation of a quantum computer.

For some physical implementations, such as the ion trap, there is no direct interaction between the atomic-level qubits. The two-qubit interaction is mediated by a special quantum subsystem which can interact with each of the qubits involved in a controlled gate: in the case of the ion trap the collective vibrational state of the ion chain in the trap controls the interaction between qubits. Such auxiliary quantum systems introduce new sources of decoherence.

The time profile of the two-qubit interaction in an ion trap is influenced by the timing of the mediating subsystem and the decoherence effects occurring during gate operation. The fact that in some implementations there is a single subsystem that mediates each two-qubit interaction can affect the fully parallel gate operations required for successful error correction. We have to expect that in any physical implementation the quantum gates could not be implemented perfectly and the realization of the associated Hamiltonians would be accompanied by systematic errors (avoidable through careful calibration) and random errors (correctable with different techniques).

(5) *The physical system must have a qubit-specific measurement capability.* At the end of a computation the result is read out by measuring specific qubits. The measurement represents an interaction between qubits and the measurement apparatus and this is an irreversible process. Realistic measurements of quantum systems are expected to have very low quantum efficiencies ($\ll 100\%$); the probability to obtain a certain value for an output qubit as a result

of one measurement is very low. The efficiency of a measurement of a quantum computer output could be "increased" either by rerunning the computation for a number of times, or by copying the value of a single read-out qubit [43] to several ancilla qubits using `CNOT` gates and measuring each of them. In the case of a non-demolition measurement a quantum efficiency of 1% could ensure reliable quantum computation if compounded with hundreds of "copy and remeasure" operations. For example, the bulk model of NMR implementation proposed [105] to perform successful quantum computations when using quantum efficiencies much lower than 1% and macroscopic quantities of copies of the same qubit (a large number of a certain molecule in solution). The final non-demolition measurement was done as an ensemble average over the whole macroscopic sample, while individual qubits were left almost undisturbed. If the measurement were performed on a time scale of $10^{-4}$ of the decoherence time, its repeated execution during the quantum computation could simplify the process of error correction.

According to DiVincenzo, physical systems involved in both quantum communication and quantum computation must satisfy two additional requirements relevant for the task of transmitting intact qubits from one place to another:

(6) *The system must have the ability to interconvert stationary and flying qubits.* The term "flying qubits" refers to the physical qubits that are transmitted from place to place while the "stationary qubits" represent the physical qubits for reliable local computation. The qubits encoded either in the polarization or in the spatial wave function of photons are the preferred flying qubits at this time. The light transmission through optical fibers is a well developed technology and reliable enough for the transmission of qubits even at relatively long distance. The unloading of the qubits from the quantum computer into a the transmission system and loading them up through a reversed process presents technical difficulties.

(7) *The flying qubits must be transmitted between specified locations without affecting their state.* The preservation of the photon quantum state during transmission through optical fibers or through atmosphere has been the main concern of many experiments in quantum cryptography [216], but not of quantum computing yet.

The physical systems under consideration at this time as quantum computing and information processing devices fulfill most of these requirements, but not all of them. The systems based on trapped ions (atoms) and on cavity quantum electrodynamics where the qubits are ions or atoms and their manipulation is performed with laser light have the advantage that quantum phenomena can be observed very clearly; measurements performed with such systems (cavity-QED is at the core of atomic clocks) are the most precise that exits at this time.

Experimentalists have been able to implement certain quantum gates [298, 270], to create entanglement with trapped cold ions [364, 271, 193], with cold atoms in optical lattices [282], to entangle up to six atoms [272], to demonstrate quantum teleportation with ions [316], to execute quantum error correction. The disadvantage of these systems is the difficulty to scale them up for computations with a large number of atoms.

The solid state systems proposed as physical implementations have the advantage of being easily scalable and of using well perfected manufacturing technologies. But their qubits

[43] The state of a qubit cannot be copied, but the value of a qubit in a given basis can be copied.

(quantum dots or Cooper pairs) are difficult to protect against decoherence due to interactions with other atoms, with impurities, or phonons[44] present in the solid lattice environment. One- and two-qubit gates have been realized experimentally with quantum dots [277, 379].

In nuclear magnetic resonance systems the qubits are atoms within the same molecule and the nuclear magnetic resonance (NMR) technique is used to manipulate them. The NMR technique is well established and the state of the qubits can be controlled with precision. As a proof of concept, Shor's algorithm was implemented on a system with seven qubits (five $^{19}$F and two $^{13}$C spin-1/2 nuclei) in a quantum computer molecule, the perfluorobutadienyl iron complex and the number 15 was factorized [94]. However, the NMR systems have their disadvantages: the cooling of the molecules is necessary in order to preserve the benefits of quantum computation and scaling up to a large number of atoms is limited by the difficulty of engineering molecules with larger and larger number of atoms (qubits).

Entanglement has also been realized with photonic systems using parametric down-conversion and conditional detection [319].

**General criteria for systems with low level decoherence.** The original criteria of DiVincenzo cannot be easily applied to some of the new emerging concepts; they can be rephrased [263] into three more general criteria based on the assumption that these can be achieved if the decoherence is kept at a small enough level. These new criteria are (i) scalability, (ii) universal logic, and (iii) correctability.

(i) For any quantum system scalability is achieved by adding new, well characterized qubits. The qubit entanglement, allowed by quantum mechanics, makes the logic space available to a quantum system of $N$ spin qubits very large; it is described by a group known as $SU(2^N)$. The space of $N$ unentangled spins is described by a much smaller group, $SU(2)^{\otimes N}$. At the same time, the size and energy consumption of a quantum computer grows linearly with the number of qubits, $N$. The scalability of a technology depends on the scalability of the classical physical systems used to define the qubits; they may be microchips, lasers, microwave electronics, cryogenic refrigerators, high vacuum systems.

(ii) Universal logic. This criterion requires the system to have a finite set of control operations. In the case of qubits, the set of universal logic gates may include nearly analogue single-qubit gates (such as spin-qubit arbitrary rotations) and any digital two-qubit entangling logic operation, such as the `CNOT` gate.

The standard "circuit model" for quantum computing has been shown to have some promising alternatives such as the adiabatic and the cluster-state quantum computation.

An adiabatic algorithm [146] follows the continuous-time evolution of a quantum system governed by a smoothly-varying time-dependent Hamiltonian. In the adiabatic quantum computation, the answer to the calculation is defined as the ground state [295] of a whole network of interactions between qubits. By slowly turning on those interactions, the qubits are being adiabatically evolved from an initial Hamiltonian with a ground state easy to construct into the final Hamiltonian whose ground state is the solution of the computation. The state

[44] Phonons are quanta of energy exchanged by collective vibrational modes in periodic solids. The atoms in the regular lattice of a solid material are connected through bonds which only allow them to vibrate about equilibrium positions like quantum harmonic oscillators. The atoms can not vibrate independently and their vibrations take the form of collective modes which propagate through the material and exchange energy only in discreet amounts. Like the photons of electromagnetic energy, the phonons obey Bose-Einstein statistics.

of the system will be very close to the desired final state if the evolution time is long enough; the required time depends on the energy difference between the two lowest states of the interpolating Hamiltonian. An adiabatic quantum computation does not require fast quantum logic operations and measurements. The universal logic criterion applies in this case with regard to the complexity of the available set of interactions, the length of time it takes to turn on those interactions and how low is the temperature the system must be kept at.

The other alternative, the cluster-state quantum computation [350] starts in one particular entangled quantum state, the so-called cluster state, of a large number of qubits; this represents the entire resource for the quantum computation. The information is written onto the cluster, processed by propagating it through the cluster. The computation is performed by changing the way in which the resulting wave function is measured; the qubits are measured in arbitrary bases and that provides the analogue component required by universal logic. The measurements are used to imprint a quantum circuit on the cluster state, but they are destroying its entanglements at the same time. The result of the computation is read out of the cluster using only one-particle measurements. The implementations for newer concepts, such as the adiabatic and the cluster state quantum computers could be simpler for some technologies; they are equivalent in power to gate-based quantum computers [351, 295].

(iii) Correctability requires that any quantum error correction protocol devised for any physical implementation should be able to maintain the desired state of the system by eliminating unwanted entropy introduced from the environment into the quantum computer while adding/droping qubits required by encoding/decoding operations. That can be done through a combination of efficient state initialization and measurement. The initialization process cools the quantum system into a low entropy state. The measurements help determine the state of the quantum system quickly and with the accuracy allowed by quantum mechanics. A quantum non-demolition measurement combines these two operations: it initializes the quantum system by projecting it to the measured (basis) state and this state remains the same even after repeated measurements.

Actual quantum computers must have the capability of simultaneously controlling the quantum systems, measuring them, and switching between these operations, while preserving the quantum systems in high isolation from their environment.

Many physical systems proposed as candidates for a quantum information processing device, such as trapped ions, optical photons, quantum dots, cavity quantum electrodynamics, show some promise in preserving quantum coherence long enough to allow the execution of the basic tasks of such a device.

## 6.2 Cold Ion Traps

The cold ion trap was the first technology proposed for quantum information processing devices; it still is very promising and very popular. A set of $N$ cold ions interacting with laser light and moving in a linear trap was proposed by I. Cirac and P. Zoller [97] and was demonstrated experimentally by C. Monroe *et. al.* [298] as a realistic physical system to implement a quantum computer. The ion traps use laser-induced coupling between the electron state and the vibrational motion of relatively heavy charged particles, ions, trapped in vacuum at very low temperature.

The system consists of $N$ spin-1/2 ions confined in the electrostatic field engineered inside the linear trap. The field is supposed to be sufficiently anisotropic and the ions sufficiently cold that they should lie "crystalized", aligned in a "string-of-pearls" formation along the longitudinal axis of the trap where the effective trapping potential is the weakest. This string of ions is the "register" of the quantum computer. The ions are laser (Doppler) cooled in all three dimensions to very low temperature of a few mK; at this point, they undergo very small oscillations around the equilibrium position. Their confined motion along the $x$-, $y$-, and $z$-directions can be described in terms of "normal" modes; each normal mode can be treated as an independent harmonic oscillator of frequency $\omega_x \ll \omega_y$, or $\omega_x \ll \omega_z$. There are a total of $n$ such normal modes along the $x$-axis (the weak axis). The state of a string of $N$ qubits in the trap can be expressed as

$$\mid q_1, q_2, \ldots, q_j, \ldots, q_N\rangle \mid n\rangle,$$

where the first term, $\mid q_1, q_2, \ldots, q_j, \ldots, q_N\rangle$, represents the states of the logical qubits ($q_j = 0$ or $q_j = 1$) and the second one, $\mid n\rangle$, represents the normal-mode vibrational state; the value of $n$, such as $n = 0, 1, 2, \ldots$ refers to the number of phonons present in the collective mode. For example, when $n = 0$ (no phonon in the normal mode), the ions are not vibrating and the state is

$$\mid q_1, q_2, \ldots, q_j, \ldots, q_N\rangle \mid 0\rangle.$$

In the lowest mode $n = 1$ (one phonon present in the normal mode), the center-of-mass (CM) mode, the ions oscillate in lock step in the $x$-direction, the trap axis, with a frequency $\omega_1$ that coincides with $\omega_x$ of the harmonic trapping potential ($\omega_1 = \omega_x$) and the state is

$$\mid q_1, q_2, \ldots, q_j, \ldots, q_N\rangle \mid 1\rangle$$

Generally, the modes are numbered in order of increasing resonance frequency. The resonant frequencies of the ions' motion are of the order of a few Megahertz, the wavelengths of the resonant radiation are larger than 100 meter; as the separation of the ions is of the order of $10\mu$m, the assumption that the external electric field is constant over the region occupied by the ion string is a good approximation. The normal mode degree of freedom is used as a "bus" to couple different ions together.

In the ion trap system, the ions themselves are the qubits. The two characteristic states of any qubit $j$, $(j = 1, \ldots, N)$ are identified with two internal states of the corresponding ion, such as the *ground state* $\mid g\rangle_j \equiv \mid 0\rangle_j$ and a particular long-lived metastable, non-degenerate, *excited state* $\mid e\rangle_j \equiv \mid 1\rangle_j$. The quantum state of each ion is a linear combination of the ground state $\mid g\rangle$ and the metastable excited state $\mid e\rangle$. A coherent linear combination of the two levels

$$\alpha_0 \mid g\rangle_j + \alpha_1 e^{i\omega t} \mid e\rangle_j$$

can survive for a time comparable to the lifetime of the excited state. The relative phase oscillates because of the energy splitting $\hbar\omega_0$ between the levels.

**Qubit manipulation in an ion trap.** Individual qubits are independently manipulated by shinning a different pulsed laser beam to each of the ions. If the laser beam directed to a particular ion is tuned to the frequency $\omega_0$ of the transition $\mid g\rangle \rightarrow \mid e_0\rangle$, then a periodic

exchange of energy can occur between the electromagnetic light field and the two-level system. These periodic exchanges are called *Rabi oscillations*[45]; they are induced between the qubit state $\mid 0\rangle$ and $\mid 1\rangle$. During the Rabi cycle the atom/ion is usually in a superposition of ground and excited state. Any one-qubit unitary transformation can be applied and any desired linear combination of $\mid 0\rangle$ and $\mid 1\rangle$ can be prepared by selecting an exact duration of the laser pulse and by appropriate choice of the laser phase.

The ions (qubits) can be read out by performing a measurement that projects their state onto the $\{\mid g\rangle, \mid e_0\rangle\}$ basis. That can be done in the following way: a laser is tuned to a transition from the ground state $\mid g\rangle$ to a short-lived excited state $\mid e'\rangle$, different from the metastable excited state $\mid e_0\rangle$. When the laser beam shines on the ions, each qubit with the value $\mid 0\rangle$ absorbs and re-emits repeatedly the laser light by oscillating between the $\mid g\rangle$ and the $\mid e'\rangle$ levels. The ions emit fluorescent light by de-exciting from state $\mid e'\rangle$ to the ground state; in this way qubits in state $\mid 0\rangle$ become "visible". The qubits in state $\mid 1\rangle$ remain "dark"; they are in the metastable excited state $\mid e_0\rangle$ which does not fluoresce. The other (excitation) levels of the ion are not affected by the laser light and do not disturb the computation process; some of them may be needed for implementing quantum gates. The laser acting on one ion will leave the internal states of all the other ions unaffected.

**Single-qubit rotations.** Rotations of a single qubit, $j$, can be performed by tuning the laser frequency on resonance with the internal transition between ground and excited state of ion $j$, $\mid g\rangle_j \rightarrow \mid e_0\rangle$ (polarization denoted $q = 0$). The Hamiltonian for this interaction is

$$\mathbf{H}_j = \frac{\Omega}{2}(\mid e_0\rangle_j \; {}_j\langle g \mid e^{-i\phi} + \mid g\rangle_j \; {}_j\langle e_0 \mid e^{i\phi}),$$

where $\Omega$ is the Rabi frequency[46] and $\phi$ is the laser phase. If the laser shines on the ion for an interaction time $t = (k\pi/\Omega)$, where $k\pi$ is the length of the laser pulse, with $k$ the laser wave vector, the evolution of the qubit is described by the unitary operator

$$\mathbf{V}_j^k(\phi) = e^{-i\mathbf{H}t} = \exp\left[-ik\frac{\pi}{2}\;(e^{-i\phi} \mid e_0\rangle_j \; {}_j\langle g \mid + \; h.c.)\right],$$

where *h.c.* stands for Hermitian conjugate terms. As a result, the qubit performs a Rabi rotation and the two internal pure states evolve into superpositions

$$\mid g\rangle_j \rightarrow \cos(\frac{k\pi}{2}) \mid g\rangle_j - ie^{i\phi}\sin(\frac{k\pi}{2}) \mid e_0\rangle_j$$

$$\mid e_0\rangle_j \rightarrow \cos(\frac{k\pi}{2}) \mid e_0\rangle_j - ie^{-i\phi}\sin(\frac{k\pi}{2}) \mid g\rangle_j.$$

[45] *Rabi oscillations* occur when laser light interacts with a two-level atomic system, such as the ground state and the excited (nondegenerate) state of an atom/ion. They are associated with oscillations of the quantum-mechanical expectations values of the level populations in the ground and excited state and of the photon numbers. That can be interpreted as a periodic change between absorption and *stimulated emission* of photons. *Spontaneous emission* of photons, a competing process, is also possible and can prevent the observation of Rabi oscillations. Rabi oscillations can be described with two different models: the Rabi model which assumes a quantum description of the atom and a classical description of the field and the Jaynes - Cummings model which assumes quantum descriptions of both atom and field.

[46] *Rabi frequency* ($\Omega$) is the rate at which the oscillations occur and is proportional to the amplitude (but not the intensity) of the light field and to the dipole moment of the level transition.

If the ion is in its ground state (qubit state $\mid 0\rangle$), a laser $\pi$-pulse will make the electron wave function evolve through half of a Rabi oscillation period and leave the ion in its excited state (qubit state $\mid 1\rangle$); the qubit rotates from the $\mid 0\rangle$ to the $\mid 1\rangle$ state. When the laser pulse duration is reduced by a factor of 2, a $\pi/2$-pulse, the ion is left in an equal superposition of the ground and excited states; the qubit rotates to the $\frac{1}{\sqrt{2}}(\mid 0\rangle + \mid 1\rangle)$ state.

**Entanglement of the motional and internal degrees of freedom of trapped ions.** The coupling of the motion of the ions inside a trap is provided by the Coulomb repulsion between ions which is much stronger than any other interaction when the typical separations between the ions is of a few optical wavelengths. For example, the average separation between calcium ions in a linear Paul trap is about $10\mu$m. The mutual Coulomb repulsion creates a spectrum of coupled normal modes of vibration for the trapped ions. The frequencies of different normal modes are well separated in the excitation spectrum.

The center-of-mass mode in which the ions oscillate in lock step in the harmonic potential well of the trap is the vibrational mode of lowest frequency ( $\omega_x$) along the $x$-axis. Initially, the ions are laser cooled at a temperature much lower than that corresponding to $\omega_x$ to ensure that each vibrational mode occupies its quantum mechanical ground state. When an ion absorbs or emits a laser photon, the center of mass of the ion recoils.

The laser pulses incident on the ion can be tuned to simultaneously cause internal spin transitions and vibrational (phonon) excitations; local (atomic) internal states are thus mapped into shared (collective) phonon states. For certain choices of the laser frequency $\omega_L$, the Hamiltonian describing the interaction between an ion and the electric field of the laser beam is resonant and the spin and the motion can be coupled efficiently (*sideband coupling*[47]) [460]. For example, for laser frequencies $\omega_L = \omega_0 \mp \omega_x$, the interaction Hamiltonians take resonant forms and the spin transitions $|\downarrow\rangle \rightarrow |\uparrow\rangle$ are accompanied by the corresponding motional mode transitions $\mid n\rangle \rightarrow \mid n \mp 1\rangle$. If the laser frequency takes the value $\omega_L = \omega_0$, the Hamiltonian is such that the spin transitions $|\downarrow\rangle \rightarrow |\uparrow\rangle$ do not change $n$, i.e., they are no longer accompanied by motional mode transitions.

In certain conditions, the so-called *Lamb-Dicke limit*, and for sufficiently low intensities, the laser beam will cause transitions that modify the state of only one of the collective modes. For example, if the laser is detuned from the resonance frequency $\omega_0$, by the quantity $\Delta\omega_L^{(j)}$, which equals minus the center-of-mass frequency, $\omega_x$, that is $\Delta\omega_L^{(j)} = -\omega_x$ (lower motional sideband), the laser excites the ion internal transition *plus* the center-of-mass motion mode as the only collective vibrational mode.

If the laser addresses the $j$-th ion in a string of $N$ ions with a frequency tuned to the lower motional sideband of the center-of-mass mode and the equilibrium position of the ion in a linear trap coincides with the node of the laser standing wave, an interaction takes place between the internal states of ion $j$ and the center-of-mass mode; the corresponding

[47]Sideband coupling means that the frequency of the laser, $\omega_L$, must be equal to the corresponding internal transition frequency, $\omega_0$, minus/plus the frequency of the trap, $\omega_x$; in this way it excites only a center-of-mass phonon. If $\omega_L = \omega_0 - \omega_x$ (*red-sideband coupling*) the spin transitions $|\downarrow\rangle \rightarrow |\uparrow\rangle$ are accompanied by motional mode transitions $\mid n\rangle \rightarrow \mid n-1\rangle$. If $\omega_L = \omega_0 + \omega_x$ (*blue-sideband coupling*) the spin transitions $|\downarrow\rangle \rightarrow |\uparrow\rangle$ are accompanied by motional mode transitions $\mid n\rangle \rightarrow \mid n+1\rangle$. When $\omega_L = \omega_0$ the spin transitions $|\downarrow\rangle \rightarrow |\uparrow\rangle$ do not change $n$.

Hamiltonian for this single ion [97] is

$$\mathbf{H}_{j,q} = \frac{\eta}{\sqrt{N}} \frac{\Omega}{2} \left( \mid e_q \rangle_j \; {}_j\langle g \mid \; a e^{-i\phi} + \mid g \rangle_j \; {}_j\langle e_q \mid \; a^\dagger e^{i\phi} \right).$$

The terms $a^\dagger$ and $a$ are the creation and annihilation operators of center-of-mass phonons, respectively, $\Omega$ is the Rabi frequency, $\phi$ is the laser phase; $\eta$ is the Lamb-Dicke parameter. The subscript $q = 0, 1$ refers to the virtual electron level excited by the laser; the transition depends on the laser polarization. These are the intermediate (virtual, electronically excited) states of the Raman process[48]. The factor $\sqrt{N}$ appears because for $N$ ions the effective mass of the center-of-mass motion is $NM$ ($M$ is the mass of the ion) and the amplitude of the mode is proportional to $1/\sqrt{NM}$. It is assumed that $(\Omega/2\omega_x)^2\eta^2 \ll 1$ and in the Lamb-Dicke limit[49] $\eta \ll 1$.

If this laser shines on the ion for a time $t = k\pi/(\Omega\eta/\sqrt{N})$ (a $k\pi$ laser pulse), the evolution of the system represents a mapping of the spin qubit to the motion qubit and is described by the unitary operator

$$\mathbf{U}_j^{k,q} = e^{-i\mathbf{H}_{j,q}t} = \exp\left[ -ik\frac{\pi}{2} (\mid e_q \rangle_j \; {}_j\langle g \mid a \; e^{-i\phi} + h.c.) \right]$$

The operator describing the evolution of the system is unitary if the laser pulse duration is proportional to $k\pi$, with $k$ the laser wave vector. This transformation does not change the system state $\mid g \rangle_j \mid 0 \rangle_c$, but transforms coherently the states $\mid g \rangle_j \mid 1 \rangle_c$ and $\mid e_q \rangle_j \mid 0 \rangle_c$ to

$$\mid g \rangle_j \mid 1 \rangle_c \rightarrow \cos(\frac{k\pi}{2}) \mid g \rangle_j \mid 1 \rangle_c - ie^{i\phi} \sin(\frac{k\pi}{2}) \mid e_q \rangle_j \mid 0 \rangle_c,$$

$$\mid e \rangle_j \mid 0 \rangle_c \rightarrow \cos(\frac{k\pi}{2}) \mid e_q \rangle_j \mid 0 \rangle_c - ie^{i\phi} \sin(\frac{k\pi}{2}) \mid g \rangle_j \mid 1 \rangle_c,$$

where $\mid 0 \rangle_c$ and $\mid 1 \rangle_c$ represent the center-of-mass states with zero phonon (vacuum), respectively, one phonon. According to the last transition, the laser pulse removes a bit of information stored as an internal excited state of the $j$th ion and transfers it to the collective state of motion of all the ions. In this way, the state of motion of the $i$th ion (where $i \neq j$) is going to be influenced by the change of internal state of the $j$th ion. The entanglement between the spin of internal atomic states and motion is evident since the final state cannot be factored out into a product of spin and motional wave functions.

**Two-qubit gate.** This gate, by its very nature, requires the entanglement of two qubits and that is difficult to realize with laser cooled trapped ions where the interaction between two two-level systems must be mediated by an auxiliary system (phonons or photons) which becomes a new source of decoherence.

[48]Raman process or Raman scattering refers to the *inelastic* scattering of photons (laser light) interacting with matter, ions in our case. A photon is absorbed and subsequently emitted via an intermediate electron state corresponding to a virtual energy level; an energy exchange occurs between the incident photon and the ion and, as a result, the energy of the emitted photon is equal to the difference between some vibrational/rotational electron energy levels of the ion.

[49]In the Lamb-Dicke limit the vibrational amplitude of the ion is considered to be much smaller that the wavelength of the light. In our case, this condition ensures that a single ion is localized in a region much smaller than the wavelength of the laser light acting upon it.

The first model of a universal two-qubit gate proposed by Cirac and Zoller in 1995 [97] assumes two ions in a linear Paul trap[50] interacting with two lasers and their state transformation is defined by

$$\mid \psi_1\rangle_i \mid \psi_2\rangle_j \rightarrow (-1)^{\psi_1\psi_2} \mid \psi_1\rangle_i \mid \psi_2\rangle_j, \qquad (\mid \psi_{1,2}\rangle = \mid 0\rangle, \mid 1\rangle)$$

where $\mid \psi_1\rangle$ and $\mid \psi_2\rangle$ are two internal states (ground state and an excited state), respectively, for each of the ions $i$ and $j$ involved; they represent the qubit basis states. The implementation of this two-qubit model phase gate would require three steps:

(1) A $\pi$ laser pulse is applied to the lower motional (red) sideband of the transition $\mid g\rangle_i \rightarrow \mid e_0\rangle_i$ of the $i$-th ion; the laser pulse has phase $\phi = 0$ and polarization $q = 0$. The evolution of ion $i$ is described by the unitary operator $\mathbf{U}_i^{1,0} \equiv \mathbf{U}_i^{1,0}(0)$.

(2) The laser directed on the $j$-th ion is next turned on for a time corresponding to a $2\pi$ pulse; this laser pulse has phase $\phi = 0$, polarization $q = 1$ and is applied to the lower motional (red) sideband of the transition from ground state to an intermediate excited state, $\mid g\rangle_j \rightarrow \mid e_1\rangle_j$, of the second ion, $j$. The lasers are applied to the lower motional sideband ($\omega_L = \omega_0 = \omega_x$) in order to excite a center-of-mass phonon only. The unitary operator corresponding to this action, $\mathbf{U}_j^{2,1}$, changes the sign of the state $\mid g\rangle_j \mid 1\rangle_c$ which is rotated through the auxiliary state $\mid e_1\rangle_j \mid 0\rangle_c$ (the subscript $c$ stands for center-of-mass mode); no other state if affected in the process.

(3) The first step is repeated and a $\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 0$ is beamed on the $i$-th ion; the unitary operator is, as above, $\mathbf{U}_i^{1,0}$.

The entire process can be described by the unitary operation

$$\mathbf{U}_{i,j} = \mathbf{U}_i^{1,0}\ \mathbf{U}_j^{2,1}\ \mathbf{U}_i^{1,0}$$

where the superscripts of each term reflect the laser pulse length (wave vector number $k$) and polarization. This unitary operation has the following effect on the possible states of the system:

$$\begin{array}{rrrr}
\mathbf{U}_i^{1,0} & \mathbf{U}_{i,j}^{2,1} & \mathbf{U}_j^{1,0} & \\
\mid g\rangle_i \mid g\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid g\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid g\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid g\rangle_j \mid 0\rangle_c, \\
\mid g\rangle_i \mid e_0\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid e_0\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid e_0\rangle_j \mid 0\rangle_c \rightarrow & \mid g\rangle_i \mid e_0\rangle_j \mid 0\rangle_c, \\
\mid e_0\rangle_i \mid g\rangle_j \mid 0\rangle_c \rightarrow & -i \mid g\rangle_i \mid g\rangle_j \mid 1\rangle_c \rightarrow & i \mid g\rangle_i \mid g\rangle_j \mid 1\rangle_c \rightarrow & \mid e_0\rangle_i \mid g\rangle_j \mid 0\rangle_c, \\
\mid e_0\rangle_i \mid e_0\rangle_j \mid 0\rangle_c \rightarrow & -i \mid g\rangle_i \mid e_0\rangle_j \mid 1\rangle_c \rightarrow & -i \mid g\rangle_i \mid e_0\rangle_j \mid 1\rangle_c \rightarrow & - \mid e_0\rangle_i \mid e_0\rangle_j \mid 0\rangle_c.
\end{array}$$

We notice that the sign of the state of the system of two ions coupled through the center-of-mass mode, $c$, changes (it acquires a phase) only when both ions $i$ and $j$ are initially excited. The state of the center-of-mass mode is restored to $\mid 0\rangle_c$ after the process.

[50]The Paul trap now being used in many applications for storing ions is a three-dimensional version of the two-dimensional mass filter patented by Wolfgang Paul (Nobel Prize in physics 1989) and his student Steiwedel. An electric quadrupole made up of two pairs of electrodes A-A and B-B (the symmetry axes of the pairs are perpendicular to each other) works as a two-dimensional mass filter if a DC voltage in addition of an RF oscillating AC voltage is applied to the electrode pair A-A. The same voltages but of opposite signs are applied to the other electrode pair B-B. In a linear trap the electrodes are parallel and create a saddle-shaped electric field; the ions are trapped by this electric field and kept levitated in the potential well created along the long axis of the electrodes. The electrode alignment and the machining requirements are critical constraints for a three-dimensional trap.

A CNOT gate can be realized using this phase coupled with an appropriate individual one-qubit rotation. This individual rotation acts on a single ion without modifying the center-of-mass motion involving both ions. The rotation can be performed using a laser with frequency on resonance with the internal transition between $\mid g\rangle_j$ and $\mid e_0\rangle_j$ of the second ion (laser detuning $\Delta\omega_L^{(j)} = 0$) and with polarization $q = 0$ while the equilibrium position of the ion coincides with the antinode of the laser standing wave. The Hamiltonian corresponding to the laser interaction with the second ion, $j$, is

$$\mathbf{H}_j = \frac{\Omega}{2}\left(e^{-i\phi} \mid e_0\rangle_j \; {}_j\langle g \mid + e^{i\phi} \mid g\rangle_j \; {}_j\langle e_0 \mid \;\right).$$

If we use a $k\pi$ laser pulse, the interaction time is $t = k\pi/\Omega$ and the interaction process can be described by the unitary evolution operator

$$\mathbf{V}_j^k(\phi) = \exp\left[-ik\frac{\pi}{2}(e^{-i\phi} \mid e_0\rangle_j \; {}_j\langle g \mid + h.c.)\right].$$

As a result of the interaction with the laser pulse the states of the second ion, $j$, evolve into superpositions

$$\mid g\rangle_j \rightarrow \cos(\frac{k\pi}{2}) \mid g\rangle_j - ie^{i\phi}\sin(\frac{k\pi}{2}) \mid e_0\rangle_j,$$

$$\mid e_0\rangle_j \rightarrow \cos(\frac{k\pi}{2}) \mid e_0\rangle_j - ie^{-i\phi}\sin(\frac{k\pi}{2}) \mid g\rangle_j.$$

The complete CNOT gate for the states $\mid \psi\rangle_i \mid \psi\rangle_j$ where $\psi = g, e_0$, is

$$G_{\texttt{CNOT}} = \mathbf{V}_j^{1/2}(\frac{\pi}{2})\; \mathbf{U}_{i,j}\; \mathbf{V}_j^{1/2}(-\frac{\pi}{2}).$$

The gate has an intermediate step whereby the atomic amplitudes corresponding to the qubits are "hidden" in a third internal atomic level $\mid e_1\rangle$ reached through a unitary transformation $\mathbf{U}_j^{k=1,q=1}$. The rotations induced by the $2\pi$ laser pulse (unitary transformation $\mathbf{U}_j^{k=2,q=1}$) proceed through this state and selectively change the sign of the atomic amplitudes. In this scheme, the two-qubit gate was realized using the collective center-of-mass mode as an auxiliary quantum degree of freedom.

**Three-qubit gate.** The gate can be implemented in the same way as the two-qubit gate, but through a five-step process involving ions $i, j$, and $l$:

1. A $\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 0$ excites the $i$-th ion (same as step (1) before).

2. A $\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 1$ is directed on $j$-th ion (same as step (2) before, but with a $\pi$ pulse).

3. A $2\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 1$ is directed on the $l$-th ion (same as step (2) before, but with ion $l$.

4. A $\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 1$ is directed on $j$-th ion (same as step (2) above).

5. A $\pi$ laser pulse of phase $\phi = 0$ and polarization $q = 0$ excites the $i$-th ion (same as step (1) above).

In this case, the sign of the state is changed only if all three ions were initially excited. This procedure can be extended to the case of many ions.

The implementation of quantum gates following the procedure proposed by Cirac and Zoller [97] is based on two key ideas:

(i) the nonlocal entanglement between individual qubits (ions) is achieved by transferring the internal atomic coherence (Rabi oscillation between ground and excited state)to and from the c center-of-mass motion involving all ions and

(ii) the gate has an intermediate step whereby the amplitudes of the atomic states corresponding to the qubits are "hidden" in a third internal atomic level $\mid e_1\rangle$; the $2\pi$ laser pulse induces $2\pi$ rotations via this state and selectively changes the sign of atomic state amplitudes. After the complete gate operation, no population would be left in the auxiliary atomic and center-of-mass levels.

Their procedure for implementing a two-qubit gate uses only product states as inputs and single-qubit measurements on the outputs [340]. In this way they avoid the preparation of Bell states to be used as inputs and performing Bell measurements on the outputs.

However, numerical calculations [340] have shown that only for Rabi frequencies much smaller than the trap frequency ($\Omega \ll \omega_x$) the evolution of the system given by the Hamiltonians of [97] corresponds to their model two-qubit phase gate. For finite Rabi frequencies the gate operation is not ideal; population is left in these state and we have an error in the gate operation. In fact, an error detection procedure could be implemented by probing the population of these intermediate states with a laser inducing fluorescence of such state after each gate operation. The full characterization of the two-qubit gate implies the preparation of various initial states and the *quantum tomography* of the corresponding output states. The quantum tomography is the process which allows the complete characterization of an unknown quantum state that can be repeatedly prepared; it consists of finding an appropriate sequence of measurements which allows us to determine the complete density matrix $\rho$ of the state.

Several research groups have shown that a controlled-NOT gate could be performed using two internal levels of a single ion and a phonon state [299, 339, 400, 91].

**Advantages and problems of ion traps.** The ion traps have several significant advantages: (i) many of the techniques required to prepare and manipulate quantum states have already been developed for precision spectroscopy work; (ii) the decoherence rates associated with the decay of excited ionic state and the damping of the ionic motion can be made small in the virtually perturbation-free environment of an ion trap (ultra-high vacuum, low electromagnetic noise level) and by using metastable atomic states; (iii) there is an experimentally demonstrated technique, quantum jumps, for reading out the result of a computation with high probability.

Major problems in ion traps such as, spontaneous transitions in the vibrational motion due to heating, internal RF transitions in the ions driven by thermal radiation, and experimental instabilities in the laser beam power, in trap related RF voltages and mechanical vibrations, and fluctuating external magnetic fields were recognized early on [298, 407]. The accuracy of

a quantum gate is decreased by the instabilities contributing to heating the ions and by the laser pulse not being exactly the right frequency and duration.

One basic feature of the Cirac and Zoller two-qubit gate implementation scheme - cooling to the ground state of motion of the trap - has proved to be very difficult to achieve. Initially, this basic step has been reliably realized in experiments for one ion [298] and two ions [427] which also were the first experimental demonstrations of a quantum logic gate.

**Experimental realization of cold ion traps.** Trapped ions have made significant experimental progress and demonstrated the essential characteristics as quantum registers of a quantum computer:

(i) a trapped ion can be initialized to a known quantum state. Early experiments (1989) of Diedrich *et al.* [123] showed that a single ion ($^{198}Hg^+$) trapped in a Paul RF trap and laser cooled in the resolved sideband regime had its kinetic energy reduced to a value where the ion was spending about 95% of its time in the ground state level of its confining potential well. A 99.9% probability of cooling to the ground state of vibration was measured for a single $Ca^+$ ion in a Paul trap by Roos *et al.* [361] and coherent quantum state manipulation was demonstrated for more than 30 periods of Rabi oscillations. A register of two ions was prepared in the ground state of their collective modes of motion almost nine years later by King *et al.* [238]; two $^9Be^+$ trapped in an elliptical Paul RF trap were cooled and probed with laser beams. The two-qubit gate of Monroe *et al.* [298] was implemented using a single trapped ion and a motional mode and two of the ion's internal states as the qubits.

(ii) final states can be efficiently detected. The fluorescence of the ions was monitored to detect the internal state $|\uparrow\rangle$ or $|\downarrow\rangle$ with discrimination efficiency of almost 100% [361].

(iii) the decoherence times of qubits are relatively long. Coherence times longer than 1 ms were observed for Rabi oscillations between internal states $S_{1/2} \leftrightarrow D_{5/2}$ of a single trapped ion of $^{40}Ca^+$ prepared by sideband cooling in the lowest motional (vacuum) state by Schmidt-Kaler *et al.* [365]. One logical qubit encoded into the decoherence-free subspace of a pair of trapped $^9Be^+$ ions is protected from collective dephasing, a major source of qubit decoherence for trapped ions; the storage time of the logical qubit increases by up to an order of magnitude [235] in comparison with a qubit encoded in a single ion.

(iv) quantum logic gates of a universal set have been experimentally demonstrated. Single-qubit gates, rotations of a single physical qubit, are easy to implement [459] and much simpler than in other systems, two-qubit gates have been realized based on slightly different concepts: entanglement of two qubits implemented as a single physical ion with conditional dynamics depending either on the wavelength of the applied laser radiation [299], or on the size of the atomic wave packet [113]; entanglement of two qubits implemented as two ions (the logical states of each qubit are two hyperfine ground states of an ion) with motion strongly coupled through their Coulomb repulsion and described in terms of the center-of-mass and *stretch*[51] normal modes [270] (where the qubit logical states are two hyperfine ground states of $^9Be^+$ ion) and [366] (where the logical states of the qubits are the ground and a metastable state and the phonon-number of the motional mode). Two-qubit gates with trapped ions, as implemented

[51] In the *stretch* normal mode two ions move in opposite directions with equal displacements at a frequency $\omega_s = \sqrt{3}\omega_c$.

so far, have their limitations. Ideally, they should be independent of temperature and need not be cooled to their ground state for accurate operation, should require no addressability and should allow ions be packed close together for stronger interaction, should be fast to minimize the effects of decoherence during gate operation and to speed up computation (gate time is limited by low laser intensity required for spectroscopically resolution of motional sidebands of the ions.

(v) states of two and four qubits have been entangled [427, 364, 293, 401].

As first proposed, quantum computing with trapped ions involved a string of ions confined in a single trap, internal electron states of ions were used as qubit basis states and quantum information was transferred between ions through their mutual Coulomb interaction. The ion trap system satisfied all the requirements mentioned above for a quantum register. The manipulation of a large number of ions in a single trap presents such serious technical difficulties that the scaling up of this scheme is limited to only tens of ions. A quantum computer based on the original Cirac-Zoller model which places all the ions in a single trap could not manipulate the $10^6$ ions needed for a full scale computation. A trap containing $10^6$ ions would be limited [121] to a center-of-mass phonon frequency $\omega_c \ll 100$ Hz, where $\omega_c \propto 1/\sqrt{N}$. At such frequency, the $10^{11}$ operations required for the Shor algorithm without error correction would be executed in over $10^9$ second or about 30 years. In the trap array model of DeVoe [121] with some $64,000$ traps, the $N = 64$ ions per trap would have a phonon frequency $\omega_c = 125$ kHz; the computation could proceed $10^3$ times faster than in the single-trap model. A logical solution to escape this limitation is to use arrays of micro-traps, eventually organized as a number of small ion-trap quantum registers, and quantum communication between them. Some proposals use photon coupling between traps [323, 121], others spin-dependent Coulomb interactions [99].

Basic experimental steps have been made towards the development of scalable ion trap quantum computing devices with a larger number of ions which can be stored and moved to different regions to perform the gates required by quantum algorithms. A recent proposal of Kielpinski *et al.* [236] of a "quantum charge-coupled device" architecture consists of a large number of interconnected ion traps and is based on techniques already demonstrated for manipulating small quantum registers. A few ions can be confined in each trap and manipulated using techniques already demonstrated experimentally or can be moved between traps by changing the voltages of the traps. The device has memory and interaction regions. Ions storing quantum information are confined to traps assigned to the memory regions. Logical operations are executed on the relevant ions which have been moved into the interaction region by applying appropriate transport voltages to the electrode segments in between. Once inside the interaction region, the ions are kept close together to allow the right amount of Coulomb coupling required by entangling gates. Gates are driven by laser beams focused on the interaction region. The trapping and transport potentials are realized through a combination of RF and quasi-static electric fields. The device can be constructed using standard microfabrication techniques. As a proof of concept, a pair of interconnected ion traps have been constructed at NIST. A qubit ($^{40}Ca^+$ ions) was coherently transported between the traps through a Ramsey-type experiment without any heating of the ion motion or shortening the ion lifetime in the trap. The ions used for quantum logic gates could be "sympathetically" cooled by another ion species used as a heat sink in the interaction region, as demonstrated experimentally [62]. Sources of decoherence during transport and positioning of ions could

be reduced by several orders of magnitude by encoding each qubit into a decoherence-free subspace (DFS) of two ions [235].

## 6.3 First Experimental Demonstration of a Quantum Logic Gate

The first experimental demonstration of a `CNOT` quantum gate operating on prepared quantum states of two qubits was reported in 1995 by Monroe, Meekhof, King, Itano, and Wineland [298] from the National Institute of Standards and Technology (NIST) in Boulder, Colorado. Their implementation followed the quantum logic gate scheme proposed by Cirac and Zoller [97]; the characteristic states of the two qubits are two internal and two external states, respectively, of a single $^9Be^+$ ion stored in a coaxial-resonator RF (Paul) ion trap. The *internal* states are two hyperfine ground states of the ionized beryllium atom, and the *external* states are two quantized harmonic oscillator states (usually denoted by $n$) associated with this ion's motion inside the trap.

The target qubit of the `CNOT` quantum logic gate, denoted $S^{(t)}$, is implemented as two $^2S_{1/2}$ hyperfine ground states of the $^9$ $Be^+$ ion. The quantization axis for the hyperfine states is defined by a static magnetic field $\mathbf{B} \approx 0.2$ mT applied along the direction $(-1/\sqrt{2})\overrightarrow{x} + (1/2)(\overrightarrow{y} - \overrightarrow{z})$, where $\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}$ are the trap axes. The hyperfine ground states are the $2s\ ^2S_{1/2} \mid F = 2, m_F = 2\rangle$ and $2s\ ^2S_{1/2} \mid F = 1, m_F = 1\rangle$ [52], usually represented as the spin - 1/2 states $|\downarrow\rangle$ and $|\uparrow\rangle$, respectively. The separation in frequency between the two levels (the hyperfine splitting) is $\omega_0/2\pi \simeq 1.25$ GHz.

Coherent transitions between these levels can be implemented with two laser beams that drive two-photon stimulated Raman transitions. In this case, the laser parameters are given by the differences between the parameters of the two laser beams at the mean position of the ion. Parameters such as the wave vector $k$, frequency $\omega_L$, and the phase $\phi$ become $k = k_{L1} - k_{L2}$ (a vector assumed to be parallel to the $x$-axis), $\omega_L = \omega_{L1} - \omega_{L2}$ and, respectively, $\phi = \phi_{L1} - \phi_{L2}$.

The intermediate, electronically excited Raman states are: $2p\ ^2P_{1/2}$, corresponding to state $\mid e\rangle$, and $2p\ ^2P_{3/2}$, corresponding to state $\mid e'\rangle$. These two levels are separated by the fine structure splitting, $\omega_F$, which is assumed to be much larger than the hyperfine splitting, $\omega_0$, $\omega_0 \ll \omega_F$; the hyperfine splitting of these two excited levels are considered negligible.

The control qubit, denoted $n^{(c)}$, is implemented as the first two quantized harmonic oscillator states $\mid 0\rangle$ and $\mid 1\rangle$ of the trapped ion. The separation in frequency of these two motional states is the vibrational frequency $\omega_x/2\pi = 11$ MHz of the trapped ion with a favored motion degree of freedom along the $x$ - axis. The two-qubit register has four basis eigenstates $\mid n^{(c)}\rangle \mid S^{(t)}\rangle = \mid 0\rangle\,|\downarrow\rangle, \mid 0\rangle\,|\uparrow\rangle, \mid 1\rangle\,|\downarrow\rangle, \mid 1\rangle\,|\uparrow\rangle$. These states can be individually accessed in the following way: a pair of off-resonance laser beams generated from a single laser are directed to the ion and they are driving stimulated Raman transitions [53] between

[52] $F$ represents the total angular momentum quantum number of the atom (nucleus + electrons) and $m_F$ is the corresponding magnetic quantum number. The Be nucleus has angular momentum 3/2, thus the total angular momentum of the ion is $F = 3/2 + 1/2 = 2$ and $F = 3/2 - 1/2 = 1$ for spin state $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively. The magnetic momentum takes values in the $m_F = +F, \ldots, -F$ range.

[53] In a stimulated Raman transition a two-color (two-frequency) laser pulse transfers the atom/molecule from ground state to a vibrational excitation state if the difference in energy matches an allowed Raman transition.

the states. The two laser beams, $R_1, R_2$, also called Raman beams, have right/left circular-polarization and linear-polarization, respectively. The beam $R_2$ is shined along the direction $(-1/\sqrt{2})\overrightarrow{x} + (1/2)(-\overrightarrow{y} + \overrightarrow{z})$ and the beam $R_1$ is perpendicular to it [238]. In this way the wavevector difference between the Raman beams is parallel to the $x$-axis and the transitions are sensitive to ion motion only along this direction. The Raman laser beams have about 1 mW power each at about 313 nm wavelength and are detuned about 50 MHz red off the ${}^2S_{1/2} \longrightarrow {}^2P_{1/2}$ transition of the Be ion. The difference in frequency of the two Raman beams could be tuned in the range 1.20 to 1.30 GHz and the Raman pulse durations.

When the difference frequency $\Delta\omega_L$ of the laser beams is set at:

(a) $\Delta\omega_L \simeq \omega_0$ (the *carrier*) - transitions $\mid n^{(c)}\rangle \mid\downarrow\rangle \longrightarrow \mid n^{(c)}\rangle \mid\uparrow\rangle$ are coherently driven between the internal states $\mid S^{(t)}\rangle$ while preserving the external state $\mid n^{(c)}\rangle$; the carrier Rabi frequency for Be is $\Omega_0/2\pi \simeq 140$ kHz.

(b) $\Delta\omega_L \simeq \omega_0 - \omega_x$ (the *red-sideband*) - transitions $\mid 1\rangle \mid\downarrow\rangle \longrightarrow \mid 0\rangle \mid\uparrow\rangle$ are coherently driven between $\mid n^{(c)}\rangle \mid S^{(t)}\rangle$ states; the red-sideband Rabi frequency is $\eta_x\Omega_0/2\pi \simeq 30$ kHz, where the Lamb-Dicke parameter $\eta_x \simeq 0.2$.

(c) $\Delta\omega_L \simeq \omega_0 + \omega_x$ (the *blue-sideband*) - transitions $\mid 0\rangle \mid\downarrow\rangle \longrightarrow \mid 1\rangle \mid\uparrow\rangle$ are coherently driven between $\mid n^{(c)}\rangle \mid S^{(t)}\rangle$ states; the blue-sideband Rabi frequency is $\eta_x\Omega_0/2\pi \simeq 30$ kHz. In the last two cases, when $\Delta\omega_L$ is tuned to one of the sidebands, the stimulated Raman transitions entangle the internal spin $\mid S^{(t)}\rangle$ states with the external motional $\mid n^{(c)}\rangle$ states and this feature is at the heart of the trapped ion `CNOT` quantum gate.

Before the actual operation of the `CNOT` gate, the ${}^9Be^+$ ion in the RF trap is Doppler precooled in all three dimensions with right circular-polarized laser beams $D_1$ and $D_2$ and then Raman cooled to the lowest ground state $\mid 0\rangle \mid\downarrow\rangle$; the two-qubit register is then prepared in an arbitrary state by applying to the ion the Raman laser pulses appropriately tuned and timed. For example, the eigenstate $\mid 1\rangle \mid\downarrow\rangle$ is prepared by applying to the ion in the lowest ground state, $\mid 0\rangle \mid\downarrow\rangle$, a $\pi$ pulse on the blue-sideband ($\Delta\omega_L \simeq \omega_0 + \omega_x$) followed by a $\pi$ pulse on the carrier ($\Delta\omega_L \simeq \omega_0$), so that

$$\mid 0\rangle \mid\downarrow\rangle \xrightarrow{\Delta\omega_L \simeq \omega_0 + \omega_x} \mid 1\rangle \mid\uparrow\rangle \xrightarrow{\Delta\omega_L \simeq \omega_0} \mid 1\rangle \mid\downarrow\rangle.$$

A third right circular-polarized laser beam prevents the optical pumping to the state $\mid F = 2, m_F = 1\rangle$.

The `CNOT` operation is implemented by sequentially shining three pulses of the Raman laser beams on the trapped ${}^9Be^+$ ion:

1. A $\pi/2$ pulse is applied on the carrier transition $\Delta\omega_L \simeq \omega_0$ and that causes the internal spin state $\mid S^{(t)}\rangle$ to go through a phase shift of $+1/4$ of a complete Rabi cycle, while leaving the external state $\mid n^{(c)}\rangle$ unchanged.

2. A $2\pi$ pulse is applied on the blue-sideband transition between the internal state $\mid\uparrow\rangle$ and an auxiliary atomic level $\mid e'\rangle$ which is not part of any qubit. This auxiliary level is another ground state, the ${}^2S_{1/2} \mid F = 2, m_F = 0\rangle$, Zeeman split from ${}^2S_{1/2} \mid F = 2, m_F = 2\rangle$, the $\mid\downarrow\rangle$ state, and shifted by about 2.5 MHz when a magnetic field of 0.18 mT was applied. The auxiliary Rabi frequency is $\eta_x\Omega_{e'} \simeq 12$ kHz.

This blue-sideband transition reverses the sign of any component of the quantum register in the $\mid 1\rangle \mid\uparrow\rangle$ state by inducing a complete Rabi cycle through the auxiliary state: $\mid 1\rangle \mid\uparrow\rangle \longrightarrow \mid 0\rangle \mid e'\rangle \longrightarrow -\mid 1\rangle \mid\uparrow\rangle$.

3. A $\pi/2$ pulse is applied on the carrier transition and the internal spin state $\mid S^{(t)}\rangle$ undergoes another phase shift of $-1/4$ of a complete Rabi cycle, but leaves the $\mid n^{(c)}\rangle$ states unchanged.

As a result of this sequence of pulses, the components of the quantum register in the $\mid n^{(c)}\rangle = \mid 0\rangle$ state are not affected by the blue-sideband transition of step (2) and the effects of the two $\pi/2$ pulses of steps (1) and (3) cancel each other. However, the components of the quantum register in the $\mid 1\rangle \mid\uparrow\rangle$ state change sign as a result of step (2) and the two Raman pulses of step (1) and (3) add constructively and rotate the target spin by $\pi$ radians (flip it).

**Qubit state detection.** The change in the population of the register after such a sequence of operations can be detected through independent measurements of the control and target qubits. The probability that after a certain operation sequence the target qubit $S^{(t)}$ is in the state $|\downarrow\rangle$, $\mathcal{P}\{\mid S^{(t)}\rangle = |\downarrow\rangle\}$, is estimated by measuring the ion fluorescence associated with the cycling transition from the $|\downarrow\rangle$ state, ${}^2S_{1/2} \mid F=2, m_F=2\rangle \longrightarrow {}^2P_{3/2} \mid F=3, m_F=3\rangle$. The transition is induced by resonant excitation with the right circular-polarized $D_2$ laser beam of 19.4 MHz radiative linewidth at $\lambda = 313$ nm. This radiation has a much lower probability of exciting the state $|\uparrow\rangle$ state; thus, the detected fluorescence is considered to be proportional to $\mathcal{P}\{S^{(t)} = |\downarrow\rangle\}$. The second measurement is supposed to evaluate the probability that the control qubit $\mid n^{(c)}\rangle$ is in state $\mid 1\rangle$, $\mathcal{P}\{\mid n^{(c)}\rangle = \mid 1\rangle\}$. The same operation sequence performed before the first measurement is repeated with the addition of a specific 'preparation' Raman laser pulse before a new detection of the target qubit $S^{(t)}$. The 'preparation' pulse has the role of mapping the $\mid n^{(c)}\rangle$ onto $\mid S^{(t)}\rangle$ in the following way:

(a) if the first measurement finds the target qubit $S^{(t)}$ to be in $|\downarrow\rangle$ state, the initial operation sequence is repeated, a laser $\pi$ pulse on the red-sideband is beamed at the ion before testing again the state of the target $\mid S^{(t)}\rangle$. If fluorescence is detected, then the control qubit $n^{(c)}$ is in state $\mid 0\rangle$; if fluorescence is not detected, the state $\mid n^{(c)}\rangle$ is $\mid 1\rangle$;

(b) if the first measurement finds the target qubit $S^{(t)}$ to be in $|\uparrow\rangle$ state, the initial operation sequence is repeated and a laser $\pi$ pulse on the blue-sideband is shone on the trapped ion before testing again the state of the control qubit, $\mid n^{(c)}\rangle$. In this case, the presence of fluorescence indicates the $\mid n^{(c)}\rangle$ state is $\mid 1\rangle$, while the absence of it indicates a state $\mid 0\rangle$. This procedure is applied before gate operation in order to verify the preparation of the states and to measure the register population and again after the operation of the `CNOT` gate to measure the population in the resulting register. When the control qubit $n^{(c)}$ is prepared in state $\mid 0\rangle$, the gate preserves the state of qubit $S^{(t)}$ with high probability; when the control qubit is prepared in state $\mid 1\rangle$, the `CNOT` gate flips the state of the target qubit $S^{(t)}$ with high probability.

The switching speed of this `CNOT` gate is limited at about 20 kHz by the $2\pi$ laser pulse in step (2); otherwise, the switching speed would be limited at about 50 MHz, the frequency separation between the control qubit vibrational energy levels. The decoherence rate was measured to be a few kHz and was due, among others, to instabilities in the laser beam power

and in the relative position ion - laser beams, to fluctuations of the external magnetic fields, to instabilities in the RF ion trap drive frequency and voltage amplitude.

**Spontaneous emission.** A fundamental limitation of the coherence of atomic qubits is due to spontaneous emission; the ground state hyperfine levels which form the qubits have a very long radiative lifetime, but the decoherence rate due to spontaneous emission during the Raman transitions between these states is a real problem. The decoherence rate, $\text{Rate}_{SE}$, due to spontaneous emission from $^2P$ levels is estimated as

$$\text{Rate}_{SE} = \sum_i \gamma_i p_i$$

where $p_i$ is the probability that an intermediate excited state is populated and $\gamma_i$ is the decay rate of that excited state.

The probability of spontaneous emission during the time $\tau_\pi$ that a $\pi$ laser pulse is acting on the carrier transition is given by

$$\text{Probability}_{SE} = \text{Rate}_{SE} \times \tau_\pi.$$

The probability of spontaneous emission for the $^9\text{Be}^+$ ion is evaluated [460] to be at least 0.001. For sideband transitions this probability will be increased by $1/\eta$

$$\text{Probability}_{SE}^{(\text{sideband})} = \text{Probability}_{SE}^{(\text{carrier})} \times \frac{1}{\eta}$$

with the Lamb-Dicke coefficient $\eta \approx 0.2$ for the $^9\text{Be}^+$ ion. Since the probabilities for errors during a gate operation must be on the order of $10^{-4}$ or smaller to be able to perform error corrections during computations, the conclusion is that the $^9\text{Be}^+$ ion would not a good choice for the ultimate qubit. A good choice ion should have a small ratio of spontaneous decay rate, $\gamma$, to fine structure splitting, $\omega_F$; only then the effects of spontaneous emission could be greatly reduced. The probability of spontaneous emission during a laser carrier $\pi$ pulse can be approximated as [460]

$$\text{Probability}_{SE} = 2\sqrt{2}\pi\frac{\gamma}{\omega_F}.$$

The following approximations were made:

- the ions have half-integer nuclear spin $I$ and a $\pi$ pulse acts on the carrier of the transition

  $$\mid F = I - \frac{1}{2}, m_F = 0\rangle \longrightarrow \mid F = I + \frac{1}{2}, m_F = 0\rangle,$$

  for which the Rabi rate is independent of the spin $I$;

- the value of the decay rate $\gamma$ corresponds to the relevant $^2P_{1/2}$ level;

- the hyperfine level splitting, $\omega_0$, and the decay rate from the intermediate state, $\gamma$, are much smaller than the fine level splitting, $\omega_F$, and the Raman lasers detunings, $\mid \Delta \mid$, respectively, $\omega_0, \gamma \ll \omega_F, \mid \Delta \mid$;

Table 32: The probability of spontaneous emission during a two-photon stimulated Raman transition for ions of interest for quantum computing with trapped ions. A $\pi$-pulse is applied on the carrier transition of the $^2S_{1/2}$ levels, $\mid F = I - 1/2, m_F = 0\rangle \longrightarrow \mid F = I + 1/2, m_F = 0\rangle$.

| Ion | Nuclear Spin I | Carrier Transition ($^2S_{1/2}$) | Probability of Spontaneous Emission |
|---|---|---|---|
| $^9Be^+$ | 3/2 | $\mid 1,0\rangle \rightarrow \mid 2,0\rangle$ | $8.7 \times 10^{-4}$ |
| $^{25}Mg^+$ | 5/2 | $\mid 2,0\rangle \rightarrow \mid 3,0\rangle$ | $1.4 \times 10^{-4}$ |
| $^{43}Ca^+$ | 7/2 | $\mid 3,0\rangle \rightarrow \mid 4,0\rangle$ | $3.0 \times 10^{-5}$ |
| $^{67}Zn^+$ | 5/2 | $\mid 2,0\rangle \rightarrow \mid 3,0\rangle$ | $2.6 \times 10^{-5}$ |
| $^{87}Sr^+$ | 9/2 | $\mid 4,0\rangle \rightarrow \mid 5,0\rangle$ | $8.0 \times 10^{-6}$ |
| $^{113}Cd^+$ | 1/2 | $\mid 0,0\rangle \rightarrow \mid 1,0\rangle$ | $5.3 \times 10^{-6}$ |
| $^{199}Hg^+$ | 1/2 | $\mid 0,0\rangle \rightarrow \mid 1,0\rangle$ | $1.8 \times 10^{-6}$ |

- the Raman beams have orthogonal linear polarization (experimentally a convenient choice).

The probabilities of spontaneous emission during a two-photon stimulated Raman transition for various ions of interest for quantum computing calculated in this approximation [460] are presented in Table 32. These values correspond to the carrier transition of the $^2S_{1/2}$ levels

$$\mid F = I - 1/2, m_F = 0\rangle \;\longrightarrow\; \mid F = I + 1/2, m_F = 0\rangle.$$

From Table 32 we see that the spontaneous emission during Raman transitions is practically suppressed for heavier ions such as $Cd^+$ or $^{199}Hg^+$. Nevertheless, even if heavy ions were to be used, fault tolerance could be harder to reach for other transitions than the carrier transitions. Transitions such as sideband and two-qubit gates have a higher probability of spontaneous emission in the Lamb-Dicke limit[54].

Among other factors, the probability of spontaneous emission is proportional to the time, $\tau_\pi$, a $\pi$ pulse takes to be executed (the probability is higher for longer $\tau_\pi$). For sideband transitions and two-qubit gates the spontaneous emission rate remains the same (as for the carrier transition), but $\tau_\pi$ scales as the inverse of the Rabi frequency which, in turn, scales as the carrier frequency, $\omega_0$, times Lamb-Dicke parameter, $\eta$, that is, $\Omega \sim \omega_0 \times \eta$. Thus,

$$\tau_\pi = \frac{\pi}{2\Omega} \sim \frac{\pi}{2(\omega_0 \times \eta)}.$$

Therefore, for fault-tolerant computation the values of $\eta$ should not be too small compared to unity; in that case the Lamb-Dicke limit would no longer be a good approximation. Since the Rabi frequencies depend on the motional states, the precise control of the motion, for example, ground state cooling, is absolutely necessary. Error resistant techniques proposed

[54]Quantum logic gates with trapped ions generally require cooling to the Lamb-Dicke limit, $\eta \ll 1$, where the spatial extent of the ion's motion is much smaller that the optical coupling wavelength, $\lambda/2\pi = 1/k$.

for fault-tolerant computation, such as adiabatic passage, spin echo or composite pulses, like those used in NMR implementations, increase the probability of spontaneous emission and are not recommended.

Decoherence associated with unwanted motion in a scaled-up trapped ion quantum computer can be reduced by sympathetic cooling of multiple ion species ($^{9}Be^{+}$ and cold $^{26}Mg^{+}$) or multiple isotopes of the same ion species ($^{112}Cd^{+}$ and cold $^{114}Cd^{+}$) mediated by their Coulomb interactions; the coherence of the internal qubit is preserved during this type of cooling. As shown by Blinov *et al.* [62] the $Cd^{+}$ system (the pair $^{112}Cd^{+}$ as qubit and $^{114}Cd^{+}$ as refrigerator ion, or the pair $^{111}Cd^{+}$ as qubit and $^{116}Cd^{+}$ as refrigerator ion) is a convenient choice because the sympathetic cooling can be accomplished without extra lasers and without strong focusing.

## 6.4 Trapped Ions in Thermal Motion

The original Cirac-Zoller model for an ion trap implementation of a quantum logic two-qubit gate [97] requires the cooling and maintenance of the trapped ions in their oscillatory center-of-mass quantum ground state. The fidelity of the quantum logic gates realized with trapped ions depends on the quantum state of the collective oscillatory degrees of freedom of these ions. The Coulomb force strongly couples the motion of the ions and, at the same time, makes them susceptible to interaction with external electric fields.

The ions in the center-of-mass mode get "heated" in the process of interacting with external electric fields and the mode becomes more and more excited, the wave function of the ions becomes more spatially spread causing a random phase shift of the ions and, ultimately, decoherence. The solution to this problem is the understanding of and preempting the causes of the trapped ions heating, on one hand, and the use of alternative designs of quantum logic gates which operate without requiring the system to be in its quantum ground state of the phonon modes on the other hand.

The simple method of avoiding the heating problem when quantum computing with trapped ions proposed by B. E. King *et al.* [238] utilizes the "higher" modes ($m > 1$), such as $(x)_{stretch}$, $(xy)_{rocking}$, and $(xz)_{rocking}$, of the ions' collective oscillations inside the trap instead of the center-of-mass mode ($m = 1$) first proposed by Cirac and Zoller [97]. The heating rate of these higher motional modes is much smaller than that of the center-of-mass mode [238]. The *heating times* [55] measured experimentally for higher modes of a two-ion system were longer than 5 msec, while the heating times for center-of-mass modes were less than 0.1 msec. The higher modes seem to be well insulated from the influence of external heating fields and can be used as a reliable "quantum information bus".

The heating of the center-of-mass mode which is still present and causes a random phase shift of the ions by spatially spreading the wave function of the ions can be reduced and kept constant by cooling with a separate species of ion. The sequence of pulses required for implementing the `CNOT` gate in this case is the same as that recommended by Cirac and Zoller [97], but the laser frequencies employed (the sideband corresponding to the stretch mode in question) are different.

[55] The heating time is defined as the time it takes the occupation number of the mode to increase by one.

This scheme has its drawbacks: (i) the laser-ion coupling to the higher modes varies for different ions and, therefore, different pulse durations are required for different ions; (ii) the energy levels of the higher modes are closer together in frequency and that makes them harder to resolve. It is possible that such laser control problems could make the quantum computer algorithm so complex that any speedup achieved through quantum parallelism were nullified.

The proposal of Poyatos, Cirac, and Zoller [339] to realize a two-qubit logic gate between two ions in a linear trap at finite temperature is based on a concept different from that of the cold ion trap of Cirac and Zoller [97]. This "hot" ion trap implements the `CNOT` gate by splitting the atomic wave packets in the same way as in atomic interferometry and creates coherent states of the ions' collective oscillations. Under ideal conditions, the operation of the logic gate is independent of the state of motion of the ions without the restriction of the ground center-of-mass mode temperature.

The `CNOT` gate proposed by Poyatos *et al.* is realized in three steps [341]:

1. *State dependent displacement of the control ion* (ion **1**): a short laser pulse, appropriately tuned, with its wave vector $k$ pointing along the main axis, $x$, of the trap is applied to ion **1** for a time $t_L = \pi/\Omega \ll t_g$, which corresponds to the strong excitation regime $\Omega \gg \omega_x$. Here $\Omega$ is the Rabi frequency, and $t_g$ is the period of the motion in the trapping potential when the frequencies for the reduced mass and center-of-mass are commensurable, $t_g \sim \omega_x^{-1}$. The laser intensity is chosen in such a way that the internal state of the ion is flipped $\mid 0\rangle \rightarrow \mid 1\rangle$, or $\mid 1\rangle \rightarrow \mid 0\rangle$. Simultaneously, the laser pulse provides a momentum "kick" to the wave packet of the ion which will move in either one of two opposite directions depending on its internal state. Thus, when the ion **1** is in state $\mid 0\rangle$, it will be changed to state $\mid 1\rangle$ and, at the same time, it will move in one direction. When the ion **1** is in state $\mid 1\rangle$, it will be changed to state $\mid 0\rangle$ and it will move in the opposite direction. The choice of direction is associated with the absorption or emission of one phonon by the ion **1**. The wave packet of the control ion (ion **1**) will be split into two wave packets moving into opposite directions (right and left). Because of the strong ion-ion Coulomb interaction, the target ion (ion **2**) will also evolve into two spatially separated wave packets dependent on the internal state of the control ion (ion **1**). If the momentum kick provided by the initial laser pulse is sufficiently strong then, after a short time, the wave packet associated with state $\mid 0\rangle$ will be spatially distinct from that associated with state $\mid 1\rangle$.

2. *Conditional flip* $\mid 0\rangle \rightleftarrows \mid 1\rangle$ *of ion* **2**: after a time $t_0 = 2\pi/(3\omega_x)$ corresponding to the maximum distance between the wave packets, a laser beam is directed to the target ion (ion **2**) and is focused on the location of precisely that wave packet that arises if the control ion **1** were in state $\mid 0\rangle$; the laser interacts with that wave packet only. The laser beam does not induce any displacement of the ion; the set up is such that either one laser beam is used and it propagates perpendicular to the $x$-axis, the confinement axis of the trap, or two laser beams are used and they propagate in opposite directions along the $x$-axis, depending on the type of internal transition intended to be induced. But, for a given laser-ion interaction time $t_1 \ll t_g$, an internal transition is induced on the target ion (ion **2**) dependent on the state of the control ion (ion **1**), such that $\mid 0\rangle_{\mathbf{2}} \rightleftarrows \mid 1\rangle_{\mathbf{2}}$. After some time, the two ions go back to oscillate as free harmonic oscillators.

3. *State dependent displacement of control ion* (ion **1**): after a time $(t_g - t_0)$ a short laser pulse is applied to the control ion (ion **1**) with a wave vector pointing in the direction of the $x$-axis as in the first step. The pulse has the same duration $t_L = \pi/\Omega \ll t_g$ and the state after the interaction coincides with the fundamental two-qubit gate under ideal conditions. This pulse reverses the effect of the first laser pulse (the operators acting on the motional state cancel out) and the control ion is brought back to its original state.

This gate implementation is easy to extend to a system of three ions, but further scaling up is difficult. The fidelity of the gate is limited, among others, by errors related to the finite size of the wave packets and the small distances between them; the focusing of the laser beam in the second step is extremely difficult. Random electric fields can induce a random relative phase shift for the two packets, as well as a heating of the ions and that leads to an increase in the size of the ions' wave packets. Proper operation of such a 'hot' ion gate could be maintained only with adequate cooling.

## 6.5 Entanglement of Qubits in Ion Traps

**Mølmer-Sørensen entanglement method.** K. Mølmer and A. Sørensen [297] proposed the application of a simple interaction Hamiltonian to two or more ions (spin qubits) in an ion trap by simultaneously illuminating the ions with laser beams of two different colors (frequencies). The action of laser fields with two different frequencies, such that $\omega_L^1 + \omega_L^2 = 2\omega_0$ makes the two-photon process resonant, but neither of the frequencies are resonant with single excitations of the ions. Thus, it is possible to apply classical laser radiation which cannot be absorbed by a single ion, but which induces two ions to undergo the $\mid g\ g\rangle \longleftrightarrow \mid e\ e\rangle$ transition simultaneously. If the laser frequencies are $\omega_L^1 = \omega_0 - \Delta$ and $\omega_L^2 = \omega_0 + \Delta$ with the detuning $\Delta$ close to, but not resonant with, the center-of-mass vibration frequency, there is a high probability that only the intermediate states with one excited ion and a vibrational quantum number raised or lowered by unity are involved in the interaction.

They showed that an interaction Hamiltonian composed of the center-of-mass vibrational energy of the ions string and the internal electronic energy of the ions can be used to create entangled states of multiple particle of the form

$$\mid \psi\rangle = \frac{1}{\sqrt{2}}(e^{i\phi_g} \mid g\ g \ldots g\rangle + e^{i\phi_e} \mid e\ e \ldots e\rangle)$$

where, the product states $\mid g\ g \ldots g\rangle$ and $\mid e\ e \ldots e\rangle$, with $N$ terms each, describe $N$ ions which are all in the same internal state $\mid g\rangle$ or $\mid e\rangle$.

Their method does not require the collective vibrational motion of the ions to be in its ground state; their scheme is based on two quantum mechanical "special effects": (i) the vibrational degrees of freedom used as intermediate states in the communication processes between ions have only a "virtual" role (they are not populated) and (ii) the transition paths involving such unpopulated vibrational states interfere destructively eliminating the dependence of decay rates and evolution frequencies on vibrational quantum number.

Sørensen and Mølmer proposed another scheme for a two-qubit (two ion) gate where each ion in the ion trap is addressed with a single laser (like in the original scheme of Cirac

and Zoller), but quantum logic operation of the gate is performed through off-resonant laser pulses as their entanglement method prescribes [400]. The two ions are illuminated with lasers detuned close to the upper and lower sidebands. The bichromatic light selects certain virtually excited intermediate states and, by choosing appropriate parameters, the desired internal state dynamics of the ions may ultimately be achieved, even if the vibrational collective motion used to couple the ions is not in its ground state.

The qubits are characterized by the states $\mid g\ n\rangle$ and $\mid e\ n\rangle$ in each ion, where $\mid g\rangle$ and $\mid e\rangle$ are the internal ion states, ground and excited, respectively, and $\mid n\rangle$ is the collective motional state of the ion string with $n$ the quantum number for the main vibrational mode of the trap; the vibrational mode can be other than the vibrational ground state and the trapped ions are called "hot". The intermediate vibrational states under consideration are $\mid g\ (n-1)\rangle, \mid g\ (n+1)\rangle$ and $\mid e\ (n-1)\rangle, \mid e\ (n+1)\rangle$. The laser addressing the first ion has a detuning close to the upper sideband, $\omega_L^{(1)} = \omega_0 + \delta$, which is close to resonance with the joint vibrational and internal excitation of the ion. The laser addressing the second ion has a detuning which is the negative of the detuning for the first ion, $\omega_L^{(2)} = \omega_0 - \delta$, and also close to resonance with the joint vibrational and internal excitation state of the ion. Here $\omega_0$ is the exact resonance frequency for the levels $\mid g\ n\rangle$ and $\mid e\ n\rangle$ and is the same for the two ions; $\mid \delta \mid$ is the detuning interval. The two laser beams shining simultaneously on the two ions induce the coupling of the states $\mid g^{(1)}\ g^{(2)}\ n\rangle$ and $\mid e^{(1)}\ e^{(2)}\ n\rangle$ using two equally probable pathways through the intermediate states $\mid g^{(1)}\ e^{(2)}\ (n-1)\rangle$ and $\mid e^{(1)}\ g^{(2)}\ (n+1)\rangle$, that is

$$\mid g^{(1)}\ g^{(2)}\ n\rangle \longleftrightarrow \{\mid e^{(1)}\ g^{(2)}\ (n+1)\rangle,\ \ \mid g^{(1)}\ e^{(2)}\ (n-1)\rangle\} \longleftrightarrow \mid e^{(1)}\ e^{(2)}\ n\rangle.$$

The detunings are chosen sufficiently close to the sideband that only one collective degree of vibrational excitation (the main vibrational mode of the trap) is taken into consideration in such a way that the intermediate states $\mid e^{(1)}\ g^{(2)}\ (n+1)\rangle$ and $\mid g^{(1)}\ e^{(2)}\ (n-1)\rangle$ are not populated in the process, they are "virtually" excited. Because of the opposite detunings, the factors containing $n$ cancel out and the coherent evolution of the internal ionic state becomes insensitive to the vibrational quantum numbers; the coherent oscillation from $\mid g\ g\rangle$ to $\mid e\ e\rangle$ can be observed even when the ions are in any superposition or mixture of vibrational states and even when the vibrational motion exchanges energy with a thermal reservoir.

The heating of the vibrational motion is an important source of decoherence for the cold trapped ion systems. The Sørensen - Mølmer scheme can be made insensitive to the interaction with the environment. Their ion trap is assumed to be operating in the Lamb-Dicke limit where the vibrational quantum numbers $n$ ensure that $\eta\sqrt{n+1} \ll 1$, with $\eta$ the Lamb-Dicke parameter. This condition is fulfilled even for $n$ values well higher than unity; the value $n = 1$ is associated with the vibrational ground state, the starting state for all the cold ion schemes.

**Experimental entanglement of multiple qubits in an ion trap.** The entanglement of up to four spin qubits using the Sørensen-Mølmer technique [297, 400, 401] was realized experimentally by a group at NIST in Boulder, Colorado [364]. The experiment was performed with two and four $^9\mathrm{Be}^+$ ions confined along the weak $x$-axis of a linear miniature radio-frequency trap [428]. Initially, both the center-of-mass and stretch modes were cooled to near their ground state. The characteristic states of the spin-$\frac{1}{2}$ qubit were two spectrally resolved ground state hyperfine sublevels

$$|\downarrow\rangle \ = \ {}^2\mathrm{S}_{1/2}\ \mid F=2, m_F=-2\rangle \ \text{ and } \ |\uparrow\rangle \ = \ {}^2\mathrm{S}_{1/2}\ \mid F=1, m_F=-1\rangle$$

with $F$ - total angular momentum quantum number and $m_F$ - quantum number associated with the projection of the total angular momentum along the quantization axis defined by an externally applied magnetic field.

The ion trap used in the experiment was run at the center-of-mass frequency, $\omega_x$, corresponding to the symmetric normal mode with the lowest energy. The next normal mode in frequency was the antisymmetric (stretch) mode[56]; the corresponding stretch mode frequencies were $\sqrt{3}\omega_x$ for two ions and $\sqrt{29/5}\omega_x$ for four ions. The Raman difference frequencies were tuned close to these asymmetric modes in the process of generating entanglement.

The ions were Doppler-cooled to the ground state temperature of both center-of-mass and stretch modes; at these temperatures the ions were strongly coupled by Coulomb interaction and formed a rigid crystal, a string lying along the trap main axis. The motional ground state of the ion string was prepared by repeating several times the following procedure: a pulse with $\theta = \pi$ driven at a difference frequency $\Delta\omega_L = \omega_0 - \omega_i$ (the red sideband) was cooling down the ions while optical pumping was bringing them from $|\uparrow\rangle$ state back to $|\downarrow\rangle$ state [234]. The frequency $\omega_0$ corresponds to the hyperfine splitting and the frequencies $\omega_i$ characterize the normal modes of the quantized collective motion of the ions. The initial state was $|\downarrow\downarrow\, 0\rangle$ for $N = 2$ ions and $|\downarrow\downarrow\downarrow\downarrow\, 0\rangle$ for $N = 4$ ions.

The ions in state $|\downarrow\rangle$ were detected by illuminating all the ions with laser light with 313 nm wavelength and left circular polarization; this light induces the transition to the hyperfine sublevel $^2\mathrm{P}_{3/2}\ \ |\ F = 3, m_F = -3\rangle$. Dipole selection rules dictate that this level could decay only to the $|\downarrow\rangle$ state and transition to the other hyperfine sublevel and qubit state $|\uparrow\rangle$ was forbidden.

The states $|\downarrow\rangle$ and $|\uparrow\rangle$ were coherently coupled through stimulated Raman transitions induced by shining two laser beams on the trapped ions. Controlled rotations of a single qubit were driven by laser beams tuned about midway between the states $^2\mathrm{P}_{3/2}$ and $^2\mathrm{P}_{1/2}$.

The Rabi flopping $|\uparrow\rangle \Longleftrightarrow |\downarrow\rangle$ was observed when the difference in frequency of the two Raman laser beams was close to the hyperfine splitting $\omega_0/2\pi \approx 1.25$ GHz; the state evolution during Rabi transitions was

$$|\downarrow\rangle \ \rightarrow\ \cos\frac{\theta}{2}\,|\downarrow\rangle + e^{i\phi}\sin\frac{\theta}{2}\,|\uparrow\rangle$$

$$|\uparrow\rangle \ \rightarrow\ \cos\frac{\theta}{2}\,|\uparrow\rangle - e^{i\phi}\sin\frac{\theta}{2}\,|\downarrow\rangle.$$

Here, the angle $\theta$ is proportional to the Raman pulse duration and $\phi$, the *ion phase*, is the phase difference between the Raman beams at the position of the ion. The phases $\phi_1$ and $\phi_2$ of two ions can be controlled independently by a combination of two techniques: changing the phase of the RF synthesizer that controls the Raman difference frequency and changing the strength of the ion trap field that controls the ion spacing. In this way, single qubit rotations can be performed on one ion without affecting the other.

At the end of the interaction, the ions had evolved from the initial state $|\downarrow\rangle$ into the maximally entangled states

$$|\,\varphi_e^2\rangle = \frac{1}{\sqrt{2}}(|\downarrow\downarrow\rangle - i\,|\uparrow\uparrow\rangle) \quad \text{for two ions}$$

[56]In the antisymmetric *stretch mode* alternating ions in a string oscillate out of phase in opposite directions.

and

$$|\,\varphi_e^4\rangle = \frac{1}{\sqrt{2}}(|\downarrow\downarrow\downarrow\downarrow\rangle + i\,|\uparrow\uparrow\uparrow\uparrow\rangle) \quad \text{for four ions.}$$

The motional quantum state of the ions was controlled by driving stimulated Raman transitions that couple spin and motion. The difference wave-vector resulting from the composition of the wave-vectors of the two Raman beams was oriented along the string of ions (the trap main axis) and exerted a dipole force only along this direction.

The lasers were detuned about $\Delta \approx 80$ GHz blue-sideband of the $2P_{1/2}$ excited state; their intensities give a Rabi oscillation frequency $\Omega/2\pi = 500$ kHz. The stretch-mode frequency for the two- and four-ion experiments was 8.8 MHz and the corresponding Lamb-Dicke parameter was $\eta = 0.23/N^{1/2}$. The two driving frequencies required for the entanglement operation were generated by frequency-modulating one of the Raman beams. The Raman difference frequencies were chosen close to the motional mode which had equal amplitudes of motion for all ions; thus, the Raman Rabi frequencies were the same for all ions under uniform illumination.

The two-step transition $|\downarrow\downarrow\, 0\rangle$ to $|\uparrow\uparrow\, 0\rangle$ was driven by laser beams with frequencies $(\omega_0 + \omega_i - \Delta)$ and $(\omega_0 - \omega_i + \Delta)$. When the detuning was sufficiently large in comparison with the transition linewidth given $\Delta > \eta\,\Omega$ (in this case 80 GHz $>$ 360 kHz), the intermediate states $|\uparrow\downarrow\, 1\rangle$ and $|\downarrow\uparrow\, 1\rangle$ were negligibly occupied (practically, no population of these states occurred through collective motion excitation). The interaction Hamiltonian in the rotating-wave approximation[57] and the Lamb-Dicke limit is given by

$$\mathbf{H} = \frac{\hbar\overline{\Omega}}{2}\,(|\uparrow\uparrow\rangle\langle\downarrow\downarrow| + |\downarrow\downarrow\rangle\langle\uparrow\uparrow| + |\uparrow\downarrow\rangle\langle\downarrow\uparrow| + |\downarrow\uparrow\rangle\langle\uparrow\downarrow|)$$

with $\overline{\Omega} = \eta^2\Omega^2/\Delta$, $\Omega$ is the Rabi frequency of the single ion internal transition $|\downarrow\rangle \longrightarrow |\uparrow\rangle$ and $\eta$ the Lamb-Dicke parameter. The Lamb-Dicke parameter is given by $\eta = (\hbar k^2/2M\omega_i)^{1/2}$ with $\hbar k$ the momentum transfer and $M$ the total particle mass involved in an excitation.

Entanglement is achieved by applying the Hamiltonian $H$ for a time $t = \pi/2\overline{\Omega}$; it creates the spin state $|\,\psi_2\rangle = (|\uparrow\uparrow\rangle - i\,|\downarrow\downarrow\rangle)/\sqrt{2}$. In this process, the intermediate states $|\uparrow\downarrow\, 1\rangle$ and $|\downarrow\uparrow\, 1\rangle$ are negligibly populated if, as mentioned above, $\Delta > \eta\,\Omega$. On the other hand, from the expression of the Hamiltonian we see that the entanglement speed is maximized for small $\Delta$. However, according to Sørensen and Mølmer [401] the technique still works for detuning values $\Delta \approx \eta\,\Omega$. For select values of the detuning $\Delta$ given by the condition

$$\frac{\Delta}{\eta\Omega} = 2\sqrt{m}, \quad \text{for any integer } m$$

the collective motion excitation disappears at the moment the entangled state is created. The NIST experiment was operated at a detuning

$$\Delta = 2\eta\Omega, \quad \text{for } m = 1.$$

[57] In the *rotating-wave approximation*, used in atom optics and magnetic resonance, terms in the interaction Hamiltonian which oscillate rapidly are neglected. This approximation is applied when the laser frequency is near an atomic resonance and its intensity is low.

The ion states after the entanglement operation were probed by illuminating them with a circularly polarized laser beam tuned to the cycling transition

$$2S_{1/2}(F=2, m_F=-2) \;\leftrightarrow\; 2P_{3/2}(F=3, m_F=-3).$$

Thus, each ion in state $|\downarrow\rangle \equiv 2S_{1/2}$ was identified by the bright fluorescent light emitted by the level $2P_{3/2}$ excited through this transition; the ions in state $|\uparrow\rangle$ remained dark.

Each experiment was repeated 1000 times under the same conditions in order to determine with good accuracy the probabilities $P_j$ that $j$ ions are in state $|\downarrow\rangle$. The probabilities, $P_{\downarrow\downarrow\ldots\downarrow\downarrow}$ and $P_{\uparrow\uparrow\ldots\uparrow\uparrow}$, that all $N$ ions ($N = 2, 4$) are in this very same state were large compared to the probabilities for the other cases; the probabilities for the intermediate cases were non-zero and that indicated that the entangled states of $N = 2$ and $N = 4$ ions were not generated with perfect accuracy.

To prove entanglement, the *fidelity* of the experimental state defined as

$$F = \langle \psi_e \mid \rho \mid \psi_e \rangle,$$

where the density matrix $\rho$ describes the experimental state and approximates the entangled state $\mid \psi_e\rangle$, must exceed 1/2 as stated by [364]. The fidelity can be written as a sum of diagonal and off-diagonal density matrix elements

$$F = \frac{1}{2}(P_(\downarrow \ldots \downarrow) + P_(\uparrow \ldots \uparrow)) + \mid \rho_{\downarrow\ldots\downarrow,\uparrow\ldots\uparrow} \mid .$$

Experimentally, the populations in states $|\downarrow \ldots \downarrow\rangle$ and $|\uparrow \ldots \uparrow\rangle$, $P_(\downarrow \ldots \downarrow)$ and $P_(\uparrow \ldots \uparrow)$, respectively, are measured directly by detecting the number of ions in state $|\downarrow \ldots \downarrow\rangle$; the far-off-diagonal density matrix elements $\mid \rho_{\downarrow\ldots\downarrow,\uparrow\ldots\uparrow} \mid$ are also measured experimentally after further manipulation of the ions. A Raman laser pulse with a certain duration and phase $\phi$ is applied and all ions are rotated together with an angle $\theta = \pi/2$ and phases $\phi_i = \phi$ relative to the entangling pulse. The parity $\Pi$ of the ion string was measured as a function of the phase $\phi$; for a string of $N$ ions the parity was estimated to be given by

$$\Pi = 2 \mid \rho_{\downarrow\ldots\downarrow,\uparrow\ldots\uparrow} \mid \cos N\phi$$

if only far-off-diagonal coherence elements are present (all the other off-diagonal coherence elements are absent). The value of the parity is $\Pi = +1$ if an even number of ions is in state $|\downarrow \ldots \downarrow\rangle$ and $\Pi = -1$ if an odd number of ions is in state $|\downarrow \ldots \downarrow\rangle$.

In the case of two ions the entanglement was estimated at $E \approx 0.5$. For two qubits Wootters [462] gives a formula to calculate "entanglement of formation" of an arbitrary state as a function of the density matrix of the system. This entanglement quantifies the resources necessary to create a given state. An entanglement of 0.5 indicates that two pairs of trapped ions entangled in the NIST experiment carry the same quantum information as a perfectly entangled pair. For the case of four ions there is no similar formula and the entanglement was estimated at $E_4 \approx 0.35$ from the term $\rho_{\downarrow\ldots\downarrow,\uparrow\ldots\uparrow}$ measured experimentally.

The decoherence was stronger in the case of four-ion experiment than in the two-ion one. The experiments showed that a four-qubit logic gate could be realized using trapped ions. The prevailing decoherence for the trapped ions is *collective dephasing* which transforms each

qubit state in the following way: $|\downarrow\rangle \longrightarrow |\downarrow\rangle$ and $|\uparrow\rangle \longrightarrow e^{i\zeta}\ |\uparrow\rangle$, where $\zeta$ is an unknown phase. The states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ of two qubits, if used as basis states, will protect the qubits against collective dephasing in a decoherence-free subspace; any superposition of these states is invariant under collective dephasing. A logical qubit can be defined using these two states in the decoherence-free subspace of two physical qubits.

**A single-ion `CNOT` gate using individual addressing.** The gate was realized by a group at Innsbruck University [366] using two $^{40}Ca^+$ ions in a linear Paul trap individually addressed with focused laser beams. They applied precise control of atomic phases and composite pulse sequences adapted from nuclear magnetic resonance techniques. The qubit basis states were the $S_{1/2}$ ground state and the $D_{5/2}$ metastable state (lifetime of about 1 second) as $\mid 0\rangle$ and $\mid 1\rangle$, respectively. After an initial Doppler cooling, the electronic qubit states of the two ions were initialized by optical pumping and the collective motional breathing mode was prepared in state $\mid n = 0\rangle$ using sympathetic sideband cooling; the ground state population was about 99%. The phonon number of the collective vibrational mode also forms a qubit with $\mid n = 1\rangle$ and $\mid n = 0\rangle$ as the corresponding logical states $\mid 0\rangle$ and $\mid 1\rangle$. The sequence of operations implementing the gate follows the Cirac-Zoller proposal:

(1) a $\pi$-pulse is applied on the blue sideband (all sideband pulses are on the blue sideband) of the first ion, which is the control ion, and maps the internal state of the ion to a corresponding state of the collective vibrational mode; thus, the quantum information carried by the control qubit is inscribed in the vibrational mode;

(2) the second ion is addressed and a single-ion `CNOT` gate is performed between this ion and the collective vibrational mode. If the collective mode qubit is in state $\mid 1\rangle$ (no vibration is present in this mode, $n = 0$) the internal state of the second ion is flipped. The single-ion `CNOT` gate operation is the result of a sequence of pulses on the target ion: a carrier $\pi/2$-pulse (Ramsey pulse), a combination of four pulses implementing different rotation angles (determined by the lengths of Rabi pulses) about different rotation axes (determined by the phases of the exciting radiation), and ending with another carrier $\pi/2$-pulse. The composite pulse sequence replaces the $2\pi$-rotation through an auxiliary excited level as proposed by Cirac and Zoller [97];

(3) a $\pi$-pulse on the blue sideband applied to the first ion restores the control qubit and the collective mode to their initial states. The gate operation will flip the internal state of the second ion when the first ion was prepared in the $\mid D_{5/2}\rangle$ state and will leave it in its original state if the first ion was prepared in $\mid S_{1/2}\rangle$ state. The state detection of each qubit in this experiment is performed with an efficiency of 98%. The fidelity of transferring the initially prepared eigenstates into the final target states is 80%, and the fidelity of creating an entangled state is $F = 0.71$. Higher fidelity of more than 90% could be obtained for reduced laser frequency noise and improved control of the addressing beam; that requires more complex composite pulse techniques and quantum control techniques which correct experimental imperfections, such as residual thermal excitation and inaccuracies in pulse shaping.

**Large-scale ion-trap quantum computing device.** The original proposal and the subsequent studies of an ion trap based quantum computation device assumed a string of ions confined in a single ion trap; such a system proved that quantum state could be efficiently prepared, manipulated, and read-out and gates could be executed with relatively

high fidelity. In the end it was understood that manipulation of a large number of ions in a single trap was limited by technical difficulties involving, among others, precise control of electric and magnetic fields, very accurate tailoring of laser pulses length and control of laser frequencies. The solutions proposed to avoid the limitations and to allow large-scale quantum computing systems involve small ion-trap quantum registers with quantum communication channels based on techniques such as photon coupling [121, 409], spin-dependent Coulomb interactions [99], or by a combination of radio-frequency and quasistatic electric fields [236].

Cirac and Zoller [99] proposed a scalable trapped ion implementation of a quantum computer where the ions are placed in separated, but closely spaced harmonic oscillator wells and are individually addressed by laser beams. A laser pulse applies a momentum kick to the two ions dependent on the state of the first ion, as in the original proposal. But this time, the combination of the wave packets undergoing a displacement in the harmonic potential of each ion and their mutual Coulomb repulsion results in a state-dependent phase shift and that can be used as a universal quantum gate. A scaled up version would contain an array of separate, single-ion traps; a "head" ion which can be moved in the vicinity of any of the ions trapped in the array would be used as the control bit in a multiple `CNOT` gate.

The large-scale ion-trap array of Kielpinski, Monroe, and Wineland [236] has a "quantum charge-coupled device" (QCCD) type of architecture. It has a large number of interconnected traps; a few ions can be confined at a time in each trap or moved from storage (memory) to interaction registers just by changing the radio-frequency (RF) and the quasistatic electric (DC) voltages applied to the trap electrodes. The electrodes are segmented and arranged in three layers: two outer layers where RF voltages are applied and a middle layer for the quasistatic fields so that ions can be individually confined in a particular region or moved along the local trap axis. Ions encoding quantum information are stored in memory registers and are transported to an interaction region when a logical gate has to be performed. Gates are driven by lasers beams focused through the interaction region where the ions are trapped close together to allow Coulomb coupling to play its role in entangling qubits.

The QCCD-type architecture proposed uses quantum manipulation techniques that have been experimentally demonstrated for small quantum registers [61]. Solid state ion traps have been built from alumina with gold evaporated electrodes [233] or with heavily doped silicon electrodes, using standard microfabrication techniques. Sympathetic cooling using a second species of ions as a heat sink in the interaction region will keep gate errors at a low level; decoherence during ion storage, gate execution, and during ion transport can be reduced by several orders of magnitude by encoding each qubit into a decoherence-free subspace of two ions [235].

Trapped ions and atoms in small experimental systems have coherence times many orders of magnitude longer than times required for initialization, qubit control, and measurement. Preserving the high-fidelity control demonstrated in these small systems while scaling to larger architectures represents s serious challenge.

## 6.6 Nuclear Magnetic Resonance - Ensemble Quantum Computing

The magnetic induction by nuclear spins, the physical phenomenon on which *nuclear magnetic resonance* (NMR) is based, was first observed in 1946 by the group of Edward Purcell [347]

and by Felix Bloch [63]. Many atomic nuclei have a non-zero magnetic moment and behave as small bar magnets in response to applied external magnetic fields. The magnetic moment of a single nucleus is hard to detect by current technologies, but in a sufficiently large population of nuclei the contributions of the nuclear spins add up and they can be observed as an ensemble using the nuclear magnetic resonance techniques. NMR has been used for more than forty years to study chemical structure, dynamics and reactions; it has become a relatively widely used medical imaging procedure under the name MRI (Magnetic Resonance Imaging).

The very long coherence times associated with the nuclear spins motivated David Cory [104, 105] and, independently, Neil Gershenfeld and Isaac Chuang [171] to propose in mid 1990s the nuclear magnetic resonance as a working concept for a quantum computer. The standard NMR spectroscopy and the NMR implementation of quantum computing use the same device and similar instrumentation, but are quite different in their goals. The liquid-state NMR has so far been used to demonstrate Shor's factoring algorithm by using a seven-qubit custom-synthesized molecule [94, 432] to factor the number 15.

One of the molecules used for quantum computing proof of concept is that of trichloroethylene, $Cl_3$ H $C_2$, where the atom of the regular $^{12}$C isotope (a spin-zero nucleus) is replaced with that of the $^{13}$C isotope (a spin-1/2 nucleus). The hydrogen nucleus (a proton) of each trichloroethylene molecule has a strong magnetic moment; when the liquid sample which contains a very large number of $Cl_3$ H $C_2$ molecules is placed in a powerful external magnetic field, the spin of each proton aligns itself with this magnetic field which determines the $z$-axis of the system.

The spins can be induced to change direction with radio frequency (RF) pulses and as soon as they tip off-axis the applied static magnetic field induces a rapid precession of the proton spins. The precession represents a rotation of the spin vector around the main axis, in this case the $z$-axis determined by the external magnetic field; the precession frequency, also called Larmor frequency, $\omega_{\rm L}$, is related to the strength $\mathbf{B}_{\rm ext}$ of the external field

$$\omega_{\rm L} = \mu_{\rm p}\mathbf{B}_{\rm ext}$$

where $\mu_{\rm p} = 42.7$ MHz/Tesla is the magnetic moment of the proton. In a typical magnetic field of strength $\mathbf{B}_{\rm ext} = 11.7$ Tesla, the precession frequency is $\omega_{\rm L} = 500$ MHz.

The precession of the proton (hydrogen nucleus) spin induces oscillating currents tuned to the precession frequency in a coil placed around the sample. These oscillating currents induce a magnetic field which is detected, thus, allowing the observation of entire ensemble of protons in the sample. This is the magnetic induction by nuclear spin [347, 63] observed in 1946. At least $10^{15}$ molecules have to be present in the sample to be able to observe the magnetic induction signal. In general, the NMR sample contains about $10^{18}$ molecules.

NMR has important applications such as molecular structure determination, molecular dynamics studies in liquid and solid state; the application of NMR to quantum computing is based on some of the same techniques as the molecular structure determination by NMR. But, instead of studying new molecules and trying to determine their structure, it uses synthesized molecules with well-defined nuclear spins which are manipulated as quantum information qubits; the molecule internal Hamiltonian and relaxation operator are precisely known and define the effective overall transformation produced by a complex series of operations applied to that sample.

In fact, the magnetic induction of spin had an early connection to information storage.

In the mid 1950s, the magnetic induction was proposed to be used for classical information storage by Arthur Anderson and coworkers [11, 12] based on the *spin echo technique.* The spin echo technique corrects for the inhomogeneity of the applied static magnetic field. When the magnetic field is not constant across the sample the precession frequency varies with the location of the spins. In this case the induction magnetic fields produced by the spins are not aligned and they do not add up to produce a good magnetic induction signal; as a result, the induction signal vanishes.

With the spin echo technique, two pulses of the RF field are applied separated by a time interval $t$ and a strong nuclear induction signal is observed at a time $t$ after the second pulse. The spin echo reverses the spins and, thus, reverses their precession, until their induction magnetic fields are all aligned again. The spin relaxation process, the main source of decoherence in this case, tends to realign the spins with the applied magnetic field (the transverse magnetization decays to zero in time), randomizes their phases and leads to complete loss of the information encoded in the spins.

The spins, in general, are susceptible only to magnetic field action but, since the spins of the nuclei are well shielded inside the atom from most sources of fluctuating magnetic fields, their relaxation times can be very long, of the order of thousands of seconds. In liquid state, the interplay between the weak nuclear interactions on one hand and the fast averaging effect due to the rapid tumbling motions of the molecules in the liquid on the other hand, makes for relaxation times of proton spins only of the order of tens of seconds, but long enough for quantum computation purposes.

## 6.7 Liquid-State NMR Quantum Computer

We discuss now the implementation of the qubits, the control Hamiltonian, and the qubit control for a liquid NMR computer.

**The qubits.** Spin-1/2 nuclei are the preferred two-state quantum systems for implementing qubits in liquid-state nuclear magnetic resonance (NMR) quantum computation. The application of the strong external magnetic field $\mathbf{B}_0$ determines the $z$-axis direction and the state space of the spin becomes a superposition of "up" and "down" states defined along this direction. The logical states $\mid 0\rangle$ and $\mid 1\rangle$ correspond to spin "up" and spin "down" states, respectively, and have different energies in the external static magnetic field; this is the so-called Zeeman splitting. The liquid NMR decoherence times are typically long because the nuclear spins interact only with magnetic fields and not with electric fields and, moreover, the atomic/molecular electron cloud shields the nucleus spin from most sources of fluctuating magnetic fields.

Each molecule can be viewed as a single quantum computer, the state of which is determined by the orientation of its spins. Quantum logic gates are constituted of radio frequency (RF) pulses; they manipulate spin orientations and spin couplings and perform unitary transformations on the state. The N-spin molecules in the NMR sample form an ensemble of N-qubit computers. This ensemble of independent quantum computers is supposed to be used in a global way without addressing them individually.

The sample containing about $10^{18}$ $N$-spin molecules is placed in a strong longitudinal static magnetic field $\mathbf{B}_0$ and has a transverse RF magnetic field applied to it, the same way

as in a conventional pulsed NMR device. The strong magnetic field $\mathbf{B}_0$ induces the Zeeman splitting of the energy levels of the spin system in each molecule. For a two-level system, such as uncoupled protons, at thermal equilibrium the ratio of the populations in the higher and the lower energy levels is 0.999999, i.e., it differs from unity by as little as $10^{-6}$. This is a small difference, but it produces a macroscopic precessing magnetization which can be detected in a pickup coil.

The time evolution of a spin-1/2 nucleus in the magnetic field $\mathbf{B}_0$ along the $z$-axis is governed by the Hamiltonian operator

$$\mathbf{H}_0 = -\hbar\ \gamma\ \mathbf{B}_0\ \mathbf{I}_z = -\hbar\ \omega_0\ \mathbf{I}_z = -\frac{\hbar\ \omega_0}{2}\ \sigma_z = \begin{pmatrix} -\hbar\omega_0/2 & 0 \\ 0 & \hbar\omega_0/2 \end{pmatrix}$$

where $\gamma$ is the gyromagnetic ratio of the nucleus, $\mathbf{I}_z$ is the angular momentum of the nuclear spin in the $z$ direction, and $\omega_0/2\pi$ (where $\omega_0 = \gamma\ \mathbf{B}_0$) is the *precession frequency* or *Larmor frequency*. Sometimes $\omega_0$ alone is called the Larmor frequency and the factor $2\pi$ is implicit. The three components of the nuclear magnetic moment, $\mathbf{I}_x$, $\mathbf{I}_y$, and $\mathbf{I}_z$ are related to the Pauli matrices

$$\sigma_x = 2\mathbf{I}_x, \quad \sigma_y = 2\mathbf{I}_y, \quad \sigma_z = 2\mathbf{I}_z.$$

The energy of the spin up state, $|\uparrow\rangle$ or $|\ 0\rangle$ given by

$$\langle 0 \mid \mathbf{H}_0 \mid 0\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} -\hbar\omega_0/2 & 0 \\ 0 & \hbar\omega_0/2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} -\hbar\omega_0/2 \\ 0 \end{pmatrix} = \frac{-\hbar\omega_0}{2}$$

is lower than the energy of the spin down state, $|\downarrow\rangle$ or $|\ 1\rangle$ given by

$$\langle 1 \mid \mathbf{H}_0 \mid 1\rangle = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} -\hbar\omega_0/2 & 0 \\ 0 & \hbar\omega_0/2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \hbar\omega_0/2 \\ 0 \end{pmatrix} = \frac{\hbar\omega_0}{2}.$$

The energy difference between the two states is

$$\Delta E = \hbar\omega_0/2 - (-\hbar\omega_0/2) = \hbar\omega_0.$$

The precession frequency is proportional to the energy difference $\Delta E$ between the "up" and "down" states

$$\omega_0 = \frac{\Delta E}{\hbar}$$

where $\hbar$ is the reduced Planck constant.

The time evolution of the state under the action of the external magnetic field [105] is described by

$$|\ \psi_t\rangle = e^{-i\omega_0 t/2}\alpha_0 \mid 0\rangle + e^{i\omega_0 t/2}\alpha_1 \mid 1\rangle,$$

where the units for $\omega_0$ are radian/ second when the time, $t$, is given in seconds.

For liquid-state NMR the typical values for the static magnetic field $\mathbf{B}_0$ are in the range $5 - 15$ Tesla. The precession frequencies [434] depend on the nuclei's magnetic moments and

can vary, for example, from 500 MHz for a proton (hydrogen nucleus) to 125 MHz for a $^{13}$C nucleus in a trichloroethylene molecule in an external magnetic field of 11.74 Tesla.

The differences in the precession frequencies help to distinguish between the types of nuclei placed in a given magnetic field; thus, they allow control and measurement of specific nuclear spins. Spins of the same nuclear species which are part of the same molecule can also have different precession frequencies. These differences between precession frequencies of the same nuclear species are called *chemical shifts* and they arise from the variable partial shielding of the applied magnetic field by the electron cloud surrounding a nucleus at a specific location inside the molecule. Different chemical environments inside a molecule determine different chemical shifts for nuclei of the same species. Large asymmetries in the molecular structure determine large chemical shifts. Typical chemical shifts range from a few tens to a few hundreds parts per million of the precession frequency. In the trichloroethylene molecule at 11.7 Tesla magnetic field the chemical shifts are between a few kHz and a few tens of kHz. They depend on the solvent, the molecule concentration in the solution, and the temperature.

The evolution in time of the initial state $\mid \psi_0\rangle$ can be expressed in terms of the Pauli matrix $\sigma_z$ as

$$\mid \psi_t\rangle = e^{i\frac{1}{2}\omega_0\sigma_z t} \mid \psi_0\rangle.$$

The operator $\omega_0\sigma_z/2$ represents the internal Hamiltonian of the spin (i.e., the energy observable,here given in units for which the reduced Planck constant $\hbar = h/(2\pi) = 1$). The Pauli matrix $\sigma_z$ can be thought of as the observable for the nuclear spin along the $z$-axis which is defined by the external static field. The observables for spin along the $x$- and $y$-axis are given by the corresponding Pauli matrices $\sigma_x$ and $\sigma_y$. The Pauli matrices are the spin observables in the laboratory frame and this representation is meaningful in the real space. In the Bloch sphere representation the time evolution under the action of the Hamiltonian $\mathbf{H}_0$, the operator associated with the static applied field $\mathbf{B}_0$, is described as a precessing motion (rotation) of the spin state vector about the $z$-axis which is the direction of the magnetic field $\mathbf{B}_0$. For a system of $n$ uncoupled nuclei the Hamiltonian $\mathbf{H}_0$ is

$$\mathbf{H}_0 = -\sum_{i=1}^{n} \frac{1}{2}\hbar\, \omega_0^i\, \sigma_z^i.$$

The internal Hamiltonian of a molecule's nuclear spins is well approximated by

$$\mathbf{H} \;=\; \frac{1}{2}\sum_i \omega_i \sigma_z^i \;+\; \frac{\pi}{2}\sum_{i\neq j} J_{ij}\sigma_z^i\sigma_z^j,$$

where the summation is over the nuclear spins, $\omega_i$ are resonance frequencies, and $J_{ij}$ are scalar spin couplings. The dominant term of the internal Hamiltonian represents the action of a strong magnetic field applied to the sample. The Zeeman effect of this magnetic field determines an axis of quantization along which the spinors $\sigma_z$ sum up; the Zeeman effect is very small and more than $10^{15}$ spins are necessary to produce an observable signal. A typical sample for liquid state NMR contains about $10^{18}$ molecules; each molecule acts as an independent processor. An NMR spectrometer records the average state of the spins.

From the point of view of a quantum computer implementation, the single-particle terms in the Hamiltonian above are used to distinguish qubits and the two-particle terms represent

the building blocks of two-qubit `CNOT` gates. The resonance frequencies $\omega_i$ in the first term depend on the efficiency of screening of the nuclear spins from the applied magnetic field by the surrounding electrons; the scalar spin couplings $J_{ij}$ in the second term are mediated by the electrons in molecular orbits that overlap both nuclear spins. Besides the interaction with the externally applied magnetic field $\mathbf{B}_0$, the nuclear spins in molecules feel the effects of a direct coupling and an indirect coupling between them.

The *direct coupling* is due to a *magnetic dipole-dipole interaction* between two neighboring nuclear spins belonging either to the same molecule or to different molecules; the spins behave like bar magnets. For molecules in a liquid solution dipolar (direct) spin couplings average out due to the tumbling motions of the molecules and they have no effect on the Larmor precession. In solids the dipolar couplings can be averaged out by technical means, for example, by applying multiple-pulse sequences or by spinning the sample at a specific angle with respect to the magnetic field.

The *indirect coupling* between nuclear spins in a molecule, also called *J coupling* or *scalar coupling*, is mediated by the electrons shared in the chemical bonds between the atoms of the molecule; the shared electron wave functions of the two bonded atoms overlap. This is a so-called *Fermi contact interaction* and its *through-bond coupling strength*, $J$, depends on the nuclear species involved and the number of chemical bonds connecting the atoms. Typical values for the frequencies associated with the $J$ coupling are a few hundred Hz for one-bond couplings and decrease to just a few Hz for three- or four-bond couplings. The effect of this scalar coupling on the nuclear spins is that of a static internal magnetic field along the $(+/-)z$-axis produced by neighboring nuclear spins, in addition to the externally applied static magnetic field $\mathbf{B}_0$.

The Hamiltonian operator associated with $J$ coupling between spins of nuclei $i$ and $j$ takes the form [434]

$$\mathbf{H}_J = \hbar \sum_{i,j}^{n} 2\pi J_{ij}\ \vec{\mathbf{I}}_z^i\ \vec{\mathbf{I}}_z^j = \hbar \sum_{i,j}^{n} \frac{\pi}{2} J_{ij}\ (\mathbf{I}_x^i\ \mathbf{I}_x^j + \mathbf{I}_y^i\ \mathbf{I}_y^j + \mathbf{I}_z^i\ \mathbf{I}_z^j), \qquad i < j$$

when the Larmor frequencies of different nuclear spin species satisfy the condition

$$\mid \omega_0^i - \omega_0^j \mid \quad \gg 2\pi \mid J_{ij} \mid .$$

The Hamiltonian operator $\mathbf{H}_J$ shifts the energy levels of $J$ coupled nuclei and modifies the Larmor frequency of spin $i$ as a function of the state of the spin $j$:

- if spin $j$ is in state $\mid 0\rangle$, the Larmor frequency of spin $i$ is shifted by $-J_{ij}/2$ and becomes $\omega_0^i - J_{ij}/2$;
- if spin $j$ is in state $\mid 1\rangle$, the Larmor frequency of spin $i$ is shifted by $+J_{ij}/2$ and becomes $\omega_0^i + J_{ij}/2$.

Here $\omega_0$ is the Larmor frequency of the un-coupled spin $i$. In the presence of an external magnetic field, the energy spectrum for an un-coupled spin $i$ has a line centered at frequency $\omega_0^i$, while the energy spectrum of spin $i$ in a coupled system with spin $j$ shows two lines separated by an interval $J_{ij}$ and centered around frequency $\omega_0^i$.

**The control Hamiltonian.** In liquid-state NMR the state of a spin-1/2 nucleus in a static magnetic field $\mathbf{B}_0$ can be manipulated by applying a radio-frequency (RF) electromagnetic field $\mathbf{B}_1$, perpendicular to the $z$-axis determined by the field $\mathbf{B}_0$ and rotating in the $x$-$y$ plane with a frequency $\omega_{\mathrm{RF}}$ near or equal to the nuclear spin precession frequency $\omega_0$.

The single-spin Hamiltonian operator $\mathbf{H}_{\mathrm{RF}}$, associated with the RF field $\mathbf{B}_1$, is similar to $\mathbf{H}_0$, the Hamiltonian associated with the static magnetic field $\mathbf{B}_0$,

$$\mathbf{H}_{\mathrm{RF}} = -\hbar\ \gamma\ B_1[\cos(\omega_{\mathrm{RF}}t + \phi)\mathbf{I}_x - \sin(\omega_{\mathrm{RF}}t + \phi)\mathbf{I}_y]$$

or,

$$\mathbf{H}_{\mathrm{RF}} = -\frac{1}{2}\hbar\ \gamma\ B_1[\cos(\omega_{\mathrm{RF}}t + \phi)\sigma_x - \sin(\omega_{\mathrm{RF}}t + \phi)\sigma_y]$$

where $\gamma$, $\phi$, and $B_1$ are the nuclear gyromagnetic ratio, the phase of the RF field, and its amplitude, respectively. The frequency of the RF field, $\omega_{\mathrm{RF}}$, is determined as $\omega_{\mathrm{RF}} = \gamma B_1$. The minus sign in front of the sine term shows that the spin evolution under the action of the RF field, $\mathbf{B}_1$, is in the same sense as its evolution under the field $\mathbf{B}_0$. In practice, $\mathbf{B}_1$ is a magnetic field applied perpendicular to the direction of the static magnetic field $\mathbf{B}_0$ and which oscillates along a fixed axis in the laboratory. Such an oscillating field can be decomposed into two counter rotating fields: one field rotates with frequency $\omega_{\mathrm{RF}}$ in the same direction as the spin precessing under field $\mathbf{B}_0$ and can be tuned at or near resonance with the spin, $\omega_{\mathrm{RF}} \cong \omega_0$; the other field component rotates in opposite direction and is very far from resonance by as much as $2\omega_0$ and thus, it has a negligible shift effect on the Larmor frequency.

The Larmor precession due to the nuclear spin interaction with a static external magnetic field and the couplings between spins of different nuclear species inside a molecule are characteristics of a particular system and can not be changed. However, the amplitude $B_1$, the frequency $\omega_0$, and the phase $\phi$ of the RF field can be varied in time and, thus, they are enabling our quantum control of an NMR system.

The effects of the static field and the resonant pulses are easier to understand if we use a "rotating frame of reference" instead of the laboratory frame. We imagine a coordinate system $x'\ y'\ z$ attached to the NMR apparatus and the sample and all of them rotating about the $z$-axis of the system which is aligned with the static magnetic field $\mathbf{B}_0$. The Hamiltonian for a spin placed in the static field, $\mathbf{B}_0$, and acted upon by the pulsed field, $\mathbf{B}_1$, can be expressed in the rotating frame as

$$\mathbf{H}_{\mathrm{rot}} = -\frac{1}{2}\hbar\ (\omega_0 - \omega_{\mathrm{rotframe}})\sigma_z - \frac{1}{2}\hbar\ \omega_{\mathrm{RF}}[\cos\phi\ \sigma_x - \sin\phi\ \sigma_y].$$

When the frequency of rotation, $\omega_{\mathrm{rotframe}}$, is equal to the Larmor frequency, $\omega_0$, the first term disappears in the expression above; thus, the effect of the static applied field in the rotating frame is cancelled, the nuclear spin appears stationary (does not precess) with respect to the field $\mathbf{B}_0$. In the rotating frame the spin precesses about the field $\mathbf{B}_1$ and this motion is called *nutation*.

In the case of a molecule containing several species of nuclear spins we have to consider a rotating frame for each of them; at resonance the frequency of each rotating frame, $\omega^i_{\mathrm{rotframe}}$, will be the characteristic Larmor frequency of the corresponding nucleus. In the aggregate rotating frame the Hamiltonian of the system takes the simplified form

$$\mathbf{H}_{\text{system}} = \mathbf{H}_J = \hbar \sum_{i<j}^{n} 2\pi J_{ij} \mathbf{I}_z^i \mathbf{I}_z^j = \frac{1}{2}\ \hbar\pi \sum_{i<j}^{n} J_{ij}\ \sigma_z^i \sigma_z^j,$$

where the terms of the form $\hbar\omega_0^i \sigma_z^i$ related to the static applied field are dropped and the remaining terms, $J_{ij}\sigma_z^i\sigma_z^j$, represent the couplings between the nuclear spins; they are invariant with respect to the magnetic fields applied.

In the rotating frame, the RF field, $\mathbf{B}_1$, applied along the $x'$-axis and oscillating at a frequency equal to that of the rotating frame (i.e., the spin Larmor frequency), $\omega_{\text{RF}} = \omega_{\text{rotframe}} = \omega_0$, is at *resonance* and appears as constant. This field has enough energy to induce a rotation of the spin about its direction, the $x'$-axis, and that represents a change of the spin state (qubit state). The angle of rotation, $\theta$, is determined by the pulse length, $\tau$ (the length of time the field is on) and its magnitude, $B_1$:

$$\theta = 2\pi\gamma\tau B_1.$$

For example, assume a spin oriented parallel to the $z$-axis of its rotating frame after the application of the static field $\mathbf{B}_0$. A so-called $\pi/2$ pulse of the $\mathbf{B}_1$ field rotates the spin by $90°$ about the $x'$-axis and brings it parallel to the $y'$-axis. Typically, the length (duration) of a $\pi/2$-pulse is $\approx 10$ microseconds when applying 50 to 300 Watts of power [433].

In the laboratory frame, the oscillating field $\mathbf{B}_1$ is decomposed into two components rotating in the $x'y'$-plane, one in the same direction as the spin at frequency $\omega_0$ and the other in the opposite direction at frequency very far off-resonance (by about $2\omega_0$); as a result, the spin will spiral around the $z$-axis down to the $x'y'$-plane under the action of the resonant component. A so-called $180°$ pulse rotates the spin by 180 degrees about the $x'$-axis and brings it along the $-z$-axis.

The total NMR Hamiltonian for multiple nuclear spins (each with its rotating frame) and for multiple rotating RF fields (one RF field for each of the $k$ nuclear spins) is made up of two terms:

$$\mathbf{H}_{\text{NMR}} = \mathbf{H}_{\text{system}} + \mathbf{H}_{\text{control}}$$

with

$$\mathbf{H}_{\text{system}} = \frac{1}{2}\ \hbar\pi\ \sum_{i,j} J_{ij}\ \sigma_z^i\ \sigma_z^j \qquad (i \neq j)$$

and

$$\mathbf{H}_{\text{control}} = \sum_{i,r} -\frac{1}{2}\ \hbar\omega_{\text{RF}}^k \{\cos[(\omega_{\text{rotframe}}^k - \omega_0^i)t + \phi^r]\sigma_x^i - \sin[(\omega_{\text{rotframe}}^k - \omega_0^i)t + \phi^k]\sigma_y^i\}$$

where the frequencies $\omega_{\text{RF}}^k$ (with $\omega_{\text{RF}}^k = \gamma_i B_1^k$) and phases $\phi^k$ of the multiple RF fields, $\mathbf{B}_1^k$, can be controlled by the user.

**Qubit control.** The qubit state changes only if the pulsed RF magnetic field is applied. The magnetic field of the resonant pulse appears like a constant field applied in the $x - y$ plane of the rotating frame, for example, along the $(x)$-axis of this frame. Now the nuclear spin rotates only about the $x$-axis, the axis of the pulse, and executes the following sequence

of motions: it starts along the $z$-axis, tips towards the $(y)$-axis, continues to $(-z)$-axis, then to the $-y$-axis and back to $z$-axis. The spin/qubit state evolves in time as:

$$| \psi_t\rangle = e^{-i \ \frac{1}{2} \ \omega_{\mathrm{RF}}^{(x)} \ \sigma_x \ t} \ | \psi_0\rangle$$

where $\omega_{\mathrm{RF}}^{(x)}$ is the nutation frequency (nutation about $x$-axis) ; it determines the strength (energy) of the pulse. If the frame rotates with a frequency $\omega_{RF}^{x} = \omega_0$, an observer in the rotating frame (Figure 146(a)) will see the spin precess about $B_1$, while an observer in the laboratory frame (Figure 146(b)) will see the spin rotate about the $z$-axis spiral down over the surface of the Bloch sphere.

If the pulse is turned on for a time interval $t = \phi/\omega_{x\mathrm{RF}}^{(x)}$ the qubit state rotates by an angle $\phi$ which is called the *phase of the pulse* with respect to the $x$-axis. Similar rotations can be implemented around any axis of the rotating frame. Rotations around the $z$-axis are simpler to implement and can be executed exactly. General rotations on the Bloc sphere can be constructed as a series of rotations around the main axes of the rotating frame.

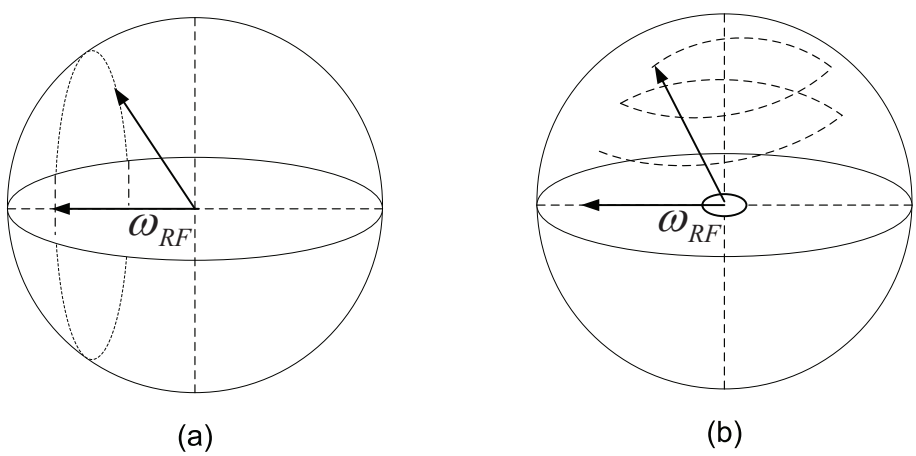



Figure 146: Spin nutation in a pulsed transverse RF field $\mathbf{B}_1$; $\omega_{RF}$ is the nutation frequency. In the rotating frame the RF field lies along a fixed axis. If the frame rotates with a frequency $\omega_{RF}$ (also denoted $\omega_{RF}^{x}$ in the case where the RF field lies along the $x$-axis) and $\omega_{RF}^{x} = \omega_0$, (a) an observer in the rotating frame will see the spin precess about $\mathbf{B}_1$, while (b) an observer in the laboratory frame sees the spin spiral down over the surface of the Bloch sphere.

If the RF field is off-resonance with respect to the spin precession (Larmor) frequency, that is, $\omega_0 - \omega_{\mathrm{RF}}^{(x)} = \Delta\omega$, the spin precesses in the rotating frame about an axis tilted away from the RF field axis, the $x$-axis in our example. The tilt angle is inverse proportional to $\Delta\omega$: if the RF pulse is far off resonance, the angle is very small. Therefore, if the Larmor frequencies are well separated, any one nuclear spin (qubit) can be selected and rotated without rotating the others.

All the axes are defined in the rotating frame or the *logical frame* with respect to the qubit implemented as nuclear spin. Each spin of interest in the molecule has its own rotating (logical) frame associated with it.

The pulsed RF field applied to one nuclear spin is felt by other nuclear spins. The spins of other nuclear species have different precession frequencies and the difference at $B_0 = 11.7$

T is large, of the order of many MHz. A pulse which is in resonance with the target spin is not constant in the rotating frame of the non-target nuclear species; it rotates rapidly and its effect is not significant. However, the spins of non-target nuclei of the same species as the target will be affected by the resonant pulse which appears as a constant field in their rotating frame; they will be rotated with the same angle as the target. The chemical shifts are too small (a few kHz) to make a difference.

How do we control a particular spin (qubit)? The practical solution to individually rotating nuclear spins of the same species is the following: (i) rotate first all these spins by a given angle by applying a hard RF pulse for a time as short as possible; and then (ii) rotate one spin which has a specific precession frequency by applying a soft pulse of sufficient duration. The power of this soft pulse is modulated in time to reduce the time needed for a rotation while minimizing crosstalk with other nuclear spins. For example, in the case of a two-spin system at equilibrium, a "hard" (nonselective) $\pi/2$-pulse along $y$-axis yields $\mathbf{I}_x^1 + \mathbf{I}_x^2$, while a "soft" (spin-selective) $\pi/2$-pulse along $y$-axis whose frequency spans the resonance frequency of only the first spin yields $\mathbf{I}_x^1 + \mathbf{I}_z^2$, leaving the second spin unaffected.

## 6.8 NMR Implementation of Single-Qubit Gates

Single-qubit gates, spin state vector rotations in Hilbert space, can be directly implemented in the rotating frame by applying RF pulses to individual nuclear spins. When an RF pulse of amplitude $\gamma B_1$ is applied to a single nuclear spin the rotating frame of which rotates at a frequency $\omega_{\rm RF} = \omega_0$, the spin evolution is determined by the unitary transformation

$$\mathbf{U} = \exp\left(-i\frac{1}{\hbar}\,\mathbf{H}_{\rm control}\;t_{\rm pulse}\right) = \exp\left(i\frac{1}{2}\gamma B_1\;[\cos(\phi)\sigma_x - \sin(\phi)\sigma_y]\;t_{\rm pulse}\right)$$

where $t_{\rm pulse}$ is the RF pulse duration. In the Bloch sphere representation, this unitary transformation describes a rotation about an axis in the $x$ - $y$ plane. The direction of the axis of rotation is determined by the phase $\phi$ of the pulse and the angle of rotation, $\theta$, is proportional to the product between $t_{\rm pulse}$ and the nutation frequency $\omega_{\rm RF} = \gamma B_1$.

A general rotation with an angle $\theta$ about an axis of rotation $\vec{\mathbf{n}}$ (with $\vec{\mathbf{n}}$ a three-dimensional vector) can be written as

$$\mathcal{R}_{\mathbf{n}}(\theta) = \exp\left(-i\frac{\theta}{2}\;\vec{\mathbf{n}}\cdot\vec{\boldsymbol{\sigma}}\right)$$

where $\vec{\boldsymbol{\sigma}} = \sigma_x x + \sigma_y y + \sigma_z z$ is a vector of Pauli matrices. For example, an RF pulse with phase $\phi = \pi$ and $\omega_{\rm RF}\;t_{\rm pulse} = \pi/2$ will perform the unitary transformation

$$\begin{aligned}
\mathbf{U} &= \exp\{i\tfrac{1}{2}\gamma B_1\;t_{\rm pulse}[\cos(\phi)\sigma_x - \sin(\phi)\sigma_y]\} \\
&= \exp\{i\tfrac{1}{2}\omega_{\rm RF}\;t_{\rm pulse}[\cos(\phi)\sigma_x - \sin(\phi)\sigma_y]\} \\
&= \exp\left(-i\tfrac{1}{2}\tfrac{\pi}{2}\sigma_x\right)
\end{aligned}$$

which is a 90° rotation about $x$-axis, $\mathcal{R}_x(90)$; a pulse with the same phase, $\phi = \pi$, but twice as long, $\omega_{\rm RF}\;t_{\rm pulse} = \pi$, realizes a transformation

$$\mathbf{U} = \exp\left(i\frac{1}{2}\omega_{\mathrm{RF}}t_{\mathrm{pulse}}[\cos(\phi)\sigma_x - \sin(\phi)\sigma_y]\right) = \exp\left(-i\frac{1}{2}\pi\sigma_x\right)$$

that is, a 180° rotation $\mathcal{R}_x(180)$. An RF pulse with phase $\phi = 0$ determines a rotation $\mathcal{R}_x(-90)$, a negative rotation about the $x$-axis (i.e., a rotation with 90° about the $(-x)$-axis.

RF pulses with phases $\phi = \pm\pi/2$ implement positive and, respectively, negative rotations about $y$-axis. For example, if $\phi = \pm\pi/2$ and $\omega_{\mathrm{RF}}t_{\mathrm{pulse}} = \pi/2$, then

$$\mathbf{U} = \exp\left(i\frac{1}{2}\omega_{\mathrm{RF}}t_{\mathrm{pulse}}[\cos(\phi)\sigma_x - \sin(\phi)\sigma_y]\right) = \exp\left(\mp i\frac{1}{2}\frac{\pi}{2}\sigma_y\right)$$

and that represents a transformation $\mathcal{R}_y(\pm 90)$, a 90° rotation about the $\pm y$-axis.

Thus, in practice, the RF field does not have to be applied along different axes in the lab frame. In the rotating frame, the relative phase of the RF field determines the direction of the nutation axis in the $x-y$ plane. The absolute phase of the first RF pulse determines only a reference value against which the phases of all the subsequent RF pulses on the same spin, as well as the value of that spin have to be estimated.

An arbitrary single-qubit unitary transformation, $U$, can be implemented as a sequence of rotations about only two axes, such as $x$- and $y$-axis, (see [284], Section 4.3)

$$\mathbf{U} = e^{i\alpha}\mathcal{R}_x(\beta)\mathcal{R}_y(\gamma)\mathcal{R}_x(\delta)$$

where $\alpha, \beta, \gamma, \delta$ are real numbers. The rotations about the $z$-axis [171] can be decomposed in terms of rotations about $x$- and $y$-axis:

$$\mathcal{R}_z(\theta) = \mathcal{R}_x(90)\mathcal{R}_y(\theta)\mathcal{R}_x(-90)$$

and

$$\mathcal{R}_z(\theta) = \mathcal{R}_y(90)\mathcal{R}_x(-\theta)\mathcal{R}_y(-90).$$

Since rotations about $z$-axis are easy to implement, they can be used to produce the respective rotations about either $x$- or $y$-axis.

## 6.9 NMR Implementation of Two-Qubit Gates

Two-qubit gates can be directly implemented through the action of the nuclear spin-spin coupling Hamiltonian, $\mathbf{H}_J$, estimated for molecules in a liquid; this Hamiltonian is invariant with respect to the laboratory or the rotating reference frame. The corresponding time evolution operator is

$$\mathbf{U}_J(t) = \exp\left(-i\frac{1}{2}\pi J_{12}\sigma_z^{(1)}\sigma_z^{(2)}t\right).$$

In matrix form

$$U_J(t) = \begin{pmatrix} e^{-i\pi J_{12}t/2} & 0 & 0 & 0 \\ 0 & e^{+i\pi J_{12}t/2} & 0 & 0 \\ 0 & 0 & e^{+i\pi J_{12}t/2} & 0 \\ 0 & 0 & 0 & e^{-i\pi J_{12}t/2} \end{pmatrix}.$$

For a time interval $t = 1/(2J_{12})$ the transformation under the action of the nuclear spin-spin coupling becomes

$$U_J(\frac{1}{2J_{12}}) = \begin{pmatrix} e^{-i\pi/4} & 0 & 0 & 0 \\ 0 & e^{+i\pi/4} & 0 & 0 \\ 0 & 0 & e^{+i\pi/4} & 0 \\ 0 & 0 & 0 & e^{-i\pi/4} \end{pmatrix}.$$

The transformation $U_J(\frac{1}{2J_{12}})$, a two-qubit 90° rotation associated with $\sigma_z^{(1)}\sigma_z^{(2)}$, is equivalent to a combination of two gates, each conditional on the logical state of qubit 1. The first gate applies a 90° rotation ($\exp(-i\sigma_z^{(2)}\pi/4)$) about the $z$-axis to qubit 2 conditional on qubit 1 being in state $\mid 0\rangle_1$. The second gate applies a $-90°$ rotation ($\exp(i\sigma_z^{(2)}\pi/4)$) about the $z$-axis to qubit 2 if qubit 1 is in state $\mid 1\rangle_1$. Now, after the two-qubit rotation, we apply a $-90°$ rotation ($\exp(i\sigma_z^{(2)}\pi/4)$) about the $z$-axis to qubit 2. As a result, the 90° rotation performed by the first gate if qubit 1 is in state $\mid 0\rangle_1$ is cancelled. If qubit 1 is in state $\mid 1\rangle_1$ these rotations add to a $-180°$ rotation ($\exp(i\sigma_z^{(2)}\pi/4)\ \exp(i\sigma_z^{(2)}\pi/4) = \exp(i\sigma_z^{(2)}\pi/2) = i\sigma_z^{(2)}$) on qubit 2.

**The controlled-phase (`CPHASE`) `gate`.** The transformation carried out by the controlled-phase gate ($G_{\texttt{CPHASE}}$) represents a 90° phase shift on each qubit and an overall phase

$$G_{\texttt{CPHASE}} = \sqrt{-i}Z_1(-90) \otimes Z_2(-90)U_J(\frac{1}{2J_{12}}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Here, the transformations $Z_1(-90)$ and $Z_2(-90)$ represent phase shifts of spins 1 and 2 or rotations with 90° about their respective $(-z)$-axes, or

$$Z_1(-90) = Z_2(-90) = \mathcal{R}_z(-\frac{\pi}{2}) = \begin{pmatrix} e^{-i(-\frac{\pi}{4})} & 0 \\ 0 & e^{i(-\frac{\pi}{4})} \end{pmatrix} = \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix}.$$

Now it is easy to calculate the transformation carried out by a $G_{\texttt{CPHASE}}$ as

$$G_{\texttt{CPHASE}} = \sqrt{-i}\begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix} \otimes \begin{pmatrix} e^{i\frac{\pi}{4}} & 0 \\ 0 & e^{-i\frac{\pi}{4}} \end{pmatrix} \begin{pmatrix} e^{-i\pi/4} & 0 & 0 & 0 \\ 0 & e^{+i\pi/4} & 0 & 0 \\ 0 & 0 & e^{+i\pi/4} & 0 \\ 0 & 0 & 0 & e^{-i\pi/4} \end{pmatrix}$$

$$= \sqrt{-i}\begin{pmatrix} e^{i\pi/2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{-i\pi/2} \end{pmatrix} \begin{pmatrix} e^{-i\pi/4} & 0 & 0 & 0 \\ 0 & e^{+i\pi/4} & 0 & 0 \\ 0 & 0 & e^{+i\pi/4} & 0 \\ 0 & 0 & 0 & e^{-i\pi/4} \end{pmatrix}$$

$$= \sqrt{-i} \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -i \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1-i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & 1-i \end{pmatrix}$$

$$= \frac{\sqrt{-i}}{\sqrt{2}} \begin{pmatrix} 1+i & 0 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1+i & 0 \\ 0 & 0 & 0 & -(1+i) \end{pmatrix} = \frac{\sqrt{-i}(1+i)}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

**The CNOT gate.** A $G_{\texttt{CNOT}}$ gate can be implemented as a controlled-phase transformation associated with a basis change of the target qubit and a phase shift on the control qubit [434]

$$G_{\texttt{CNOT}} = iZ_1(180)\ Y_2(-90)\ G_{\texttt{CPHASE}}\ Y_2(90)$$

where,

$$Z_1(180) = \mathcal{R}_z(180) = \begin{pmatrix} e^{-i\frac{\pi}{2}} & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = i \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Y_2(-90) = \mathcal{R}_y(-90) = \begin{pmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

$$Y_2(90) = \mathcal{R}_y(90) = \begin{pmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Then,

$$G_{\texttt{CNOT}} = i^2 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$= -\frac{1}{2} \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

A `CNOT` gate can be implemented as a `CPHASE` gate specified by a unitary matrix with diagonal elements $\{1, 1, 1, -1\}$ preceded and followed by $\mathcal{R}_y^2 \pm 90°$. The single-spin rotations

are executed by spin selective RF pulses while the two-spin rotations are implemented by allowing the system to evolve freely for a specific time interval.

Let us assume that in their initial states the two spins are oriented along their $z$-axes. First, a spin-selective RF pulse on spin 2 is applied along its $y$-axis. The frequency of the RF pulse is centered at $\omega_0^2/2\pi$ and has a spectral bandwidth that covers the frequency range $\omega_0^2/2\pi \pm J_{12}/2$ but not the $\omega_0^1 \pm J_{12}/2$ frequency range which belongs to spin 1. The pulse rotates the spin 2 from $z$- to $x$-axis.

The duration of the RF pulse must be as short as possible, so that we can safely make the assumption that meanwhile the spin 1 has not changed its orientation under the action of the spin-spin coupling. Next, the spin system is allowed to evolve freely for an interval of $1/(2J_{12})$ seconds during which they will execute 90° rotations about their respective $z$-axis. Spin 1 will remain parallel to its $z$-axis. As spin 2 is concerned, its precession frequency is shifted by $\pm J_{12}/2$ depending on whether spin 1 is in state $\mid 1\rangle$ or $\mid 0\rangle$; therefore, it will arrive in $J_{12}/2$ seconds at either $+y$- or $-y$-axis, depending on the state of spin 1. At this moment, a spin-selective pulse on spin 2 applied along its $x$-axis rotates spin 2 by 90° and brings it back along the $+z$-axis if spin 1 is in state $\mid 0\rangle$, or brings it along the $-z$-axis if spin 1 is in state $\mid 1\rangle$.

The effects of the NMR pulse sequence that implements a `CNOT` gate can be visualized in the Bloch sphere representation for logical initial states $\mid 0\rangle$ and $\mid 1\rangle$ only if the intermediate state are products of such states. If, for example, the initial state of the two spins is $1/\sqrt{2}(\mid 00\rangle + \mid 10\rangle)$, the spin 1 can be represented as pointing along the $x$-axis and spin 2 pointing along $z$-axis, but the spin-spin coupling, $(J_{12})$, leads to a superposition of states, a maximally entangled state, which can no longer be represented as a simple combination of arrows in the Bloch sphere.

If the input spin states are superpositions, then extra $z$-axis rotations have to be added to the sequence of pulses to give all elements in $G_{\texttt{CNOT}}$ the same phase. If the spin-spin interaction Hamiltonian is not of the form $\sigma_z^i \sigma_z^j$, but contains also transverse components (such as $\sigma_x^i \sigma_x^j$ and $\sigma_y^i \sigma_y^j$), the pulse sequences needed to implement the `CPHASE` and `CNOT` gates are more complicated.

When two spins are not directly coupled to each other, a `CNOT` gate can still be implemented if the two spins are connected by a network of couplings (Figure 147). If this is a nearest-neighbor coupling network, as the linear chain of qubits in Figure 147(a), we can swap the states of two neighboring qubits at a time via a sequence of `CNOT` gates until the states of the two qubits we originally wanted to involve in a `CNOT` gate are in the immediate vicinity of each other. For example, we want to execute a `CNOT` gate between the qubits 2 (control) and 4 (target) which belong to the nearest-neighbor coupling chain of six qubits. Let us denote the gate as $\texttt{CNOT}_{24}$. The qubits 2 and 4 are not directly coupled to each other, but each of them is coupled to qubit 3.

The solution is to transfer the state of qubit 2 to the nearest neighbor of qubit 4 through permutations of qubit states. The states of qubits 2 and 3 can be `SWAP`ed by performing a sequence of `CNOT` gates: $\texttt{CNOT}_{23}$ $\texttt{CNOT}_{32}$ $\texttt{CNOT}_{23}$. After that the required $\texttt{CNOT}_{34}$ is applied and the qubits 2 and 3 are swaped back (or just relabelled). The net effect is a $\texttt{CNOT}_{24}$ gate. In a linear chain of $n$ qubits with only nearest-neighbor couplings we need to perform at most $\mathcal{O}(n)$ permutations (`SWAP` operations) to perform a `CNOT` gate between any two qubits. In a network with qubits coupled via a "bus" degree of freedom (Figure 147(b)), a `CNOT` gate

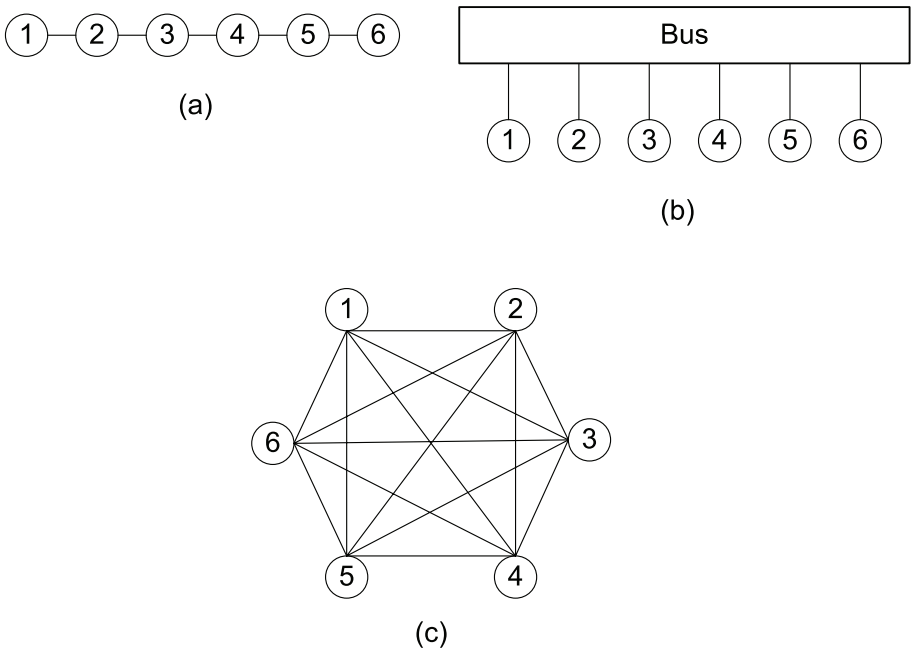



Figure 147: Possible coupling networks between six qubits. (a) A nearest-neighbor coupling network. Linear chains or two-dimensional such coupling networks are proposed for solid-state implementations. (b) A linear network coupling via a "bus" is used in ion trap implementations; since physical interactions decrease with distance the "bus" degree of freedom ensures the good coupling only of a limited number of qubits. (c) A full coupling network; the size of such networks is also limited due to physical interactions becoming weaker with distance.

between any two qubits can be executed through a series of `SWAP` operations between the intermediary qubits.

In a full-coupling network where a qubit is coupled to several other qubits (Figure 147(c)), a `CNOT` between only two of them can be performed after removing the effect of the remaining couplings. In a quantum computer there are situations, such as maintaining a state in the memory, or selectively coupling certain qubits, when we have to control single spins in order to cancel the couplings effects over a given time interval.

**Spin refocusing technique.** The $J$-coupling between nuclear spins in a molecule can not be physically turned off, but the unwanted spin couplings can be "removed" using the so-called *spin refocusing* technique. The refocusing technique was developed by NMR spectroscopists before the advent of quantum computation and it works in the following way: a 180° RF pulse is applied to one of two coupled spins in the middle of the desired time interval.

We make the following assumptions: the two spins are in a weak-coupled regime with a positive coupling constant, $J_{12}$, and, if we consider the Bloch sphere representation, the spin 1 is in one of its logical states, $|\,0\rangle$ (spin up), or $|\,1\rangle$ (spin down) and spin 2 is in the $xy$-plane. The coupling between them can be interpreted as either an increase in precession frequency of spin 2 if spin 1 is up, or a decrease if spin 1 is down. Therefore, if spin 1 points along the $z$-axis, spin 2 in its own rotating frame precesses counterclockwise (from $x$-axis to $y$-axis) around spin 1.

If spin 1 points along $-z$-axis, spin 2 precesses clockwise (from $x$-axis to $-y$-axis) around spin 1. If we wish to cancel the coupling effects over a period of 4 milliseconds, we wait for 2 milliseconds (2-ms delay) while spin 2 precesses counterclockwise/clockwise if spin 1 is up/down with the same angle. At the end of the 2 ms we apply the refocusing RF pulse which rotates spin 1 by 180° around the $x$-axis and produces an *inversion* of state for spin 1, respectively, a reversal in the direction of precession for spin 2. In the following 2 ms spin 2 precesses clockwise/counterclockwise if spin 1 is down/up. The rotation angles in the first 2 ms and the next 2 ms are equal, but in opposite directions and, as a result, spin 2 returns to its initial orientation. Another 180° RF pulse applied to spin 1 at the end of the 4-millisecond interval reverses the inversion of spin 1 and returns the two-spin system to its initial state.

The refocusing can be represented as

$$X_1(180)\ U_J(t)\ X_1(180)\ U_J(t) = U_J(-t)\ U_J(t) = I.$$

We can replace $X_1(180)$ with $Y_1(180)$ with the same result

$$Y_1(180)\ U_J(t)\ Y_1(180)\ U_J(t) = U_J(-t)\ U_J(t) = I.$$

We can apply the pulses to either one of the two qubits to cancel the coupling between them, but if we apply pulses to both qubits the coupling is not removed.

Refocusing schemes have been specifically designed for the purpose of quantum computing; the pulse sequence can be simplified after a careful study of the quantum algorithm to be implemented. In actual implementations, control techniques are always complicated by experimental artifacts and undesired Hamiltonian terms which must be addressed. Two important factors, cross-talk and undesired couplings are intrinsic to NMR. A pulse applied to a selected spin (qubit) will affect other spins (cross-talk) to some extent. The frequency selectivity of the RF pulses can be significantly improved if shaped and composite pulses are used. Unintended phase shifts $\mathcal{R}_z(\theta)$ can be calculated in advance for each spin-pulse combination and can be corrected for with the corresponding $\mathcal{R}_z(-\theta)$ inserted in the pulse sequence.

The coupling terms $J_{ij}\sigma_z^i\sigma_z^j$ act during time intervals of free evolution under the system's Hamiltonian; their effect can be removed using refocusing techniques, but only if the single spin (qubit) rotations are perfect and instantaneous. In practice, the RF pulses have a finite duration allowing the spin-spin coupling to distort the single-qubit rotations. Coupling effects can be neglected for short, high-power pulses used in heteronuclear spin systems where, typically $J_{12} < 300$ Hz and $\omega_1 \approx 50$ kHz. But coupling effects can no longer be neglected for low-power pulses used in homonuclear spin systems for which $J_{12}$ and $\omega_{\mathrm{R}F}$ are of the same order of magnitude. Moreover, two spins could become entangled when acted upon by two simultaneous RF pulses. This problem can be corrected if the RF pulses are shaped and composite.

Due to its isolation inside the atom the nucleus has extremely long coherence times, up to thousands of seconds. Modern spectrometers enable complex operations and make an NMR quantum computer experimentally attractive. In an NMR quantum computer the qubit is embodied by the spin of an atomic nucleus and the unitary operations are the result of magnetic field pulses applied to individual varieties of nuclear spins in a strong magnetic field, while the couplings between spins are enabled by the chemical bonds between neighboring atoms. The initial state is prepared by polarizing the nuclear spins in a strong magnetic field

and then by applying "effective pure state" preparation techniques. The result of a quantum computation will be read out by measuring the voltage signal induced by the precessing of the nuclear magnetic moment.

For liquid and solid NMR devices the spins are identifiable based on the internal Hamiltonian determined by the chemistry of the sample. When the molecular system is carefully engineered, the spins (qubits) can be spatially localized.

In contrast to other physical implementations, NMR is an ensemble phenomenon and only an aggregate signal from many individual molecules can be practically observed. The initial states are not pure states, but thermal statistical mixtures. The techniques presently used to prepare an effective pure state reduce the signal exponentially in the number of qubits unless the initial polarization is high enough.

## 6.10 The First Generation NMR Computer

The computational results from individual members of the ensemble are not accessible due to the averaging resulting from signals summation over the ensemble of a liquid NMR system. Nevertheless, Gershenfeld and Chuang [171] have shown in 1997 that NMR can be applied to perform quantum computation using ordinary liquids at room temperature and normal pressure and standard commercial instrumentation.

The first generation devices based on liquid NMR and commercial technology have been used to manipulate systems of up to seven qubits [432]. The simplest instance of Shor's algorithm, the factorization of $N = 15$ (prime factors 3 and 5) was implemented using a custom-synthesized "quantum computer" molecule, a perfluorobutadienyl iron complex. The qubits were five $^{19}$F and two $^{13}$C spin-1/2 nuclei. The experiments were performed in 2001 at IBM Almaden Research Center by a group of IBM scientists and Stanford University graduate students lead by Isaac Chuang.

The defining step of the algorithm implementation was the computation of the function $f(x) = a^x \mathrm{mod} N$ for $2^n$ values of $x$ in parallel. Like in the classic case, the function was computed using the identity $a^x = a^{2^{n-1}x_{n-1}} \ldots a^{2x_1} a^{x_0}$, where $x_k$ are the binary digits of $x$. Thus, the computation of $f(x)$ represents a serial multiplication by $a^{2^k} \mathrm{mod} N$ for all $k$ (where, $0 \leq k \leq n-1$) for which $\mid x_k\rangle = \mid 1\rangle$. The powers of $a^{2^k}$ can be computed on a classical machine by repeatedly squaring $a$; for $N = 15$ the values of $a$ may be $2, 4, 7, 8, 11, 13$, or 14.

(i) When $a = 2, 7, 8, \text{or } 13$, $f(x) = a^4 \mod N = 1$ with $x = 4$ and $k = 2$; then only two qubits, $x_0$ and $x_1$ control the multiplications leading to the computation of $f(x)$.

(ii) When $a = 4, 11, \text{or } 14$, $f(x) = a^2 \mod N = 1$ with $x = 2$ and $k = 1$; then only $x_0$ can be used for the computation of $f(x)$.

A register with $n = 3$ qubits (though even $n = 2$ qubits would had been enough) was used for detecting the periods of the modular exponentiation function and the quantum Fourier transform. Another register of $m = 4$ qubits was used to store $f(x)$. As reported in [432] the first three qubits were three fluorine ($^{19}$F) nuclei and the second register of four qubits was made up of two fluorine ($^{19}$F) and two carbon ($^{13}$C) nuclei. This experiment was performed at room temperature (30°) and the sample was in thermal equilibrium; the molecules and the initial state of each qubit was a statistical mixture of spin states $\mid 0\rangle$ and $\mid 1\rangle$. This mixed

state was converted into a 7−spin *effective pure state* through a variant [431] of the temporal averaging procedure [249].

An effective pure state (or pseudo-pure state as Cory *et al.* [105] call it) is assumed to have a density matrix that can be shifted by adding a multiple of the unit matrix to it so as to obtain a scalar multiple of the density matrix of a pure state. In the original scheme [249] the temporal averaging for an $n-$qubit system involved the summation of $2^n - 1$ density matrices, each of them obtained from the density matrix, $\rho_{\text{eq}}$, of the mixed state at thermal equilibrium by cyclicly permuting all populations except the ground state population.

A new version takes a weighted sum of diagonal density matrices, $\rho_i$, obtained from $\rho_{\text{eq}}$ by performing a few `CNOT` operations to rearrange the $2^n - 1$ populations. The exact density matrix of the effective pure state, $\rho_{\text{effps}}$, may require up to $2^n - 1$ experiments, but a good approximation of $\rho_{\text{effps}}$ may be obtained through much fewer experiments. The $\rho_{\text{effps}}$ obtained with this new variant of temporal averaging proved to be a good approximation of the desired initial state of the seven qubits $\mid \psi_{\text{initial}}\rangle = \mid 0000001\rangle$.

The effective pure state of the seven qubits was created in the following way: the state of the five $^{19}$F spins was made effective pure through the summation of nine experiments, where each experiment used a different sequence of two-qubit `CNOT` gates and single-qubit, spin-flip, `NOT` gates. These nine experiments were executed four times, each time with different additional `CNOT` and `NOT` gates, such that the result was an effective pure state of the two $^{13}$C spins. The summation of the 36 experiments and a `NOT` operation on the 7th spin led to the effective pure state.

The algorithm was realized with a sequence of RF pulses selecting different spins; the time intervals between the pulses allowed for free evolution under the Hamiltonian of the system. Each pulse sequence was designed to produce a certain transformation of the spin state required by a computational step in the algorithm.

For $a = 7$, the pulse sequence for the step calculating the function $f(x) = 7^x \mod 15$ was about 400 ms long, the inverse QFT was implemented in 120 ms, and the complete sequence of pulses for the Shor algorithm was about 720 ms long. After the execution of the complete sequence of pulses, the state of the first three qubits was estimated using NMR spectroscopy. The spectra showed that qubit 1 was in state $\mid 0\rangle$, while qubits 2 and 3 were in a mixture of $\mid 0\rangle$ and $\mid 1\rangle$. The 3-qubit register was in a mixed state of $\mid 000\rangle$, $\mid 010\rangle$, $\mid 100\rangle$, and $\mid 110\rangle$ which, in decimal notation, correspond to $\mid 0\rangle$, $\mid 2\rangle$, $\mid 4\rangle$, and $\mid 6\rangle$.

With the periodicity of these terms at 2, $r = 2^3/2 = 4$ and the greatest common divisor, gcd (which can be efficiently computed on a classical computer) was: $\gcd(7^{4/2} \pm 1, 15) = 3, 5$. A comparison between the measured NMR spectra and the spectra simulated based on a parameter-free decoherence model showed that the degree of unitary control in this first NMR quantum computation experiment was very high.

As the experiments have proven, NMR is a good tool for demonstration of experimental and theoretical techniques for precise control and modelling of quantum computers, but its scalability is limited by the size and stability of the custom-synthesized molecule containing an even larger number of spin-1/2 nuclei as qubits.

We continue our presentation of techniques for the implementation of quantum information processing systems with an analysis of quantum dots based on the solid-state technology which is well understood and very flexible.

## 6.11 Quantum Dots

The physical implementations of quantum computation must achieve some very demanding conditions: (i) a well-defined two-level quantum system to play the role of the qubit, (ii) precision control of the Hamiltonian operations on the qubit quantum system, and (iii) a very high degree of quantum coherence. Atomic physics implementations such as the ion trap and cavity quantum electrodynamics systems fulfil these conditions, but it is still unclear if they could be scaled up to the levels required by a full-scale quantum computer.

At this time, it is believed that solid-state systems could offer solutions for a scalable quantum computer technology. As noted by Burkard *et al.* [79], in 2000 the number of proposals for solid-state implementations of quantum computers was almost equal to the other proposals put together. Solid state physics in itself is extremely versatile: on one hand, most phenomena possible in physics can be embodied in a solid-state system; on another hand, solid-state physics allied with computer technology has enabled the creation of a wide variety of artificial structures and devices.

Quantum dots, also known as nanocrystals or "artificial atoms", play a central role in all these solid-state proposals. Quantum dots are atom-like compound semiconductors structures where charge carriers are confined in all three spatial dimensions. A semiconductor quantum dot may contain from a few thousands to tens of thousands of atoms arranged in a nearly defect-free three-dimensional crystal lattice. The carriers in the quantum dot interact strongly with lattice vibrations and could be strongly influenced by defects, surfaces, or interfaces.

The size of a quantum dot is of the order of the Fermi wavelength in the host semiconductor material, typically ranging from 2 to 10 nanometers (10-50 atoms) in diameter. Quantum dots can be lithographically defined or self-assembled. The self-assembled quantum dots have smaller sizes and stronger confinement potentials than the lithographical ones. The confinement is usually defined by electrical gating of a two-dimensional electron gas.

The quantum dots have extraordinary tunability; their characteristics are extremely sensitive to the dot's size and composition and these can be controlled by specific engineering techniques. At such small sizes semiconductors behave differently and that explains why quantum dots have unusual properties.

In bulk semiconductor material the electrons have a range of energy levels so close together that they are described as *continuous*. Some of the energy levels are forbidden to electrons and this energy region is called the *bandgap*. The overwhelming majority of the electrons of a bulk semiconductor occupy almost completely the *valence band* formed by the energy levels below the bandgap. Only a minuscule percentage of the semiconductor electrons may by found in the energy levels above the bandgap, in the so called *conduction band*.

Electrons in the valence band could be stimulated to jump in the conduction band only if given enough energy through heat, voltage, or photon flux. The minimum energy required to raise an electron in the conduction band is equal to the energy bandgap. The pair formed by the exited electron (raised in the conduction band) and the hole left in the valence band is called an *exciton*. An electron in the conduction band falls back to its valence energy level after a very short time; the radiation emitted in this de-excitation process has a frequency corresponding to the transition energy. Most transitions are from the lowest levels of the conduction band to the highest levels of the valence band and the transition energy is equal to the bandgap. Since in a bulk semiconductor the bandgap is fixed, the transitions occur at

fixed emission frequencies, the so-called *resonant frequencies*.

Quantum dots are also made of semiconductor material, but because their sizes are so small their electron energy levels can no longer be treated as continuous. The electron energy levels of a quantum dot must be treated as *discrete*, that is, with a finite, small distance between them. This discrete level structure, through its *quantum confinement*[58] effects, determines properties in quantum dots totally different from bulk semiconductor material.

The addition/removal of a few atoms to/from the quantum dot changes the discrete electron energy levels and thus, alters the boundaries of the bandgap and changes its width. Because of the small size, the bandgap width can be also modified by changing the geometry of the quantum dot surface. The bandgap in a quantum dot will always be energetically larger than the bandgap in the corresponding bulk semiconductor. Since the transition energy is larger, the emission frequency is higher and, respectively, its wavelength is shorter or "blue shifted". By controlling the bandgap of a quantum dot containing a single electron one can determine and manipulate the state of that electron.

The number of electrons in the conduction band of a quantum dot in a GaAs heterostructure can be precisely controlled one by one starting from zero[418]. This is the direct result of the so-called *Coulomb blockade*[59]. In the experiments of Tarucha *et al.* [418] a gated vertical quantum dot (diameter of about 500 nm) was made from a double-barrier heterostructure with well defined tunnel junctions and was cooled down to 50 mK. The number of electrons was varied by changing the gate voltage and the quantum dot displayed an atom-like shell structure associated with a 2-dimensional harmonic potential. More than that, in the presence of a magnetic field applied parallel to the tunnelling current, the electrons filling the shell structure showed pairing into spin-degenerate single-particle states like an atom.

Quantum dot systems are considered viable candidates for a quantum computer implementation. The qubit can be defined in several ways:

(i) as a two-level system, the ground state and the excited state of a single-electron quantum dot controlled by optical resonant pulses [19],

(ii) as the spin of the excess electron in a single-electron quantum dot [277],

(iii) as two conduction band spin states of two electrons optically excited from two ground spin states in the presence of a magnetic field applied perpendicular to the confinement direction of the quantum dot[60][217], or

(iv) as the localization of the excess electron charge in one or the other of two coupled single-electron quantum dots [417].

**A CNOT gate based on dipole-dipole interaction between two single-electron quantum dots.** Adriano Barenco, David Deutsch, and Arthur Ekert [19] were the first to propose an implementation of a quantum `CNOT` gate based on the electric dipole-dipole

[58]Quantum confinement occurs when at least one dimension of a nanocrystal approaches the size of an exciton in bulk crystal, called the exciton Bohr radius. A quantum dot has all dimensions close to the Bohr exciton radius; in compound semiconductors the calculated Bohr exciton radius varies between 2.7 nm for CdS and 13 nm for GaAs.

[59]The Coulomb blockade is a result of the discreteness of the electron charge and manifests itself as the increased resistance at small bias voltage of a low-capacitance tunnel junction. (See Glossary for more information).

[60]The optical excitation also creates a single-hole in valence band spin state which together with the two electron in the conduction band are referred to as *trion* state.

interaction between two single-electron quantum dots embedded in a semiconductor. In their simple model, the ground state and the excited state of the single electron in each quantum dot constitute the computation basis states $\mid 0\rangle$ and $\mid 1\rangle$ of a qubit; the qubit quantum state can be controlled by optical resonant pulses.

The first quantum dot, the control qubit, and the second quantum dot, the target qubit, are assumed to have different resonant frequencies $\omega_1$, respectively, $\omega_2$ obtained be controlling their bandgaps through one of the methods mentioned above. Assume that we apply adiabatically an external static electric field $E_0$ to the two qubits. In this way electron transitions between levels are avoided and we observe only a quantum-confined Stark effect in both quantum dots: the charge distribution in the ground state (state $\mid 0\rangle$) is shifted in the direction of the applied electric field while the charge distribution in the first excited state (state $\mid 1\rangle$) is shifted in the direction opposite to the applied electric field.

This change in charge distribution gives rise to dipole moments corresponding to the state the single electron is in each of the two quantum dots. For example, if the qubit state is encoded in a single electron per quantum dot we can assume that for selected coordinates, the dipole moments may be assigned the values $+d_i$ in state $\mid 0\rangle$ and $-d_i$ in state $\mid 1\rangle$, where $i = 1, 2$ refers to the control and target dots, respectively.

The electron in each quantum dot creates an electric field which may shift the energy levels in the other dot without causing transitions. This is a good approximation because the dipole-dipole interaction term, $\mathbf{D}_{12}$, is much larger than the other terms in the total Hamiltonian of the two quantum dots system

$$\mathbf{H} = \mathbf{H}_1 + \mathbf{H}_2 + \mathbf{D}_{12}.$$

The dipole-dipole interaction operator, $\mathbf{D}_{12}$, is diagonal in the four-dimensional state space spanned by the eigenstates of the free Hamiltonian of the system, $\mathbf{H}_1 + \mathbf{H}_2$, where $\mathbf{H}_1$ and $\mathbf{H}_2$ represent the Hamiltonians of the individual quantum dots.

The shift in energy levels due to the dipole-dipole interaction induced by the presence of a static electric field corresponds to a shift in the resonance frequencies of the transitions between ground level and excited level in the two quantum dots. The average shift in resonance frequencies, $\overline{\omega}$, is proportional to $d_1 d_2$, the product of the two dipole moments, and to $1/R^3$, the inverse of the cube power of the distance, $R$, between dots.

In the presence of a static electric field, the resonance frequency for transitions between the ground and excited state, respectively, between the states $\mid 0\rangle$ and $\mid 1\rangle$ of one quantum dot depends on the state of the neighboring dot. As seen in Figure 148, the resonant frequency for the first dot becomes $\omega_1 + \overline{\omega}$ or $\omega_1 - \overline{\omega}$ as the second dot is in state $\mid 1\rangle$ or $\mid 0\rangle$, respectively. At the same time, the resonant frequency of the second dot becomes $\omega_2 + \overline{\omega}$ or $\omega_2 - \overline{\omega}$ if the first dot is in state $\mid 1\rangle$ or $\mid 0\rangle$. Thus, if we apply a $\pi$ laser pulse at frequency $\omega_2 + \overline{\omega}$ the second dot will make a transition from state $\mid 0\rangle$ to state $\mid 1\rangle$ if and only if the first dot is in state $\mid 1\rangle$. At this resonant frequency the first and the second quantum dots become the control qubit and, respectively, the target qubit of a `CNOT` gate.

A system of two single-electron quantum dots could be useful for quantum information processing if the decoherence time is greater than the time scale of the optical interaction. The decoherence time for resonant frequencies in the infrared regime is estimated at about $10^{-6}$ s. Impurities and thermal vibration (phonons) can reduce it to about $10^{-9}$ s, or even lower if their effects are not minimized by fabrication or by cooling the crystal. The time

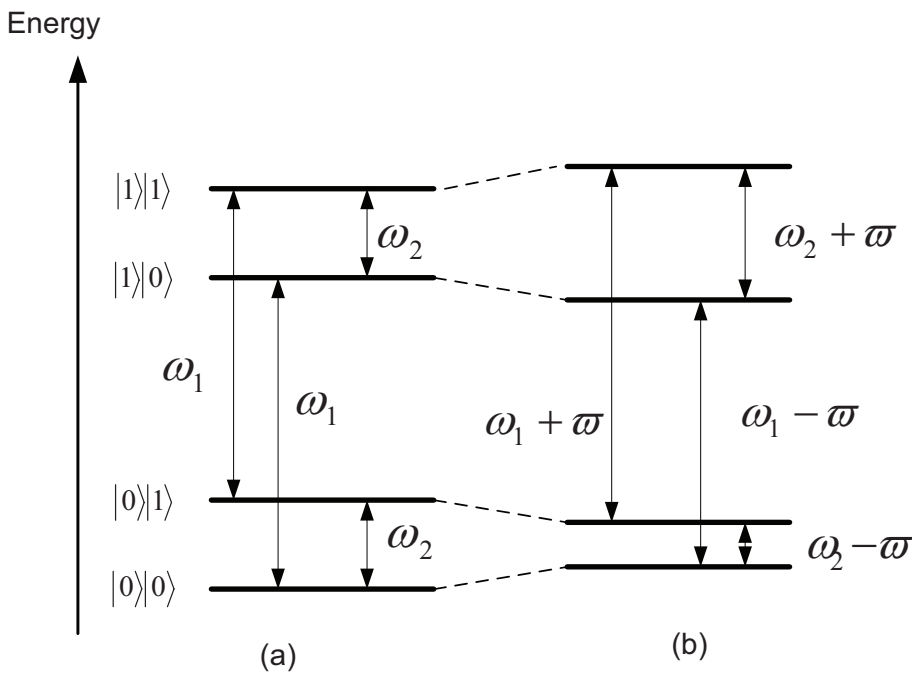



Figure 148: Energy levels of a two quantum dots system. (a) In the absence of an external static electric field the resonance frequency for the first dot is $\omega_1$ and for the second dot is $\omega_2$. (b) In the presence of a static electric field, the dipole-dipole interaction between the two quantum dots shifts their energy levels; the average shift in resonance frequencies is $\overline{\omega}$. The resonant frequency for the first dot becomes $\omega_1 + \overline{\omega}$ or $\omega_1 - \overline{\omega}$ as the second dot is in state $\mid 1\rangle$ or $\mid 0\rangle$, respectively. At the same time, the resonant frequency of the second dot becomes $\omega_2 + \overline{\omega}$ or $\omega_2 - \overline{\omega}$ if the first dot is in state $\mid 1\rangle$ or $\mid 0\rangle$.

scale of the optical interaction is defined by the length of the $\pi$ laser pulse at about $10^{-9}$ s. The $\pi$ pulse must by monochromatic and selective enough to initiate the desired transition. Thus, the length of the pulse must be greater than the inverse of the pulse carrier frequency and the inverse of the dipole-dipole interaction, $1/\omega$, which for this simple model is about $10^{-12}$ s.

**A coupled quantum dot system.** In the coupled quantum dot system proposed by Daniel Loss and David DiVincenzo [277] the qubit is realized as the spin of the excess electron in a single-electron quantum dot. The two-qubit gate operates by electrical gating of the tunnelling barrier between neighboring quantum dots, not by spectroscopic manipulation as in the previous system. For single-electron quantum dots patterned as two-dimensional electron-gas structures the gating of the tunnelling barrier between neighbors can be controlled using a split-gate technique [273].

The control is achieved in the following way: if the barrier voltage is set "high" the tunnelling is forbidden between the quantum dots and their states are preserved without evolution in time. If the barrier is pulsed to a "low" voltage the spins of the electrons in neighboring quantum dots will be subject to a transient coupling (Heisenberg coupling) of the form

$$\mathbf{H}_s(t) = J(t)\mathcal{S}_1 \cdot \mathcal{S}_2$$

where, $J(t) = 4t_0^2(t)/u$ is the time-dependent exchange constant related to the turning on and off of the tunnelling matrix element $t_0(t)$; $u$ is the charging energy of a single dot and $\mathcal{S}_1$ and $\mathcal{S}_2$ are the spin-1/2 operators for dot 1 and 2, respectively. The expression above is based on several assumptions: (i) $\Delta E \gg kT$, where $\Delta E$ is the single-particle level spacing of the dots, $k$ is Boltzmann's constant, and $T$ is the temperature, thus the single-particle states of the dots higher than the ground state can be ignored; (ii) the switching time, $\tau_s$, for pulsing the gate potential "low" is $\tau_s > k/\Delta E$ to prevent transitions to higher orbital levels; (iii) the charging energy, $u$, of a single dot is $u > t_0(t)$ for all $t$ to ensure that the Heisenberg exchange approximation is correct. The quantum computation can be executed if (iv) the decoherence time is much longer than the switching time, $\Gamma^{-1} \gg \tau_s$, where $\Gamma$ is the decoherence rate. In quantum dots the decoherence times of the spin degrees of freedom, which are insensitive to environmental fluctuations of the electrical potential, are expected to be longer than the charge degrees of freedom. Nevertheless, the spin-1/2 degrees of freedom are affected by any magnetic coupling to the environment.

For a given duration of the spin-spin coupling, so that $J_0\tau_s = \pi$, the effect of the pulsed Hamiltonian (pulsed voltage "low") is that of a unitary time evolution operator $\mathbf{U}_s(t)$ when applied to the initial state of the two spins $\psi(t) = \mathbf{U}_s\psi(0)$. For a specific duration $\int J(t)dt = J_0\tau_s = \pi(\text{mod } 2\pi)$ of the pulsed Heisenberg spin-spin coupling, this operator is a spin "swap" operator: $\mathbf{U}_s(J_0\tau_s = \pi) = \mathbf{U}_{sw}$ such that $\mathbf{U}_{sw} \mid ij\rangle = \mid ji\rangle$, where $\mid ij\rangle$ represent the basis states of the two spins in the $S^z$ basis with $i, j = 0, 1$. As such, this swap operator conserves the total angular momentum of the system and does not function as a useful two-qubit quantum gate. But, if the pulsed coupling is applied for half the given duration the resulting operator is $\mathbf{U}_{sw}^{1/2}$. A sequence of $\mathbf{U}_{sw}^{1/2}$ and single-qubit operations act like a `CNOT` gate

$$\mathbf{U}_{\text{CNOT}} = e^{i(\pi/2)S_1^z}\ e^{-i(\pi/2)S_2^z}\ \mathbf{U}_{sw}^{1/2}\ e^{i\pi S_1^z}\ \mathbf{U}_{sw}^{1/2}$$

where terms of the form $e^{\pm i\pi}S_i^z,\ i = 1, 2$ are single-qubit operations and they can be realized by applying local magnetic fields $\mathbf{B}_i$ pulsed exclusively onto spin $i$. This expression for the `CNOT` gate is given in the basis where it has the form of a conditional phase-shift; a simple basis change for qubit 2 converts it to a standard `CNOT`.

An exchange constant of the order of $J_0 = 80\ \mu$eV could be obtained with pulse durations of $\tau_s \approx 25$ ps and decoherence times of $\Gamma^{-1} = 1.4$ ns. Then, a $\pi/2$ single-qubit rotation could be realized over a time interval of the order of a pulse duration $\tau_s$ of 25 ps by applying a magnetic field $B \approx 0.6$ T, a realistic value for solid state.

The two implementations discussed so far are based on electron spin coupling; now we discuss implantations based on electron charge.

**Quantum dot qubit based on charge.** The quantum dot implementation of qubits based only on the electrical effects [417] uses a field-effect-transistor structure for qubit state detection purposes. This implementation is closely related to conventional large scale integrated (Si) circuits and, in this sense, easier to manufacture and operate, but the decoherence phenomena are considered extremely difficult to overcome.

The basic idea is to insert an excess electron into two coupled quantum dots and to consider the localization of this excess electron in one or the other of the quantum dots as the qubit

states $|\,0\rangle$ and $|\,1\rangle$, respectively. A superposition state can be realized by varying the voltage on the gate electrodes attached to the quantum dots via oxide barriers.

The quantum dots are small in size (for example, diameter $d \leq 10$ nm in silicon) and their electron energy levels are discrete. When the discrete energy level in one quantum dot matches that of the other quantum dot in a coupled quantum dot system, the corresponding electron can transfer from one dot to the other one inside the system. The electrons in two quantum dots of the coupled system are separated by an energy barrier which can be easily tunnelled through when electron levels in the two dots match in energy. If this barrier is low and the coupling between the dots is strong, the two coupled quantum dots behave as a large single quantum dot where, possibly, only one electron resides. This large single quantum dot is itself characterized by a Coulomb blockade phenomenon like any regular single quantum dot. The electron tunnelling in and out of a quantum dot is dominated by this classical effect related to the discrete nature of the electron charge. When the potential barrier delimiting the quantum dot is relatively high, the tunnelling to and from the dot is weak, and the number of electrons on the dot is a well defined integer, $N$. A current flowing as a result of a sequence of single electrons tunnelling through the dot will change the electron number by one at a time.

Due to the Coulomb repulsion between the electrons in the dot, the addition of one electron is associated with an energy cost. The extra energy required is $\sim e^2/C$, where $e$ is the electron charge and $C$ is the capacitance of the dot. No new current will flow until the extra energy is provided by increasing the source-drain voltage applied on the dot. This characteristic charging energy must be larger than the thermal energy of electrons for the Coulomb blockade to be observed. For a quantum dot of 1 femtofarad ($10^{-15}$F) capacitance this condition is satisfied at temperatures below 1 Kelvin which can be realized in a so-called dilution refrigerator[61].

When an excess electron is injected into a system of two coupled quantum dots of different size, the system can be treated as a two-state system; the energy levels of the coupled system as a whole show the localized state of the wave function reflecting the different energy levels of the two independent isolated quantum dots. At a certain value of the applied gate bias voltage, the *resonant gate voltage*, the energy levels of the excess electron in each of the original quantum dots coincide and the electron transfers to the other dot through a process called *resonant tunnelling*. In this way the coupling between the dots removes the degeneracy of the electron energy levels in the single dot quantum states; the coupling creates new states associated with delocalization of the electron such that its even-parity and odd-parity wave functions spread over the two coupled dots.

The perfect localization of the electron in one or the other of the coupled dots represent the qubit basis states ($|\,0\rangle, |\,1\rangle$). By adjusting the gate bias, the electron state can be controlled and changed from a well localized one ($|\,0\rangle$ or $|\,1\rangle$) to an intermediate delocalized state where the electron transfers from one dot to the other within the coupled-dot system in a very short time (of the order of 1 picosecond). One dot is set close to the channel where from the excess electron is inserted into the coupled-dot system and the other dot is set near the gate electrode; the gate voltage controls the discrete energy-levels of the coupled dots. The presence of an electron inside the dot on the channel side represents qubit state $|\,1\rangle$ and the

[61] The operation of the dilution refrigerator is based on the phase transition a liquid mixture of two helium isotopes, $^3$He and $^4$He, undergoes at temperatures below 0.86 K, the triple critical point. (For more information see the Glossary).

presence of the electron inside the dot on the gate electrode side represents the state $|0\rangle$.

The size of the two coupled dots is determined by the semiconductor material in which they are imbedded and by two requirements:

- The electron once inserted into the system should be prevented to return to the channel by applying a threshold gate voltage; and
- The discrete energy levels of the two dots should coincide at an applied gate voltage low enough to maintain the coherence of the coupled-dot state.

For coupled dots imbedded into the conventional silicon metal oxide semiconductor field effect transistor these requirements translate into an asymmetry of the coupled dots with the larger dot close to the channel and the smaller dot near the gate electrode [417].

When two such coupled-dot qubits are positioned side by side, the electric field created by the charge distribution of the electron in a qubit changes the potential well profiles and the energy levels of the neighboring qubits. The electron state in a qubit is affected by the states of the electrons in the neighbor qubit, whether this is a state $|0\rangle$, a state $|1\rangle$, or an arbitrary superposition of $|0\rangle$ and $|1\rangle$. The electric field generated by one coupled-dot qubit slightly changes the relative positions of the electron energy levels of the other coupled-dot qubit; it moves them off resonance and in this way makes impossible the transfer of the localized excess electron between the dots of this coupled-dot qubit. Thus, by changing the charge distribution in a selective way in one coupled-dot qubit, we can control the total charge distribution of the electron, respectively the state of the electron, in the other qubit.

A `CNOT` quantum gate can be easily implemented by positioning two coupled-dot qubits close to each other and connecting them to the same channel while keeping their gate electrodes separated. It is assumed that the electron transfer between two neighboring coupled-dot qubits can be neglected. Assume that the electron in the target qubit is localized in its larger dot at a lower gate bias while the control qubit is in state $|1\rangle$ (the electron charge is localized in the larger dot placed near the channel). As the gate voltage applied to the target qubit is raised, the energy level of the electron localization in the larger dot exceeds that of the smaller dot. When the gate voltage reaches the resonance value, the wave function of the lowest energy state and that of the excited state spread over the two dots (of the target qubit) with equal weight and the electron transfers from the larger to the smaller dot, changing the state of the target qubit.

The energy levels of the single dots are split by the coupling of the dots and become separated by a small energy difference proportional with a frequency $\nu$ whose value is determined by the frequencies associated with the energy levels in the individual dots and by the coupling between the dots. This energy difference depends on the size of the dots and the semiconductor system (e.g., Si in $SiO_2$ matrix or GaAs in AlGaAs matrix) they are made of. When the control qubit is in state $|1\rangle$ the resonance gate bias of the target qubit is shifted with this small energy difference towards lower bias region, $V_{\text{res}}^{|1\rangle}$; the target qubit moves between the dots, it changes state. However, when the control qubit is in state $|0\rangle$ (the electron charge localized in the smaller dot of the control qubit) the energy level in the smaller qubit of the target qubit is raised and the resonance gate bias shifts towards a higher bias region, $V_{\text{res}}^{|0\rangle}$, where,

$$V_{\text{res}}^{|0\rangle} > V_{\text{res}}^{|1\rangle}.$$

Thus, if we apply a gate bias $V_{\text{res}}^{|1\rangle}$ to the target qubit when the control qubit is in state $|\,0\rangle$, the electron in the target qubit is prohibited to move between the dots. The two coupled-dot qubits operate like a `CNOT` gate. It can be shown that the electron moves completely between the coupled dots of a pair with a period $\pi/(2\nu)$. The charge transfer is realized when we apply a bias voltage at which the ground state and excited levels of the coupled dots come in the closest vicinity in the time period $\pi/(2\nu)$.

## 6.12 Fabrication of Quantum Dots

Quantum dot structures which contain carriers or excitons confined in all three dimensions to a nanometer-sized region of a semiconductor are possible to fabricate due to progresses in lithography, colloidal chemistry, and epitaxial growth. Quantum dots are made of different materials in different sizes: self-assembled quantum dots, semiconductor vertical or lateral dots, single molecules trapped between electrodes, metallic or superconducting nanoparticles, semiconducting nanowires, even carbon nanotubes between closely spaced electrodes.

The realization of an appropriate (two-level system) quantum dot circuit containing just a single conduction electron where electrical voltages are used to control one-electron quantum states is a real challenge. In reducing the number of electrons, the gates used to deplete the dots influence (decrease) the tunnel barriers; the current peaks associated with the Coulomb blockade become unmeasurably small, but not necessarily due to empty dots.

Few-electron quantum dots have been realized in self-assembled structures[329], in small semiconductor vertical pillars defined by etching[256], or as lateral quantum dots defined in a *two-dimensional electron gas* (2DEG) by surface metal gates on top of a semiconductor heterostructure. The first two types of quantum dots are hard to integrate into circuits with a controllable coupling between the elements. Multiple 2DEG quantum dots can be integrated simply by increasing the number of gate electrodes and the coupling between them can be controlled by setting the gate voltages.

**Vertical dots.** The vertical dots can be lithographically defined; the technology is similar to silicon-based very large system integration Tiwari *et al.* used for manufacturing the single-electron memory device[421]. The quantum dot is inside the pillar and has the shape of a two-dimensional disc. The pillar is a three-terminal field-effect transistor where the current can be switched on and off by changing the gate voltage. In fact there is a quasi-periodic set of voltages where the current switches. Since the switch can be triggered by a small fraction of a single electron charge these devices are called *single-electron transistors*.

The vertical quantum dots single-electron transistors proposed by Kouwenhoven *et al.* [256] are miniaturized resonant tunnelling diodes. The pillars are etched from a semiconductor double-barrier heterostructure and a metal gate electrode deposited around the pillar in the region of the quantum dot. In lithographically patterned quantum dots the confinement along the growth (vertical) direction is provided by the quantum potential well of the material, but the lateral confinement has to be provided by an electrostatically induced potential barrier. The gate potential together with the surface potential confine the electrons in the lateral $x$-

and $y$-directions while the double-barrier heterostructure defines the electron confinement in the $z$-direction. The confinement potential is repulsive, can be approximated by a harmonic potential and has a $r^2$-dependence, different from the $r^{-1}$-dependence of the attractive atomic potential.

The quantum dot, 12$\mu$m thick, a few hundred nanometers in diameter and containing up to 100 electrons, is created in the central well which is made of undoped $In_{0.05}Ga_{0.95}As$. The tunnel barriers bordering the dot are made of undoped $Al_{0.22}Ga_{0.78}As$ with thicknesses of 9.0 nm (upper one) and 7.5 nm (lower one). The conducting source and drain contacts are made of Si-doped n-GaAs with a variable concentration of Si dopant (increasing away from the tunnel barriers). The variation in the heterostructure material and in doping determines a particular conduction band profile for the quantum dot. For example, in this particular case, the introduction of 5% indium (In) atoms in the well lowers the bottom of its conduction band 32 meV bellow the Fermi level of the contacts [62]. The confinement potential in the $z$-direction (vertical) is such that the lowest quantum state and the second quantum state in the well are, respectively, 6 meV below and 63 meV above the Fermi level of the contacts and the quantum dot is a good approximation of a two-state quantum system. Due to the presence of In, this quantum dot system contains electrons without applying voltages.

**Self-assembled quantum dots.** Such structures form spontaneously during the epitaxial growth process and are characterized by smaller sizes. Their confinement is stronger and is provided in all three dimensions by a high bandgap in the surrounding material. The fabrication of self-assembled quantum dots is based on in-situ growth techniques such as molecular beam epitaxy or metalorganic chemical-vapor deposition. For example, a beam of indium (In) atoms (gallium or arsenic atoms can also be used) is sent onto an ultra clean gallium arsenide (GaAs) surface heated at high temperature [329]. The atoms diffuse across the clean surface and start arranging themselves to form a continuous epitaxial layer of $In_xGa_{1-x}As$.

The atomic arrangement on the surface develops so as to minimize the sum of surface, interfacial, and elastic energies of the epitaxial film. The elastic strain energy of the epitaxial film grows quadratically with the film thickness rapidly increasing the total energy of the film. If the epitaxial film has a lattice parameter even a few percent different from that of the GaAs substrate, nanometer-size islands can form on the surface of the film to minimize its total energy. Since the film relaxation is elastic, no defects are introduced during the island formation process. As the film strain due to lattice misfit between GaAs and $In_xGa_{1-x}As$ increases from 0 to 7%, the In concentration $x$ increases from 0 to 1. Islands of InAs shaped as truncated pyramids nucleate randomly on the clean surface. They are converted into quantum dots by covering them with a GaAs epitaxial layer. The process can be repeated to create a three-dimensional array of quantum dots.

In general, the confining potential can not be accurately estimated since the exact shape and composition of the islands are not well known and depend on the growth technique, the band structure of the strained material in the epitaxial film becomes more complex, and piezoelectric forces are present. Therefore, an *effective confining potential* inside the quantum dots is often defined. This effective quantity is estimated from a combination of spectroscopic

[62] The Fermi level is the last energy level occupied by electrons at the zero absolute temperature, $T = 0$ K. According to the Pauli exclusion principle, at $T = 0$ K the electrons of an atom fill all available levels starting with the lowest one. (For more information see the Glossary).

studies of the ground state and the excitation energies in quantum dots.

Capacitance spectroscopy and transmission spectroscopy in far-infrared can be performed on quantum dots fabricated in a special field-effect structure based on metal-insulator-semiconductor field-effect transistor (MISFET) design to study the ground (charging) state and, respectively, the excitations of the $n$-electron state in the dot. The typical *capacitance* vs *voltage* characteristic of a quantum dot shows six maxima where each maximum corresponds to the insertion of one electron at a time, from one up to six, into the quantum dot.

The first two maxima correspond to the first two electrons occupying the lowest, spin-degenerate state, also called the $s$ shell. After the insertion of the first electron, the second electron is harder to insert into the quantum dot due to its Coulomb interaction with the first electron (Coulomb blockade) which raises the energy of the first level. This raise in energy is compensated by a raise in the gate voltage. The addition of the third electron to the dot requires an energy equal to the quantum energy gap between the first and the second energy level of the dot plus the energy displacement of the first level.

The four electrons which occupy the second shell, the $p$ shell correspond to the second group of four maxima. These four maxima are equally spaced as a consequence of the shape of the quantum dot confining potential which is symmetrical in the $xy$ plane and has a stronger component in the $z$ direction, the growth direction.

As long as only two electrons ($n = 2$) are inserted into the quantum dot, the far-infrared spectroscopy data agree well with a parabolic model of the confining potential and its effective curvature can be precisely determined. The insertion of a third electron (a lithium atom-like quantum dot) changes the character of the electron energy spectrum and the data can no longer be explained or predicted with the simple parabolic model. Direct measurements must be performed on specific quantum dot systems.

A fully tunable qubit quantum system with a combination of both a single-electron charge degree of freedom, which is controlled with electrical voltages, and the spin degree of freedom, which ensures a long coherence time, can be realized with a quantum dot integrated with a quantum point contact, QPC, as proposed by Elzerman *et al.* [143]. The quantum dot is formed in the two-dimensional electron gas (2DEG) of a GaAs/AlGaAs heterostructure by applying negative voltages to the metal surface gates M, R, and T. The 2DEG is depleted below the gates and a potential minimum, the quantum dot itself, is created in the middle.

The gate voltages can be tuned such that the dot contains either zero or one electron; the insertion of one electron is controlled by the voltage applied to the plunger gate P. The gates T and R create a tunnel barrier between the dot and the drain; this barrier can be made opaque to completely isolate the dot from the drain. The gates T and M set the barrier between the dot and the reservoir at a certain tunnel rate (for example, $\Gamma \approx (0.05\text{ms})^{-1}$).

The QPC defined by the gates R and Q is set to operate in the tunnelling regime where the current flowing through the QPC is very sensitive to electrostatic changes in the vicinity. An electron tunnelling into or off the dot changes the electrostatic potential around the dot and produces a change in the QPC current. The QPC can be used as a charge detector with a resolution much better than a single electron charge and a measurement timescale of about $8\mu$s, almost ten times shorter than $1/\Gamma$; it permits to determine whether an electron is present in the dot or not.

The procedure for measuring the spin state when an electron of unknown spin orientation has been detected inside the quantum dot requires placing the quantum dot device inside a

dilution refrigerator and applying a magnetic field of 10 T in the plane of the 2D electron gas [256]; in this case, the Zeeman splitting of the electron levels in the dot was $\Delta E_Z \approx 200$ $\mu$eV, larger than the thermal energy (25 $\mu$eV), but smaller than the orbital energy level spacing (1.1 meV) and the charging energy (2.5 meV). The procedure is controlled with voltage pulses on the plunger gate P. When the dot is empty, both spin-$\uparrow$ and spin-$\downarrow$ electron energy levels, respectively, $E_\uparrow$ and $E_\downarrow$, where $E_\downarrow > E_\uparrow$, are above the Fermi energy of the reservoir, $E_F$. A voltage pulse (about 10 mV) applied on P moves both levels below $E_F$ and tunnelling into the dot becomes energetically favorable for an electron of either spin $\uparrow$ or spin $\downarrow$.

The tunnelling can happen after a typical time interval of $\sim \Gamma^{-1}$. For as long as the pulse amplitude is maintained at 10 mV the electron stays trapped inside the quantum dot and the Coulomb blockade prevents a second electron of entering the dot. After a time interval $t_{\text{wait}}$ the pulse amplitude is lowered to about 5 mV and the electron levels are repositioned with respect to $E_F$ according to their spin.

If the electron spin is $\uparrow$, its energy level is below $E_F$ and the electron stays trapped inside the quantum dot. If the electron spin is $\downarrow$, its energy level is above $E_F$ and the electron tunnels out of the dot after a typical time of $\sim \Gamma_\downarrow^{-1}$; the Coulomb blockade is cancelled and an electron with spin $\uparrow$ can tunnel into the dot, usually after a time $\sim \Gamma_\uparrow^{-1}$, where $\Gamma_\uparrow + \Gamma_\downarrow = \Gamma$. After a time $t_{\text{read}}$ the pulse ends and the dot is emptied.

During $t_{\text{read}}$ the current at the QPC tracks the charge on the quantum dot: it goes up when an electron tunnels off the dot and goes down by the same amount when an electron tunnels into the dot. Therefore, the QPC current goes up and then down if an spin-$\downarrow$ electron is inside the dot at the beginning of $t_{\text{read}}$. If a spin-$\uparrow$ electron is inside the dot, the current at QPC should be flat. The spin measurement fidelity can be significantly improved by lowering the temperature of the electrons and by making the charge measurement faster. The single spin energy relation time was estimated at about 1 ms for such quantum dots in a 8 T magnetic field.

Two such lateral quantum dots built next to each other and integrated with two QPCs as charge detectors can serve as a fully tunable two-qubit quantum system[142]. The QPCs provide information about the charge on the dot when an electron is pushed out of it by a more negative electrode voltage, even when no current flows through the dot. The QPCs are sensitive enough to detect inter-dot transitions. The charge read-out can affect the spin state indirectly via the spin-orbit interaction. This effect could be much reduced until spin-to-charge conversion is initiated if the charge detectors are switched only during the read-out stage of a cycle.

## 6.13 Quantum Dot Electron Spins and Cavity QED

Quantum computing implementation based on semiconductor quantum dot electron spins offers two significant advantages:

(1) the quantum dot arrays could be scalable to $\geq 100$ coupled qubits, and

(2) the spin decoherence times for conduction band electrons in III-IV and II-VI semiconductors are "very" long as demonstrated by recent experiments. For example, in moderately $n$-doped GaAs spin conduction band electron spin decoherence times of $\sim 130$ ns were measured at temperature $T = 5$ K and magnetic field $B = 0$ T [237].

Microcavities are optical resonators (cavities) with dimensions in the micrometer or sub-micrometer range in all directions; while the cavities are very large compared to the optical wavelength, the microcavities are of the order of a wavelength. The coupling of the quantum dot electron spins through a cavity-QED mode could bring another advantage, that of long-distance, fast interactions between qubits, as first proposed by Pellizzari *et al.* [323].

In the scheme of Sherwin *et al.* [379] the qubits consist of the lowest electronic states of specially engineered quantum dots and are coupled by terahertz (THz) ($10^{12}$ Hz) cavity photons. Each quantum dot contains exactly one electron. The lowest two energy levels form the states $\mid 0\rangle$ and $\mid 1\rangle$ of a qubit and the third energy level, denoted $\mid 2\rangle$ serves as an auxiliary state to perform conditional rotations of the state vector of the qubit. The spacing between and the absolute values of these electron energy levels of the semiconductor quantum dot are controlled with the voltages applied to the electrical gates of the dot via the Stark effect[63]. A large number of such quantum dots is enclosed in a three-dimensional microcavity with fundamental resonance of wavelength $\lambda_C$ much longer than the diameter of a quantum dot. The electron in each quantum dot is acted upon with a continuous-wave laser of fixed wavelength $\lambda_L$, where $\lambda_L \neq \lambda_C$.

The THz photons in the cavity act as a data bus and can couple an arbitrary pair of quantum dots. The state of an electron in a quantum dot can be coherently manipulated by tuning the transition energies $E_{10}$ and $E_{20}$ between the dot states $\mid 0\rangle \rightarrow \mid 1\rangle$ and, respectively, $\mid 0\rangle \rightarrow \mid 2\rangle$. The transition energies are tuned by varying the electric field across the quantum dot, that is, by varying the electrode voltages. The transition $E_{10}$ is tuned in and out of resonance with $\hbar\omega_C$ and $\hbar\omega_L$ at electric fields $\mathcal{E}_C$ and $\mathcal{E}_L$, respectively. The transition $E_{20}$ is tuned in and out of resonance with the two-photon transition with energy $\hbar\omega_L + \hbar\omega_C$ at the electric field $\mathcal{E}_{L+C}$.

**A CNOT gate.** The operation of a `CNOT` gate using quantum dot spins and cavity - QED is based on a series of voltage pulses applied across the electric gates of a pair of qubits. The pulses begin at a baseline value of the electric field $\mathcal{E} = 0$ and then rise to one of the target values $\mathcal{E}_C$, $\mathcal{E}_L$, or $\mathcal{E}_{L+C}$. The cavity always begins with no photons.

In the first step, a $\pi$ pulse with amplitude $\mathcal{E}_C$ and duration $\pi/(2g_{01})$ is applied to the control qubit, where the factor $g_{01}$ represents the vacuum Rabi frequency for the transition $\mid 0\rangle \rightarrow \mid 1\rangle$[64]. If the control qubit is in state $\mid 0\rangle$, the pulse does not affect it, but if it is in

[63] In a semiconductor heterostructure like that used to manufacture the quantum dots, the effect of splitting and shifting of an atomic spectral line into several components in the presence of an electric field can be significantly enhanced by bound excitons.

[64] The vacuum Rabi frequency for the transition $\mid 0\rangle \rightarrow \mid 1\rangle$ is

$$g_{01} = q z_{01} \mathcal{E}_{\text{vac}}$$

with $q$ the electron charge, $z_{01}$ the dipole matrix element of the $\mid 0\rangle \rightarrow \mid 1\rangle$ transition, and $\mathcal{E}_{\text{vac}}$ the amplitude of the vacuum electric field in the cavity. This amplitude is given by

$$\mathcal{E}_{\text{vac}} = \left(\frac{\hbar\omega_C}{2\varepsilon_0 \varepsilon V}\right)^{1/2}$$

where, $\varepsilon$ and $V$ are the dielectric constant and the volume of the cavity and $\varepsilon_0$ is the dielectric constant of the vacuum.

state $\mid 1\rangle$, the pulse rotates it into state $\mid 0\rangle$ with phase $i$ and a single photon is transferred to the cavity.

In the second step, a $2\pi$ pulse of amplitude $\mathcal{E}_{L+C}$ and duration $\pi/\tilde{\Omega}(\mathcal{E}_{L+C})$ is applied to the target qubit; the two-photon effective Rabi frequency $\tilde{\Omega}(\mathcal{E}_{L+C})$ refers to a Rabi oscillation between states $\mid 0\rangle$ and $\mid 2\rangle$ involving both cavity and laser photons at electric field $\mathcal{E}_{L+C}$. The target qubit in state $\mid 1\rangle$, it is not affected, but if the target qubit is in state $\mid 0\rangle$ it acquires a phase $-1$ if the cavity contains one photon.

In the third step, a $\pi$ pulse of amplitude $\mathcal{E}_C$, identical to the one in the first step, is applied to the control qubit. If a photon is present in the cavity, the control qubit absorbs it, returns to the state $\mid 1\rangle$ and acquires another phase $i$. The end result after these three steps is a conditional two-qubit gate in which the state vector of the two-qubit system acquires a phase $-1$ if and only if both the control and the target qubit are initially in state $\mid 1\rangle$. A true `CNOT` gate is realized by applying to the target qubit $\pi/2$ and $3\pi/2$ pulses with amplitude $\mathcal{E}_L$ and durations $\pi/(4\Omega_{L,01})$ and $3\pi/(4\Omega_{L,01})$[65], respectively, before and after the sequence of $\pi$ and $2\pi$ pulses previously discussed.

The fidelity of such a `CNOT` gate can be ensured only if:

(i) all the state vector rotations occur while the electric field is at its required value $\mathcal{E}$; therefore the rise and fall times $\delta t$ of the electric pulses must be short compared to the period of the Rabi oscillation at the respective field $\mathcal{E}$;

(ii) the Hamiltonian changes are adiabatic in comparison with the rise and fall times of the electric pulses ($\hbar/\mathcal{E}_{10} \ll \delta t$), so that the probability of transition $\mid 0\rangle \rightarrow \mid 1\rangle$ induced by the ramping of the electric field should be minimized;

(iii) the timing between the pulses must be adjusted to compensate for the quantum mechanical phases, $\exp(-i\mathcal{E}_{10}t/\hbar)$, accumulated by inactive qubits in their excited states;

(iv) the height and duration of the electric pulses are adjusted to account for the ac Stark shifts[66] in the energy levels of the quantum dots which are induced by the laser field.

If the dominant mechanism for decoherence is due to acoustic-phonon emission, calculations made by Sherwin *et al.* [379] for a specific GaAs/al$_x$Ga$_{1-x}$ quantum dot, a lossless dielectric cavity, a laser electric field of about 30 kV/m, and $2\pi$ and $\pi$ electric voltage pulses of 25 ns and 3ns, respectively, suggest that several thousand `CNOT` operations can be performed before the quantum computer decoheres. Here, the $2\pi$ pulse applied to the target qubit which is interacting with both a laser and a cavity photon is the longest operation; the $\pi$ pulse applied to the control qubit is only 3 ns.

**Self-assembled InAs quantum dots.** Imamoglu *et al.* [217] have proposed self-assembled InAs quantum dots embedded in microdisk structures with a high cavity quality factor. The quantum dots have a strong confinement along the $z$-direction which ensure that the lowest energy eigenstates consist of $\mid m_z = \pm 1/2\rangle$ conduction band states and $\mid m_z = \pm 3/2\rangle$ valence band states; the in-plane confinement is also strong enough to ensure that the electron will always be in the ground state orbital.

[65]The Rabi frequency $\Omega_{L,01}$ refers to a Rabi oscillation between states $\mid 0\rangle$ and $\mid 1\rangle$ involving one laser photon at electric field $\mathcal{E}_L$.

[66]Shift of atomic levels caused by a time-varying (ac) electric field of the order of $10^{10}$ V/cm, typically, that of a laser. In a two-level atom, the excited state shift is of equal magnitude but opposite sign of ground state shift.

The quantum dots are doped such that for each of them the valence band is completely filled with electrons and the conduction band contains a single electron. A magnetic field is applied along the $x$ direction, $B_x$ and the quantum dot qubit is defined by the two spin states of the electron in the conduction band

$$| m_x = -1/2\rangle = |\downarrow\rangle \quad \text{and} \quad | m_x = 1/2\rangle = |\uparrow\rangle$$

with energies $\hbar\omega_\downarrow$ and $\hbar\omega_\uparrow$, respectively. In this scheme each quantum dot is individually addressed by lasers and the two spin eigenstates are Raman coupled through strong laser fields and a single microcavity mode.

An arbitrary single-qubit rotation can be initiated by applying two laser fields $\mathcal{E}_{L,x}(t)$ and $\mathcal{E}_{L,y}(t)$ polarized along the $x$ and $y$ directions, respectively. The laser frequencies $\hbar\omega_{l,x}$, $\hbar\omega_{L,y}$ and the frequencies of their Rabi oscillations $\Omega_{L,x}$ and $\Omega_{L,Y}$ satisfy exactly the Raman-resonance condition between the $|\downarrow\rangle$ and $|\uparrow\rangle$ states. The $\pi/r$ ($r$ any real number) laser field pulses have a short duration and the laser polarizations must have nonparallel components in order to create a nonzero Raman coupling. Such arbitrary single-qubit rotations can be carried out in parallel on several quantum dots inside the microcavity.

A two-qubit operation can be implemented by a selective coupling between the control and the target qubit that is mediated by the microcavity field. A combination of a cavity mode with energy $\hbar\omega_{\text{cav}}$ and $x$-direction polarization and a laser field with $y$-direction polarization could induce a Raman transition between the two conduction band states. As a first step in the implementation of a `CNOT` operation, two laser fields of frequencies $\omega_L^i$ and $\omega_L^j$ are shined onto two quantum dots to establish a near two-photon resonance condition for both the control $i$ and target $j$ qubits, respectively.

The *exact* two-photon resonance condition for the control quantum dot, $i$, is

$$\Delta\omega_\uparrow^i = \Delta\omega_\downarrow^i$$

with

$$\Delta\omega_\uparrow^i = \omega_\uparrow^i - \omega_v^i - \omega_{\text{cav}}^i$$

and

$$\Delta\omega_\downarrow^i = \omega_\downarrow^i - \omega_v^i - \omega_L^i$$

where $\omega_v^i$ defines the energy, $\hbar\omega_v^i$, of the relevant valence band states for the control qubit.

The *exact* two-photon resonance condition for the target quantum dot, $j$, is given by a similar set of equations

$$\Delta\omega_\uparrow^j = \Delta\omega_\downarrow^j$$

with

$$\Delta\omega_\uparrow^j = \omega_\uparrow^j - \omega_v^j - \omega_{\text{cav}}^j \quad \text{and} \quad \Delta\omega_\downarrow^j = \omega_\downarrow^j - \omega_v^j - \omega_L^j.$$

The energy of the relevant valence band states for the target qubit, $\hbar\omega_v^j$, corresponds to $\omega_v^j$.

A *near* two-photon resonance condition for both the control qubit ($i$) and the target qubit ($j$) can be realized if the two-photon detuning between the two qubits

$$\Delta_{ij} = \Delta_i - \Delta_j = 0.$$

The detunings for the individual qubits are defined as

$$\Delta_i = \omega^i_{\uparrow\downarrow} - \omega_{\text{cav}} + \omega^i_L \quad \text{and} \quad \Delta_j = \omega^j_{\uparrow\downarrow} - \omega_{\text{cav}} + \omega^j_L$$

with

$$\omega^i_{\uparrow\downarrow} = \omega^i_{\uparrow} - \omega^i_{\downarrow} \quad \text{and} \quad \omega^j_{\uparrow\downarrow} = \omega^j_{\uparrow} - \omega^j_{\downarrow}.$$

The detunings for each qubit must satisfy the following conditions

$$\Delta_i \ll \omega^i_{\uparrow\downarrow} \quad \text{and} \quad \Delta_j \ll \omega^j_{\uparrow\downarrow}.$$

A fully controllable long range transverse spin-spin interaction between electrons in two distant quantum dots can be established through Raman coupling via a common cavity mode if the laser field frequencies $\omega^i_L$ and $\omega^j_L$ are chosen such that the condition $\Delta_{ij} = 0$ is fulfilled.

A two-qubit gate such as the *conditional phase-flip operation* between two spins $i$ and $j$ can be implemented by combining a transverse spin-spin coupling Hamiltonian with one-qubit rotations.

A critical issue for the quantum dots made with semiconductors in group III-V, such as GaAs and AlAs, is the presence of nuclear spins in the semiconductor substrate; they create an inhomogeneous magnetic field which reduces the decoherence time and produce decoherence through dynamic spin-diffusion induced by nuclear dipole-dipole interactions. The group IV semiconductors, such as Si and Ge are free of nuclear spin effects. If the silicon semiconductor is used, the quantum dot can be replaced with a single impurity, such as a single phosphorous atom [229, 230] which binds a donor electron at room temperature. The quantum information can be encoded in the donor electron, or in the state of the single $^{31}$P nuclear spin. The electrostatically defined quantum dots and silicon-based impurities have exchange interactions at extremely short range. Photonic connections between quantum dots assisted by cavity quantum electrodynamics with optical microcavities can help solve this problem. Progress in optical control and measurement of qubits in self-assembled quantum dots potentially enable very fast quantum computation. Though large wafers of spins trapped in quantum dots or impurities are routinely fabricated, the scaling of a system of coupled spins remains a difficult task.

## 6.14 Quantum Hall Effect

Topologically defined quantum gates represent a crucial development in quantum computation; they are considered to define not only very powerful fault-tolerant quantum error correcting schemes among qubits, but also a very special method of quantum computation for which the appropriate physical implementation remains to be determined.

The anyons, a type of quantum excitation with fractional, non-Abelian statistics are considered as playing a part in topological condensed matter system. Quantum logic implemented as braiding of trajectories in a (2D + Time)-space of such quasiparticle is expected to offer a way to robust quantum computation. Theoretical calculations and some the experimental indicate that some fractional quantum Hall effect states with filling factors such as $\nu = 5/2$

and $\nu = 12/5$ could be states of non-Abelian anyons. These incompressible fractional quantum Hall effect states are the only experimentally established topological phases. Universal topological quantum computation requires both a topological phase and quasiparticles with non-Abelian braiding statistics and with topological properties in a special class.

We will review first some concepts related to the quantum Hall effect, then the fractional quantum Hall effect and some of its theoretical models. The quasiparticles introduced by these models to explain the fractional filling factors match the properties of the anyons introduced by Frank Wilczek in 1982.

**Quantum Hall effect.** The Hall effect was discovered by the American physicist Edwin Hall in 1879 while working on his doctoral thesis; the results were published in 1880 [196, 197]. This discovery was made almost eighteen years *before* the electron was identified as an electric charged particle.

The classical effect manifests itself as a potential difference, the *Hall voltage*, on the opposite sides of a conductor when an electric current is flowing through the conductor and a magnetic field is applied perpendicular to the current. A century later, in 1980, Klaus von Klitzing discovered the quantum Hall effect, QHE; he observed that the Hall resistance measured at a temperature of 1.5 Kelvin and in a magnetic field of the order of 15 Tesla, the Hall resistance variation with the strength of the applied magnetic field in semiconductor devices used as low-noise transistors is no longer linear but "stepwise" [243].

In such conditions, the Hall effect is changing its characteristics. The devices contain electrons which, though trapped close to the interface between two distinct parts of the transistor (junction), are highly mobile along this surface. At very low temperature the electrons move in two dimensions only, as if on a plane surface. The Hall resistance values where the steps occur do not depend on the properties of the material, they are given by a physical constant divided by an integer

$$R_{\text{Hall}} = \frac{h}{e^2}\,\frac{1}{i} \approx 25\,\frac{1}{i}\ \text{k}\Omega \quad \text{with} \quad i = 1, 2, 3, 4, 5, 6, 8, \text{and } 10,$$

where $h$ is the Planck constant and $e$ is the electron charge. The Hall resistance is *quantized* and the effect becomes the *integer quantum Hall effect* because the factor $i$ takes integer values. At the same time, it was observed that at magnetic field intensities where the Hall resistance takes quantized values the Ohmic resistance disappears, the material behaves like a superconductor.

The behavior of the Hall resistance was explained by the fact that the degenerate electron gas in a very thin layer at the interface between semiconductor and metal-oxide (the inversion layer in a metal-oxide-semiconductor field effect transistor) becomes fully quantized under the action of the electric and magnetic fields, when the device is operated at temperatures close to absolute zero and strong magnetic fields.

The electric field perpendicular to the surface orders the electron motion normal to the interface and the magnetic field produces the so-called *Landau quantization* of the motion parallel to the interface. In the absence of the magnetic field, the density of electron states in the interface plane is constant as a function of energy. In the presence of the magnetic field the states group themselves in Landau levels separated by the cyclotron energy $\hbar\omega_c$. The

energy levels are identical to those of the quantum harmonic oscillator

$$E_n = \hbar\omega_c(n + \frac{1}{2}) \qquad \text{with} \qquad n \geq 0,$$

where $n$ is the main quantum number and $\omega_c$ is the cyclotron frequency, $\omega_c = eB/M^*$. Here, $e$ and $M^*$ are the charge and effective mass, respectively, of the particle rotating under the action of a magnetic field of strength $B$. The presence of the Landau levels (the Landau quantization) can be felt when the mean thermal energy $k_BT$ (with $k_B$ the Boltzmann constant and $T$ the absolute temperature) becomes much smaller that the energy level separation, $\hbar\omega_c$.

The maximum number of electrons filling $n$ Landau levels ( assuming $T = 0$ K and the Fermi level lies just above the Landau level with quantum number $n$) is directly proportional to the strength of the applied magnetic field

$$N_e^L = n\frac{\Phi}{\Phi_0},$$

where $\Phi = B \times A$ is the magnetic flux through the two-dimensional electron gas plane of area $A$ and $\Phi_0$ is the "fundamental quantum of flux"[67], $\Phi_0 = h/e$ with $h$ the Planck constant and $e$ the electron charge. The ratio $N_\Phi = \Phi/\Phi_0$ gives the number of flux quanta enclosed by the 2-dimensional electron gas for a given value of the applied magnetic field.

The filling factor, $\nu$, is defined as

$$\nu = \frac{N_e^L}{N_\Phi}$$

and is equal to the quantum number $n$ of the maximum occupied Landau level when the position of the Fermi level is between $n$ and $n + 1$.

The inverse of the filling factor

$$\frac{1}{\nu} = \frac{N_\Phi}{N_e^L}$$

gives the average number of flux quanta enclosed by each electron at a given value of the applied magnetic field.

When the magnetic field increases, more electrons can fill a certain (higher energy) Landau level; thus, the Landau levels move relative to the Fermi level. When the Fermi level happens to be in the gap between two Landau levels, a region where there are no allowed electron states, the transport is dissipationless and the Ohmic resistance falls to zero, while the Hall resistance takes the value

$$R_H = \frac{1}{\nu}\frac{h}{e^2}.$$

For filling factors $\nu = integer$, this value is the same as the classical Hall resistance. The difference is that while in the classical effect the Hall resistance varies monotonously with the strength of the applied magnetic field, in the *quantum Hall effect* the Hall resistance exhibits relatively stable levels, or plateaux; it can not change from its quantized value corresponding to each plateau for a range of magnetic field strengths for which the Fermi level stays in a gap between Landau levels. When the magnetic field reaches a value where the Fermi energy

[67]The fundamental quantum of flux, or flux quantum represents the smallest unit of magnetic flux which can be enclosed by an electron orbit.

overlaps with a Landau level, electrons can scatter to new states, the Hall voltage changes and the Hall resistance changes by a finite value.

The quantum Hall effect plateaux correspond to gaps between electron states which extend across the sample and carry the current. At the same low temperatures the plateaux in pure, high mobility semiconductor heterojunctions are much narrower than in dirtier, lower mobility samples where the gap between states is larger due to impurities.

## 6.15 Fractional Quantum Hall Effect

In 1982, Daniel Tsui, Horst Störmer, and Arthur Gossard [425] discovered experimental evidence of a new phenomena called the fractional quantum Hall effect, FQHE. In experimental studies of the quantum Hall effect performed at even lower temperatures and even more powerful magnetic fields they observed an unexpected new step in the Hall resistance of a GaAs – $Al_{(1-x)}Ga_xAs$ heterojunction; the 2-dimensional electron gas resulting from donors located inside the AlGaAs was formed on the undoped GaAs side of the junction. The new quantized plateau was observed at Hall resistance

$$R_H = 3\frac{h}{e^2} \quad \text{for a filling factor} \quad \nu \approx \frac{1}{3}.$$

More such steps with heights obtained by dividing the constant $h/e^2$ by different fractional numbers have been observed above and between the integer plateaux. Some of the fractional filling factors experimentally observed in the lowest Landau level are:

$$\frac{1}{3}, \frac{1}{5}, \frac{1}{7} \qquad \frac{2}{3}, \frac{2}{5}, \frac{2}{7} \qquad \frac{3}{5}, \frac{3}{7}, \frac{3}{13} \qquad \frac{4}{7}, \frac{4}{9} \qquad \text{and} \qquad \frac{5}{9}, \frac{5}{11}.$$

**Quantum fluid and quasiparticles.** Less than a year after the experimental discovery of the fractional quantum Hall effect, Robert Laughlin [268] explained the new experimental facts as the result of the condensation of the two-dimensional electron gas in the heterojunction into a new type of collective ground state, a kind of quantum fluid incompressible due to long-range Coulomb interaction between electrons. Electrons are fermions and as such, are not able to condensate as a fluid; to behave like the quantum fluid proposed by Laughlin, the electrons are assumed to combine first with *flux quanta* (the magnetic flux tubes they orbit around when placed in a magnetic field) and form quasiparticles.

The $\frac{1}{3}$ step observed by Tsui, Strömer, and Gossard was explained by assuming that each electron in the two-dimensional electron gas captured three flux quanta and thus formed a sort of composite particles which behaved like bosons and could be condensed at very low temperatures and very strong magnetic fields. The ground state of this electron gas forms a quantum fluid and its elementary excitations are fractionally charged quasielectrons and quasiholes. This quantum fluid is incompressible for a given filling factor $\nu = n/m$ ($n$ and $m$ are integers with $m$ being odd) and has an energy gap for adding quasiparticles.

If only one electron is added to this ground state, it becomes excited and a number of quasiparticles is created. The quasiparticles are not true particles but manifestations of the interplay of the electrons in the quantum fluid.

Laughlin's quantum fluid explained the disappearance of the Ohmic resistance at the Hall resistance steps and all the fractional filling factors found experimentally. This quantum fluid is said to be incompressible because compressing or expanding it implies injecting particles with a cost in energy.

Assume the FQHE system is a full Landau level with collective excitation occurring at $\hbar\omega_c$. Its response to the compressive stress of an applied magnetic field is the following: first, the system generates Hall currents without compressing until the stress reaches a critical value and then it collapses by an area quantum $m2\pi a_0$ and nucleates a quasiparticle which is surrounded by a vortex of Hall current rotating in a sense opposite to that induced by the stress.

The states in the $n$th Landau level, $\mid m, n\rangle$, are eigenstates of the total angular momentum, with eigenvalues $\mathcal{M} = 3m + 2n$. Laughlin explained the $\frac{1}{3}$-effect in terms of the states in the lowest Landau level with $n = 0$ [267]. The states of the lowest Landau level, $\mid m, n = 0\rangle$, are described qualitatively as states of cyclotron motion about the origin with angular momentum $m$ and orbit of radius $\sqrt{2m}$.

The quantization of the electron separation in the quantum Hall effect has a physical origin: in a strong magnetic field, particles with the same charge sign do not repel one another, as if the Coulomb interaction were absent. But, they orbit each other with a speed proportional to the Coulomb force between them, at a distance $r$ determined by a negative binding energy $e^2/r$. The quantization of the interelectronic spacing follows from quantization of angular momentum.

Laughlin found that the first fractional QHE state observed experimentally at filling factor $\nu = \frac{1}{3}$ was correctly described by a many-body wave function he explicitly constructed; that wave function had a lower energy than the single particle energy. The Laughlin wave function can be used to describe states at any filling factor with an odd number denominator of the type $\nu = \frac{1}{(2m+1)}$, such as $\frac{1}{3}, \frac{1}{5}, \frac{1}{7}$. As $m$ increases, the energy difference between states is smaller and the plateaux become weaker. Laughlin also suggested that elementary excitations from the stables states are quasiparticles with fractional electric charge $\pm e^* = \pm e/m$. Thus, a quasihole is formed in the incompressible quantum fluid as a two-dimensional "bubble" of a size such that $1/m$ of an electron is removed. Laughlin's result was verified by others, for example [13]. The radius of a quasiparticle was estimated [268] as $R_{(\text{quasiparticle})} = \sqrt{2}a_0$ assuming an ion disc of charge $1/m$, where the magnetic length $a_0 = \sqrt{\hbar/m\omega_c} = \sqrt{\hbar c/eB}$. The energy required to create a quasiparticle, assuming accumulation of charge $1/m$ inside an ion disc, is approximately

$$\Delta E^{\text{disc}}_{(\text{quasiparticle})} = \frac{3}{2\sqrt{2}\pi}\frac{1}{m^2}\frac{e^2}{a_0}.$$

The energy to produce a quasiparticle does not depend on the position of the point where it is generated, as long as its distance from the boundary of the Hall sample is larger than its size. Elementary excitations of the quantum fluid are generated by the adiabatical transport of a flux quantum, $\Phi_0 = hc/e$, through the electron gas plane.

Nevertheless, states at filling factors $\nu = \frac{p}{q}$, where $p > 1$ or where $q$ is an even number, can not be described with Laughlin's wave function.

Haldane [194] and Halperin [199] developed a *hierarchical model* to account for filling factors of the type $\nu = \frac{p}{q}$ with $p > 1$ and $q$ odd. The model describes a hierarchy of

incompressible quantum fluid states with fractional filling factors given by continued fractions of the type

$$\cfrac{1}{m+\cfrac{\alpha_1}{p_1+\cfrac{\alpha_2}{\ddots+\cfrac{\alpha_n}{p_n}}}}$$

where,

$$m = 1, 3, 5, \ldots, \qquad \alpha_i = \pm 1, \qquad \text{and} \qquad p_i = 2, 4, 6, \ldots$$

These filling factors are rational numbers with an odd denominator; thus, a given fluid state at filling factor $\nu = [m, p_1, \ldots, p_n]$ cannot occur unless its parent state at $\nu = [m, p_1, \ldots, p_{n-1}]$ also occurs. Incompressible fluid states derive from parent incompressible states at simpler rational filling factors $\nu$. The most stable fluid states correspond to the simplest rational numbers with small values of $m$ and $p_i$ where the fluid densities are highest.

In the hierarchical model the Laughlin state at filling factor $\frac{1}{3}$ is a ground state and a *parent state* for quasielectrons and quasiholes which are excited out of it and condensed into higher order fractions, the so-called *daughter states*. For example, from the $\nu = \frac{1}{3}$ parent state one can obtain the daughter state with filling factor $\nu = \frac{2}{5}$ when quasielectrons are excited out of it, or the daughter state with filling factor $\nu = \frac{2}{7}$ when quasiholes are excited out. In turn, quasiparticles excited out of the daughter states condense into more daughter states and the process repeats itself with all the daughter states forming a hierarchy of states. The Haldane scheme is able to reproduce all quantum Hall effect steps corresponding to experimentally observed fractional filling factors with odd denominators

$$\nu = \frac{2}{7}, \quad \frac{1}{3}, \quad \frac{2}{5}, \quad \frac{2}{3}, \quad \frac{4}{5}, \quad \frac{4}{3}, \quad \text{and} \quad \frac{5}{3}.$$

They correspond to filling factors given (in Haldane notation) by $[3, 2]$, $[3]$, $[3, -2]$, $[1, 2, -2]$, $[1, 2]$, and $[1, 4]$, all derived from $m = 1$ and $m = 3$ hierarchies.

The hierarchical model has some problems: it does not account for the relative strength of quasielectrons or quasiholes states; after a few levels of the hierarchy, the number of quasiparticles will be larger than the number of electrons initially in the system; the system is not well defined between fractions; and the quasiparticles have to carry fractional charge.
The fractional quantum Hall effect represents only an indirect demonstration of the quantum fluid and its fractionally charged quasiparticles which carry the current.

The presence of quasiparticles has been observed directly in experiments like that of de-Picciotto *et al.* [114] who measured the quantum shot noise which probes the temporal behavior of the current and provides a direct way to measure the charge of the quasiparticles.

The charge transport in the FQHE regime can be explained based on the assumption that one-dimensional interacting electrons propagate along the edges of the plane where the two-dimensional electron gas is confined (the edges of the Hall sample) and behaves like a liquid [450]. The quantum shot noise generated by the weak backscattering of the current is proportional to the quasiparticle's charge and some other physical quantities given for a certain measurement. The shot noise measurements show that the current in the FQHE regime, at filling factor $\frac{1}{3}$ is carried by quasiparticles with charge $e/3$.

**Topological order of FQHE states.** The ground state degeneracy of the FQHE states depends on the topology of the space and that suggests that the FQHE states possess what

Wen [449] calls topological orders, that is, *the number of degenerate ground states is a topological invariant for any given quantum Hall phase.* This topological invariance is related to the observation that when the bulk flow of the quantum Hall effect fluid is restricted to the incompressible ground state, it becomes irrotational and incompressible.

The ground state degeneracy of the FQHE states does not change under arbitrary perturbations, therefore, it can be used to characterize different phases in phase space. Different FQHE states which are macroscopically distinct, can give rise to the same Hall conductance; if that is the case, they can be distinguished, at least partially, by their different ground state degeneracies (as estimated on torus and Riemann surfaces). The ground state degeneracy of the FQHE state is directly determined by the fractional statistics of the quasiparticles and not by the filling factor, $\nu$ [449].

Kitaev [240] analyzes the apparent paradox: the FQHE ground state is characterized by topological quantum order in its degeneracy and, at the same time, is separated by a finite energy gap from the excited states which makes all spatial correlation functions to decay exponentially. He shows that, though correlations can disappear at small scale, parts of the microscopic system keeps track of the topology through long-range entanglement manifested in the excitation properties.

**Fractional statistics.** The adiabatic theorem in quantum mechanics states that if we slowly vary the parameters in the Hamiltonian of a system, then a state occupying an isolated energy level, $E$ remains fully in that level - there are no "quantum jumps". If, after a sequence of adiabatic changes, the Hamiltonian returns to its starting value, the system returns to its original state, apart from a phase factor; the state vector (wave function) has a change in its phase.

Michael Berry showed [53] that in addition to the dynamical component given by $e^{-iEt/\hbar}$, which keeps track of the evolution in time, $t$, of any stationary state, there is also a geometrical phase component, $e^{i\gamma}$, which remains finite in the adiabatic limit. This is the so called Berry phase and is given by

$$\gamma = \oint i\langle\psi(t) \mid \frac{d\psi(t)}{dt}\rangle dt$$

where the integral can be taken over time, or over any parameter used to describe the path of the state vector (wave function) in the state space. The integral depends only on the geometry of this path; it does not depend on the coordinate used to describe the path or on the way the evolution occurred in time (provided that the circuit is traveled slow enough for the adiabatic approximation to be applicable).

For a particle of spin $S$ in a magnetic field $\mathbf{B}$, the geometric phase factor depends only on the eigenvalue, $s$, of the spin component along $\mathbf{B}$

$$e^{i\gamma(C)} = e^{-is\Omega(C)}$$

where $\Omega(C)$ is the solid angle subtended by $C$ when $\mathbf{B} = 0$; the phase can be changed by varying the field $\mathbf{B}$ round a circuit $C$. For fermions, particles of spin-1/2 $s = \frac{1}{2}$ such as electrons, protons, neutrons, a rotation with angle $\Omega = 2\pi$ (a whole turn of $\mathbf{B}$ around $C$) produces a phase factor $-1$. For bosons, particles of integer spin $s$ such as photons, the same rotation produces a phase factor $+1$. These quantum particles obey Fermi-Dirac and,

respectively, Bose-Einstein statistics. Upon exchange of two particles the two-particle state vector, $\mid \psi(1,2)\rangle$, acquires a phase and becomes

$$\mid \psi(2,1)\rangle = e^{i\pi\Theta} \mid \psi(1,2)\rangle.$$

A new exchange brings the two particles to the initial state

$$\mid \psi(1,2)\rangle = e^{i\pi\Theta} \mid \psi(2,1)\rangle = e^{i2\pi\Theta} \mid \psi(1,2)\rangle.$$

Quantum state vectors are single-valued and that requires

$$e^{i2\pi\Theta} = +1$$

where $\Theta =$ integer. In particular, for fermions $\Theta_{\text{fermions}} = 1$ and for bosons $\Theta_{\text{bosons}} = 0$. Such particles obey exchange statistics which are *integer statistics*.

In quantum mechanics the term statistics defines the effect of exchange of identical particles/quasiparticles. In 3D space, under adiabatic conditions, the state vector of a quantum particle executing a full loop around another particle is not distinct from the state vector of the same particle executing a closed circuit without a particle inside; both acquire the same phase factor.

In 2D space (a plane) the closed circuit with particle inside is topologically distinct from a circuit without a particle inside. The evolution in time of particles exchanging places while confined to a plane represents *braiding* of their trajectories in (2D + time). The interchange of the quasiparticles introduced by Wilczek in 1982 can give any phase; the factor $\Theta$ in the phase is no longer required to be an integer, it can be *any* real number and the particles obeying such *fractional statistics* wherefrom their name, *anyons* [456].

Wilczek's quasiparticles are composites made of charged particles tied to magnetic flux tubes in two dimensions and obey fractional statistics: when the positions of two quasiparticles are interchanged the state vector of the pair changes, it acquires a complex phase factor $(-1)^{\pm 1/m_k}$ (similar to Berry's geometrical phase) with the sign depending on the sense of rotation as the quasiparticles pass each other and $m$ is a rational number $m_k \geq 1$; this number is related to $\Theta$ above.

**The composite-fermion model.** The model was introduced by J. K. Jain [218] in 1989 is a unified approach that explains the integer QHE and the fractional QHE as two different manifestations of the same type of correlations in the phase factors present at fractional filling factors, $\nu = \frac{p}{2mp\pm 1}$ with $p$ and $m$ integers, and present at integer filling factors, $\nu = p$.

The composite-fermion consists of an electron, or hole, bound to an even number, $2p$, of magnetic flux quanta (vortices); a composite-fermion is interpreted as a particle. Electrons capture vortices to form composite fermions because they minimize their interaction energy in this way. The fractional QHE of electrons in external magnetic field $\mathcal{B}$ is the manifestation of the integer QHE of composite-fermions in an "effective" magnetic field $\mathcal{B}^*$.

The composite-fermions have integer charge like electrons, but because they move in an effective magnetic field they appear to have a fractional topological charge. As composite-fermions move along closed paths, the vortices bound to them produce Berry phases which cancel part of the Aharonov-Bohm phases originating from the motion in the applied magnetic field. Composite-fermions sense an effective magnetic field, $\mathcal{B}^*$, much smaller than the applied

field $\mathcal{B}$; the effective field can be even zero. Now, the composite-fermions form Landau-like levels in the reduced $\mathcal{B}^*$ and occupy $\nu^*$ of them and that distinguishes the composite-fermions from electrons.

Composite-fermions are collective, topological quantum particles; even a single composite-fermion is a collective bound state of all the electrons. Their topological character is due to the fact that the quantum phase associated with a closed loop around a vortex is exactly $2\pi$, independent of the size or shape of the loop. Composite-fermions represent the non-perturbative reorganization that takes place when a 2D electron gas is subjected to a strong magnetic field at very low temperature.

The composite-fermion model of the QHE predicts correctly all the observed fractions (with odd denominator), their relative intensities and the order they appear in when the sample purity increases. This model predicts all these fractions at the first level, in contrast to the standard hierarchical model which has to go down many levels to obtain some of the observed fractions; otherwise there is a close analogy between the states predicted by the two models and how they are obtained.

The statistics of composite particles is in general fractional [456] and is due to the special topology of the two-dimensional space where the electrons reside in the case of the fractional QHE. The relative motion of the particles confined to the 2D plane is limited to going around each other, they can not get out of the plane and move above or below each other.

During the evolution in time of the system, the paths (world lines) of the particles form braids[68] in the three dimensions defined by the 2D plane + time, see Section 1.18. Qubit braiding alone or together with some single-qubit gate and measurements could provide universal logic for quantum computation. The states of the FQHE were thought to have ordinary Abelian statistics associated with usual phase factors of $\pm 1$ when braiding quasiparticles corresponding to FQHE states with odd denominator filling factors (similar to exchanging identical fermions or bosons among themselves).

**Non-Abelian quasiparticles.** Considering the special topology of the quasiparticles, Moore and Read [300] suggested that some states could have non-Abelian statistics. Following the intuitive approach pioneered by Laughlin, they came up with several wave functions as verifiable solutions of the Hamiltonian for the quantum Hall system and they showed that the quasiparticle excitations of the states associated with those wave functions were non-Abelian.

In $(2+1)$ dimensions braiding sometimes changes the nature of the phase associated with the interchange of quasiparticles. The wave function of a set of excitations with specified positions and quantum numbers becomes a vector and each exchange of these quasiparticles gives rise to a unitary matrix transformation, instead of a simple alteration of its phase; that represents a *non-Abelian* action on this state vector.

In the case of quasiparticles with non-Abelian braiding statistics, the exchange of two particles does more than changing the phase of the wave function, it may rotate it into a different one in the space spanned by the state of the system. If we assume that a system of $a = 1, 2, \ldots, m$ quasiparticles at positions $x_1, x_2, \ldots, x_n$ is in a set of $m$, $m > 1$ degenerate states, $\psi_a$, then by exchanging particle 1 and 2 the state vector becomes

$$\psi_a \to A^{12}_{ab}\psi_b$$

[68]The braid group, sometimes denoted by $\mathcal{B}_n$, replaces the permutation group in describing the particle statistics in two dimensions.

If particles 2 and 3 are exchanged, then the state vector becomes

$$\psi_a \to A^{23}_{ab}\psi_b$$

For some pairs of quasiparticles the matrices $A^{12}_{ab}$ and $A^{23}_{ab}$ do not commute; such quasiparticles obey non-Abelian matrix statistics. If a system has a large set of non-Abelian states, the repeated application of the braiding transformation $A^{ij}_{ab}$ would allow the approximation of any unitary transformation to arbitrary accuracy.

The simplest of the wave functions proposed by Moore and Read corresponds to a spin-polarized $p$-wave pairing Bardeen-Cooper-Schrieffer state for a fixed number of composite fermions. The non-Abelian nature of this Moore-Read state comes from the collective degeneracy of the quasiparticles in this state. If quasiparticles are moved around each other, the state of the entire collective ensemble changes in a way that depends only of the topology of the move and the result is a unitary transformation in Hilbert space. As we know, unitary transformations in Hilbert space are the operations executed by a quantum computation.

The physical system that may serve as a platform for topological quantum computation is the fractional QHE with Landau level filling factor $\nu = 5/2$ associated with the quantized Hall plateau with transverse conductivity $\sigma_{xy} = \frac{5}{2}\frac{e^2}{h}$ first observed in 1987 by Willett *et al.* [457] at a temperature of about 100 mK. The state is observed on a regular basis in experiments on low-disorder samples of GaAs at very low temperature ($T = 15$ mK )and magnetic field of about 5 Tesla. The temperature dependence of the Hall resistance $R_{xx}$ at $\nu = 5/2$ measured in such experiments [464, 136] suggested an energy gap $\Delta_{5/2} = 310$ mK. This state is assumed to belong to the first excited Landau level with $N = 1$.

The state with $\nu = 5/2$ is easily destroyed and replaced by a strongly anisotropic phase when a strong magnetic field component $B_{\|}$ parallel to the 2D electron gas plane is added to the applied perpendicular field. The application of the transverse magnetic field initiates a phase transition of the system from a gapped quantum Hall (incompressible) phase into an anisotropic compressible phase.

Shortly after its discovery, Moore and Reed [300] developed a theory predicting that the elementary excitations of the $\nu = 5/2$ state are non-Abelian anyons. Then Nayak and Wilczek [304] found the corresponding braiding group representation.

Numerical calculations by Rezayi and Haldane [357] indicate that the 5/2 state belongs to the non-Abelian topological phase characterized by a so-called Pfaffian[69] quantum Hall wave function.

The existence of non-Abelian quasiparticles at filling factor $\nu = 5/2$ depends on two important hypotheses:

(1) the Coulomb repulsion in the second Landau level (where this state belongs) has a form favoring pairing of the quasiparticles and

(2) the electrons are fully-polarized.

Numerical calculations (for example, [357, 327] offer a strong evidence that the first premise is satisfied, especially when the finite layer thickness is taken into account in the electron-electron Coulomb interaction; when the 2D layer of electrons is assumed to have a thickness $d \sim 4l$ (where $l = \sqrt{\hbar c/eB}$ is the magnetic length) the exact ground state of the Coulomb

[69]The Pfaffian is the square root of the determinant of an antisymmetric (skew-symmetric) matrix. The Pfaffian is non-vanishing only for $2n \times 2n$ matrices.

Hamiltonian is very well approximated by the Moore-Read Pfaffian wave function [327] which assumes pairing.

**Experimental evidence of non-Abelian statistics.** Experiments set up to observe the quasiparticles with charge $e/4$ support the first hypothesis, but cannot distinguish and rule out some paired states which could have the same quasiparticle charge, but are Abelian and not fully-polarized. In a two-dimensional electron system under high magnetic field and at very low temperature where the FQH effect is present, electron-electron interactions lead to gaps in the bulk excitation spectra and the current can only be carried by states that propagate around the edges of the system.

A quantum point contact in a $2D$ electron system represents a constriction where counter-propagating edge states are brought close enough together so that quasiparticles can tunnel between them. The weak quasiparticle tunnelling depends [451, 453] strongly on the voltage difference between edges (or the current through the quantum contact point) and on temperature (the tunneling decreases slowly, as a power law, with increasing temperature). Its dependence on temperature provides a measurement of the effective quasiparticles charge, $e^*$, and the strength of the Coulomb interaction, $g$, both of them specific to the state. Quantum point contact measurements provide a discriminating probe of FQH effect wave functions.

The quantum point contact tunnelling experiments on a high-mobility GaAs two-dimensional electron system [348] indicate that various candidate states at $\nu = 5/2$ when compared with the experimental tunnelling data predict a quasiparticle charge $e/4$, but their Coulomb interaction strength parameter,$g$, can differ. The data are most consistent with charge $e/4$ and factor $g = 1/2$, values corresponding to two non-Abelian states, particle-hole conjugate of the Moore-Read state (anti-Pfaffian) and a $U(1) \times SU_2(2)$ edge state [452]. But, at the same time, the Abelian state with charge $e/4$ and $g = 3/8$ is not excluded by these experiments.

Direct evidence of effective quasiparticle charge $e/4$ in the state with fractional filling $\nu = 5/2$ in the second Landau level of the quantum Hall effect was found in the experiments of Dolev at al [132]. In their measurements, the current tunnels across a constriction (quantum point contact) between two opposite edge states of a Hall bar GaAs/AlGaAs sample; the quasiparticle charge is estimated from the current shot noise. Their conclusion is that a quasiparticle charge of $e/4$ at filling $\nu = 5/2$ indicate a paired composite fermion state, consistent with a spin-polarized Moore-Read Pffafian state [300] or . This experiment is not a direct test of the non-Abelian nature of the quasiparticles, but if the Moore-Read state is the correct one for filling factor 5/2, the observed quasiparticles with charge $e/4$ obey non-Abelian statistics.

**Quasiparticle charge.** Experiments of quasiparticle interferometry are expected to allow direct observation of their fractional charge and, eventually, their non-Abelian statistics. The two point-contact interferometer as proposed in [88] consists of a Hall bar with two weak barriers and can be used to study quantum interference effects in strongly correlated (electron) systems. In the quantum Hall phase, the electrons are in the bulk of the Hall bar specimen while the low energy excitations, the quasiparticles, lie on the edges. At the constrictions corresponding to the point-contacts, the quasiparticles carrying fractional charge and statistics can tunnel from one edge to another creating a tunnelling current which modifies the quantum Hall conductance.

The presence of two tunneling sites results in interference between the tunnelling events taking place at individual point-contacts (similar to the double-slit experiments discussed in [284]; the tunneling current becomes sensitive to the phase of the tunneling events. The interference of the tunnelling quasiparticles can be acted upon in three ways:

(1) changing the magnetic field at constant filling factor (it leads to Aharonov-Bohm oscillations);

(2) changing the number of quasiparticles enclosed by the interfering orbitals (it leads to fractional statistical oscillations);

(3) varying the source-drain voltage (it leads to so-called "Fabry-Pérot" oscillations).

Theoretically, it has been shown that these three effects can be studied separately and in doing so, they provide means of measuring the fractional charge and fractional statistics of the quasiparticles.

Interference in integer quantum Hall effect systems has been reported by several groups, the most recent with device designs analog to the optical Mach-Zehnder interferometer [220, 362, 84, 83]. Extensive measurements of fractional quantum Hall states interference performed recently [87, 86] have shown well defined Hall conduction periodic structures over a range of integer and fractional quantum Hall effect filling factors.

The proposed application of Fabry-Pérot interferometers to detection of non-Abelian braiding statistics will provide a verification of the microscopic ground state of certain observed fractional QHE states, such as the even-denominator $\nu = 5/2$. The interferometric measurement of charge has the advantage of a simple analysis - a direct comparison of the interference period at different quantum Hall state filling factors to that at state 5/2.

Interferometer devices constructed to accommodate the less robust fractional quantum Hall states [458] will permit interference of fractional state excitations with small coherence lengths; that is considered a crucial step towards experimentally proving the non-Abelian nature of the $\nu = 5/2$ state specifically associated with charge $e/4$. A charge $e/2$ is a potentially possible value for a $\nu = 5/2$ state with Abelian properties.

A two-point contact interferometer has the geometry of a Hall bar with two constrictions, such that the quasiparticle tunneling occurs between two edge currents on the opposite sides of the bar. The tunnelling current is sensitive to the total topological charge of anyons trapped inside the interferometer loop. Such an interferometer could form a universal gate set together with the short-range interaction between two anyons transported sufficiently close to each other where they could exchange a virtual quasiparticle, and with topological quantum computation operations.

**Electron spin polarization.** As for the second hypothesis concerning fully-polarized electrons, calculations [301, 124] indicate that a fully spin-polarized state is lower in energy than an unpolarized one. That is seen as excluding the hypothesis that the $\nu = 5/2$ state with charge $e/4$ could be a completely unpolarized state. However a partially-polarized state can not be excluded [124] and that could be either Abelian or non-Abelian in nature.

So far, there is little experimental evidence to support the fully-polarized state hypothesis. Experiments set up to directly measure the spin polarization at $\nu = 5/2$ through the resistively detected nuclear magnetic resonance (RD-NMR) technique have so far been unsuccessful. This technique relies on the hyperfine interaction in the following way: a finite nuclear spin polarization in the nuclear species present in the FGH effect sample creates an effective

magnetic field, $B_N$, which modifies the Zeeman energy, $E_{\text{Zeeman}} = g\mu_B(\mathcal{B} + B_N)$ and with it the spectrum of hyperfine levels. For example, for electrons in GaAs ($g = -0.44$), the effective magnetic field induced is $B_N = -5.3$T if the three spin-3/2 nuclear species $^{69}$Ga, $^{71}$Ga, and $^{75}$As present in the sample are all fully polarized.

Similar experiments at $\nu = 1/2$ have shown that the system remains partially polarized up to magnetic fields of about 8 Tesla [424]; the 5/2 state is typically observed at half the magnetic field. However, the results of all these experiments at low magnetic fields could also be explained assuming fully or partially spin-polarized, even unpolarized states; that leaves us with a confusing picture of the nature of the 5/2 state.

For the quantum Hall state with filling factor $\nu = 12/5$ recently observed experimentally [464], there is no direct evidence yet that it has non-Abelian statistics, though calculations indicate that it could be a state of non-Abelian anyons (a Read-Rezayi state). At the same time, it is also possible that it belongs to the conventional Abelian hierarchy of odd-denominator states. If the $\nu = 12/5$ state proves to be a Read-Rezayi ($\mathcal{Z}_3$ parafermion) state [352] than it has braiding statistics that allow universal topological quantum computation. Calculations show that the state is close to a phase transition between Abelian hierarchy state and the non-Abelian Read-Rezayi (parafermion) state and the system could be pushed over the phase boundary from an Abelian to a non-Abelian phase by engineering minor changes in the experimental Hall sample. However, besides the well quantized plateau observed in the original experiments [464] very little is known about the $\nu = 12/5$ state and its spin polarization and quasiparticle charge.

**Application of FQHE to quantum computing.** The quasiparticles in $\nu = 5/2$ state can not support universal computation; the transformations generated by braiding operations with such quasiparticles are not sufficient to implement all possible unitary transformations [160]. The braiding of two such quasiparticles has the effect of a 90 degree rotation in the Hilbert space of a system of multi-quasiparticles. Arbitrary unitary transformations can not be constructed through composition of 90 degree rotations. However, braiding together with a single-qubit $\pi/8$ phase gate and a two-qubit measurement form a universal gate set. The two extra gates can be implemented through some non-topological operations [75] and are not topologically protected.

A single-qubit phase gate can be implemented by bringing two quasiparticles (out of the four making a qubit) together at a given distance and waiting for some time before pulling them apart. The choice of distance and waiting time determines the phase gate. The other gate under consideration for the universal set is a non-demolition measurement of the total topological charge of any four quasiparticles out of the eight forming two qubits [160]. This is done with an interference measurement and depending on the total topological charge of the four quasiparticles, the two possible trajectories interfere with a phase $\pm 1$. In this way we measure the total parity of the two qubits. A non-Abelian topological $\nu = 12/5$ state would have the advantage of supporting universal topological quantum computation through quasiparticle braiding alone.

## 6.16 Alternative Physical Realizations of Topological Quantum Computers

In the previous section we have seen that quantum information can be processed by braiding the anyons in a $2D$-structure with an exponentially large topologically protected Hilbert space. Topological quantum computing introduced by Kitaev in 1997 and further developed by Friedman, Kitaev, Larsen, and Wang has the potential to revolutionize fault-tolerance and justifies the interest in alternative physical realizations of such systems.

Universal topological computation requires both a topological phase and quasiparticles with non-Abelian braiding statistics and with topological properties in a special class. A topological phase of a physical system is defined [305] as the system's ground state in the presence of multiple quasiparticles or in a non-trivial topology with a stable degeneracy resistant to weak (but finite) local perturbations. The stability of the ground state degeneracy to local perturbations necessitates the existence of an excitation gap.

The possible non-Abelian fractional quantum Hall states, such as $\nu = 5/2$ and $12/5$ are the first candidates. Incompressible fractional quantum Hall states are, so far, the only experimentally established topological phases. There are other physical systems, besides the fractional quantum Hall states, which are assumed to contain topological phases, such as transition metal oxides and ultra-cold atoms in optical traps.

The transition metal oxides are already known for displaying strong collective phenomena (high-$T_c$ superconductivity,huge magnetoresistance and thermoelectricity). At least for one material in this family, the $Sr_2RuO_4$ (strontium ruthenate), there is strong evidence that it is an odd-parity, spin-triplet ($p$-wave) paring superconductor at low temperature, $T_c \approx 1.5$ K. Half-quantum vortices in a thin film of this metal oxide superconductor are expected to exhibit non-Abelian statistics.

The ultra-cold atoms in optical traps represent another type of physical systems where topological phases could be present. The Hamiltonian can be controlled by tuning the lasers which define an optical lattice or by tuning through a Feshbach resonance. Topological phases can be generated in ultra-cold gases using two techniques: (1) by fast rotating dilute boson gases to make quantum Hall systems of bosons and (2) by using a gas of ultra-cold fermions with a $p$-wave Feshbach resonance to form a polarized chiral $p$-wave superfluid. Topological phases are hard to detect directly in the transition metal oxide or in ultra-cold atomic system unless the phase breaks parity and time-reversal symmetries spontaneously or in an external magnetic field.

CNOT gate based on parity measurement. In a system of $\nu = 5/2$ anyons, two quasiparticles form a two-level system equivalent to a spin-1/2 and can encode a qubit. That is the most efficient way of encoding a logical qubit, but in this case, due to charge superselection rules, a qubit can not be prepared in a superposition of basis states, e.g., $\mid 0\rangle \pm \mid 1\rangle$. The solution adopted by [75] is to represent a logical qubit by a group of four quasiparticles and to identify the logical basis states $\mid 0\rangle_{\text{logical}}$ and $\mid 1\rangle_{\text{logical}}$ with the physical states $\mid 0,0\rangle$ and $\mid 1,1\rangle$. In the four-quasiparticles representation, a qubit can be prepared in a superposition of basis states.

But, any two-qubit logical state that can be prepared by topological operations has the form of a product of two logical one-qubit states; therefore, in the four-quasiparticle representation, no entangled states of two logical qubits can be prepared topologically. This

apparent drawback could be circumvented: a state approximating the logical entangled target state is prepared using noisy non-topological operations and then it is purified following a purification protocol involving only topological quantum computing operations [75]. The entanglement of two logical qubits could be implemented through topological braiding operations and a non-destructive measurement of the total topological charge of four quasiparticles without destroying their pairwise correlations. In the case of $\nu = 5/2$ anyons, the total charge measurement is equivalent to a parity measurement.

The Pauli operators can be implemented by braiding operations; for example, if the qubit is represented by four quasiparticles, $1, 2, 3, 4$ in succession, the action of the Pauli operator $\sigma_z$ corresponds to braiding quasiparticle 1 around quasiparticle 2, or to braiding quasiparticle 3 around quasiparticle 4. The braiding of quasiparticle 2 around quasiparticle 3 implements the action of the Pauli operator $\sigma_x$. The action of Pauli operator $\sigma_y$ corresponds to braiding quasiparticle 1 around quasiparticle 3. Any four-quasiparticle braiding gate commutes with the parity operator.

A two-qubit `CNOT` gate for such multi-quasiparticle qubits was proposed to be implemented based on parity measurements [470]. In the case of $\nu = 5/2$ anyons what is measured is in fact the charge of four quasiparticles and that is equivalent to measuring their spin parity. The total topological charge of the quasiparticles can be measured using an interferometric technique. This implementation of a `CNOT` gate an ancilla qubit besides the target and control qubits.

The `CNOT` operation is executed through spin parity non-demolition measurements on groups of four neighboring quasiparticles each: two quasiparticle of the control qubit and the neighboring two quasiparticles of the ancilla qubit and then the other two quasiparticle of the ancilla qubit and the neighboring two quasiparticles of the target qubit. For example, assume anyons $1, 2, 3, 4$ form the control qubit, anyons $5, 6, 7, 8$ form the ancilla qubit, and anyons $9, 10, 11, 12$ form the target qubit. The parity measurement of anyons $3, 4, 5, 6$ gives the parity of the control and ancilla qubits while the parity measurement of anyons $7, 8, 9, 10$gives the parity of the ancilla and target qubits. The Hadamard rotations required by the `CNOT` operation are implemented through topological braiding of anyons of ancilla, respectively, target qubit. The last measurement is of anyons 5 and 6 of the ancilla.

This implementation uses only local braiding operations between the nearest-neighbor anyons composing a qubit, while that of [75] requires long-range anyon braiding operations between qubits. Both implementations would require some un-protected operations, but that could be avoided if $\nu = 12/5$ non-Abelian anyons were used.

The topological quantum computation tries to use the emergent properties of many-particle systems to encode and manipulate quantum information in a way which is fault-tolerant. The physics of topological phases evolves alongside topological quantum computation and it is influenced by it. So far, the incompressible fractional quantum Hall states, such as $\nu = 5/2$ and $12/5$ are the only experimentally established topological phases. Universal topological quantum computation has an additional requirement which is quasiparticles with non-Abelian braiding statistics and topological properties belonging to a special class.

## 6.17 Photonic Qubits

Over the past several years photons and an all-optical architecture have emerged as a leading approach to a physical implementation of quantum information processing. An optical qubit is often realized as one photon in two optical modes, such as horizontal and vertical polarization. But qubits can be encoded in any of the other degrees of freedom of single photons, such as time bin, or path. Quantum information may also be encoded in the continuous phase and amplitude variables of many-photon laser beams. Single photons are relatively free of decoherence.

**Single-photon source.** Single photons can be prepared by conditional measurements on a biphoton state generated in the process of *parametric down-conversion*, a process investigated experimentally in 1986 [214, 181]. The parametric down conversion is the elementary quantum process of the spontaneous decay of a photon of frequency $\omega_0$ into two new photons of frequencies $\omega_1$ and $\omega_2$ such that: $\omega_0 = \omega_1 + \omega_2$. The emission of the two photons is simultaneous and the process obeys the conservation relations for momentum and energy. The two photons are in a highly entangled quantum state known as biphoton, have the same polarization and travel in well defined directions.

In the 1986 experiment a coherent beam of light from an argon-ion laser (351.1 nm line) was traversing a (8 cm long) nonlinear crystal of potassium dihydrogen phosphate. Some incident photons split into two lower frequency photons, the so-called *signal* and *idler* photons. The signal and the idler photons were directed by two mirrors to pass through a beam splitter BS and the superposed beams interfere and were detected by two photodetectors.

The rate at which the photons were detected in coincidence was measured when the beam splitter was displaced from its symmetry position by various small distances $\pm c\delta\tau$ . It was observed that the signal and idler photons: (i) had no definite phase; (ii) were mutually incoherent: they exhibited no second-order interference when brought together at either one of the two photodetectors; (iii) exhibited fourth-order interference effects as demonstrated by the coincidence counting rate between the two photodetectors, the so called *cosine modulation*.

The individual frequencies $\omega_1$ and $\omega_2$ had large uncertainties (though $\omega_0$ was very well defined. The large uncertainties were determined by the pass bands of the interference filters which were of the order of $5 \times 10^{12}$ corresponding to a coherence time for each photon of 100 fsec [70]. The two-photon probability amplitudes at the two photodetectors were expected to interfere only if they overlap to this accuracy. Detection of a photon in one of the emission channels causes the collapse of the non-local photon pair and the projection of the quantum state of the other channel into a single-photon well defined state. Most single-photon sources use entangled photons produced through parametric down-conversion.

**Linear optics quantum computation.** Single-qubit logic gates acting as polarization rotations, are easy to implement using birefringent wave plates; polarizing beam splitters can achieve conversion between polarization and path encoding. One of the earliest proposals for quantum information processing uses photons to implement a quantum logic (Fredkin) gate [294]. In fact, optical systems represent the only realistic implementation of long-distance quantum communication; they are applied to quantum cryptography (quantum key distribution) [216].

[70] A femtosecond is $10^{-15}$ seconds

The realization of entangling, two-qubit logic gates required for universal quantum computation, is a major difficulty. An optical `CNOT` gate can be implemented only if a $\pi$ phase shift is applied to the target qubit when the control qubit has value 1; this conditional phase shift is applied after a Hadamard gate creates a superposition state for the target qubit and a second Hadamard gate flips its value. Normally, the phase shift would require optical nonlinearities stronger than what is available in conventional nonlinear materials. Otherwise, electromagnetically induced transparency (EIT) and atom-photon interactions enhanced by an optical cavity (cavity quantum electrodynamics) were considered as the only available option.

In 2001, Knill, Laflamme and Milburn (KLM) [251] showed that quantum information processing could efficiently (with polynomial resources) be performed with passive linear optics. What linear optics quantum computation requires are single-photon sources, beam splitters, phase shifters, photo-detectors, and feedback from photo-detector outputs. Photonic (bosonic) qubits are defined by states of optical modes, each mode being characterized by the number of photons. For example, the basic states of a photonic qubit encoded in two modes, $m_1$ and $m_2$ are given by $\mid \mathbf{0}\rangle \rightarrow \mid 0\rangle_{m_1} \mid 1\rangle_{m_2}$ and $\mid \mathbf{1}\rangle \rightarrow \mid 1\rangle_{m_1} \mid 0\rangle_{m_2}$. As any other implementation, an optical quantum computer must have the means to prepare state, apply quantum gates and extract the result. The initial state is a vacuum state, a state in which there are no photons in any of the used optical modes. Photons are added to the initial state using a single photon[71] source. Quantum states containing a definite number of photons (Fock states) are required in quantum cryptography for increased capacity and security in communication channels and for efficient linear quantum computation.

With the injection of a single photon in a well defined state, the state of any given mode can be set to the one-photon state $\mid \mathbf{1}\rangle$ in this way. The state preparation should be non-deterministic, that is, its probability of success is non zero and it is known whether it succeeded or not.

Simple optical elements like beam splitters and phase shifters generate evolutions which preserve the total photon number and their effects on the creation operator of each mode can be described by unitary matrices. The unitary matrix associated with phase shifter $P_\theta$ is $u_{(P_\theta)} = e^{i\theta}$ and the unitary matrix associated with beam splitter $B_{\theta,\phi}$ is

$$u_{(B_{\theta,\phi})} = \begin{pmatrix} \cos(\theta) & -e^{i\phi}\sin(\theta) \\ e^{-i\phi}\sin(\theta) & \cos(\theta) \end{pmatrix}$$

A sequence of beam splitters and phase shifters is able to implement the operation associated with any such unitary matrix up to a global phase. The state space of a photonic qubit is preserved in the interaction with phase shifters and beam splitters; their effect can be expressed in the qubit basis using Pauli matrices $\sigma_x, \sigma_y$ and $\sigma_z$. For example, a phase shifter $P_\theta^{(1)}$ corresponds to the application of $e^{-i\sigma_z(\theta/2)}$ operation up to a global phase and a beam splitter $B_\theta^{(12)}$ applies the operation $e^{-i\sigma_y\theta}$. Thus, one-qubit rotations can be implemented with linear optics [251].

A two-qubit gate such as the conditional sign flip in one mode was proposed to be implemented by using two ancilla modes with one prepared photon and post-selection based

[71] The single photon character is usually stated by showing that the amount of energy flowing during a certain characteristic time, such as coherence time, or time-of-flight between source and detector, is small compared to $h\nu$ [181]. The detection process appears discrete.

on measuring the ancillas. The output is accepted only if the ancilla photon counters detect one photon in only one of the ancilla modes and none in the other. For any other detection pattern, the CNOT gate is not applied. Knill, Laflamme and Milburn showed that measurement induced nonlinearity was sufficient to implement a controlled-NOT gate and to obtain efficient quantum computation.

In the case of a nondeterministic CNOT gate, the control and target qubits, encoded in photon polarization, for example, and two ancilla photons simultaneously enter an optical network of beam splitters (a beam splitter implements a Hadamard gate) where the paths of the four photons are combined. The CNOT operation is considered to have been applied to the state of the emerging control and target qubits conditional on a single photon being detected at both ancilla detectors. As such, the probability of a successful CNOT gate was estimated to be $P < 1/16$; the probability of a successful computation decreases exponentially with the number of applied CNOT gates.

The success probability of a two-qubit gate can be increased by reducing it to a state preparation problem by using quantum teleportation. The control qubit (mode (1)) in unknown state and one of the ancilla (mode (2)) photons are subjected to a Bell measurement. Before the measurement, the control qubit is in unknown state $\mid c\rangle = \alpha_0 \mid 0\rangle^{(1)}$ and the ancilla (mode (2)) photon is prepared in a maximally entangled state, $\mid \varphi\rangle^{(23)} =\mid 01\rangle^{(23)}+ \mid 10\rangle^{(23)}$ with the second (mode (3)) ancilla. The measurement is performed in the Bell basis $\{\mid 01\rangle^{(12)}\pm \mid 10\rangle^{(12)}, \mid 00\rangle^{(12)}\pm \mid 11\rangle^{(12)}\}$. In the process the unknown input state of the control and target qubits can be teleported to the first and second ancilla, respectively.

The nondeterministic CNOT can be independently repeated on the ancillas until the gate succeeds and then we proceed with the teleportation [176]. A successful gate could be detected by measuring the ancilla photons in some appropriate state. The control and target qubits are preserved until the gate works and CNOT operation can be implemented deterministically. Quantum teleportation had been realized experimentally with single photons in 1997 [69].

The techniques applied include near-deterministic non-destructive parity measurements to assess photon loss and sign, a method for creating entanglement by local measurements of uncorrelated photons shared with beam splitters, and nearly unconditional quantum teleportation and Bell state measurements with linear optics. The Knill-Laflamme-Milburn scheme for efficient linear optics quantum computation with photonic (bosonic) qubits shows that these techniques can be implemented with several types of operations: (i) preparation of $X, Y$, or $Z$ eigenstate (eigenvalue $+1$ or $-1$); (ii) measurements as $X, Y$, and $Z$; (iii) rotations $X_{180}, Y_{180}$, and $Z_{180}$; rotation $X_\phi$; (iv) rotation $Z_{90}$; (v) rotation $(Z^{(1)}Z^{(2)})_{90}$ on two physical qubits which encode one logical qubit. The 180° rotations of logical qubits can be applied by using the corresponding 180° rotations directly on the physical qubits.

The computation result is read out by measuring a mode with a photo-detector associated with a photon counter; the measurement determines if one or more photons are present in that mode but destroys it/them in the process. The sources of errors for linear optics quantum computation are: photon loss at the single-photon source or during processing, detector inefficiency which can be considered as loss of photons and phase errors.

In principle, the probability of photon loss can be estimated from the characteristics of the optical devices used by the implementation, but the loss itself introduces a so-called "erasure" failure, whereby we no longer know what happened to the state of the qubit. The probability for erasure as well as phase errors can be reduced if several photons are used to encode logical

qubits and error correcting codes are applied. A linear optics implementation of a quantum computer as proposed by [251] requires good mode matching, good synchronization of pulses or delay lines which can be rapidly controlled, tunable phase shifters and beam splitters, single-photon sources, and fast, high efficiency single-photon detectors.

The linear optics quantum computation scheme proposed by Knill, Laflamme, and Milburn was not economical in its use of optical components or ancilla photons. Several groups have been working on improving the theoretical efficiency of the gates and reducing their complexity [331, 253, 210]; improved `CNOT` gates have been demonstrated experimentally the following years by Pittman *et al.* [332], O'Brien *et al.* [312], Gasparoni *et al.* [169], among others. Encoding against measurement errors was also demonstrated experimentally.

A proof of principle of a high fidelity realization of a two-qubit $Z$-measurement quantum error correction scheme was provided for a two-qubit code of the type $\mid \varphi\rangle_L = \alpha \mid 0\rangle_L + \beta \mid 1\rangle_L = \alpha(\mid 00\rangle + \mid 11\rangle) + \beta(\mid 01\rangle + \mid 10\rangle)$ where the encoded state is prepared from an arbitrary input state [313]. The input into the target mode of a nondeterministic optical `CNOT` gate was a single qubit prepared in an arbitrary state $\mid \varphi\rangle = \alpha \mid 0\rangle + \beta \mid 1\rangle$ using a half wave plate and a quarter wave plate. An ancilla qubit prepared in real equal superposition $\mid 0\rangle + \mid 1\rangle$ with a half wave plate is input into the control mode of the gate. The resulting two-qubit encoded state $\mid \varphi\rangle_L$ generated from the possible inputs $\mid \varphi\rangle = \mid 0\rangle, \mid 1\rangle, \mid 0\rangle \pm \mid 1\rangle, \mid 0\rangle \pm i \mid 1\rangle$, was measured using two-qubit quantum state tomography and found with high fidelity. The two-qubit measurements were performed using a pair of analyzers each consisting of a quarter wave plate and a half wave plate followed by a polarizing beam splitter and a single-photon counter; each analyzer could perform any one-qubit projective measurement.

A combination of two-qubit coding and fast feed forward control techniques was used by [333] to demonstrate quantum error correction for measurement errors in linear optics quantum computation. In this experiment the qubit state $\mid \varphi\rangle$ and the ancilla state $\mid 0\rangle$ were represented by the polarization states of two single-photons from a parametric down-conversion pair. The logical state $\mid \varphi\rangle_L$ was encoded probabilistically using linear optics and post-selection. The feed-forward-controlled bit flip was performed with an electro-optic polarization rotator triggered by the output of single-photon detectors. An intentional $Z$ measurement was performed on one of the photons and the success of the error correction procedure was verified by comparing the corrected polarization state of the remaining photon with the initial state $\mid \varphi\rangle$.

The encoding was done with a single polarizing beam splitter (PBS) which transmits horizontally ($\mid H\rangle$) polarized single photons and reflects the vertically ($\mid V\rangle$) polarized ones. A qubit photon in arbitrary state $\mid \varphi\rangle = \mid 0\rangle + \mid 1\rangle$ was sent into one of the input ports of the PBS and the ancilla in state $\mid 0\rangle$ was sent into the other input. Coincidence basis measurements on the two outputs of PBS made sure that one photon will exit each output port. Single photons from pairs produced by a down-conversion source were used as qubit and ancilla. The encoding operation for an arbitrary qubit state $\mid \varphi\rangle$ can be understood as a two-photon quantum interference effect that uses a beam splitter and post-selection to generate polarization entanglement, in the coincidence basis, from an initial product state of two single photons.

The fidelity of the logical qubit state $\mid \varphi\rangle_L$ was determined by the quality of the two-photon interference effects. The $Z$ measurement was performed with a second PBS controlled in such a way that a photon with polarization corresponding to state $\mid 0\rangle$ would be transmitted to a

designated single-photon detector, while photons with polarization corresponding to state $\mid 1\rangle$ would be reflected to another designated detector. The detector recording state $\mid 1\rangle$ photons triggered the application of a high voltage pulse on the polarization rotator whose special axis orientation would cause a bit flip in the computational basis. The polarization state exiting was measured by a different detector. The coincidences between events in this detector and the events in $Z$-measurement detectors represented the coincidence-basis operation of this specific purpose encoding device which was proven to be able to accomplish two-qubit encoding.

**Some new contributions to linear optics implementations.** In 2004, Michael Nielsen [310] has shown that efficient linear optical quantum computation is possible without the teleportation and $Z$-measurement error correction steps used by KLM scheme. Nielsen combines KLM non-deterministic gates with fewer optical components with the cluster-state model proposed by Raussendorf and Briegel [350].

The cluster state is a massively entangled state of a network of many qubits; the computation is executed as a sequence of single-qubit measurements with the measurement outcome determining the basis for the measurement on the next qubit; the answer state is left in the output qubits in the network. The preparation of the cluster state can be probabilistic and non-deterministic linear optics `CNOT` gates can be used in the process. The advantage is that much of the large overhead required for error encoding used to make non-deterministic `CNOT` gates is no longer necessary.

A first experimental demonstration of an all-optical one-way implementation of Deutsch algorithm was performed in 2006 on a four-qubit cluster state [416]. The entangling operations required by the computation were realized by using the cluster resource and the nonlinearity induced by detection. The qubit states $\mid 0\rangle_i$ and $\mid 1\rangle_i$ were embodied by the horizontal and the vertical polarization state, respectively, of one photon populating a spatial mode $i = 1, \ldots, 4$. The input photons were produced in pairs in entangled states of the form $(1/sqrt2)(\mid 00\rangle - \mid 11\rangle)_{ab}$ and $(1/sqrt2)(\mid 00\rangle + \mid 11\rangle)_{ab}$ following two passages of an ultraviolet pump-laser beam through a nonlinear beta-barium-borate crystal. The cluster state $\mid \varphi\rangle_c = (1/2)(\mid 0000\rangle + \mid 0011\rangle + \mid 1100\rangle - \mid 1111\rangle)_{1234}$ was prepared through post-selection, a four-photon coincidence event at the single-photon detectors facing each spatial mode.

The algorithm was executed by using quarter wave plates, half wave plates, polarizing beam splitters and photocounter pairs for polarization measurements in arbitrary bases of the photons in mode $i$. In this implementation of Deutsch's algorithm the application of a `CNOT` gate is replaced with as a series of rotations, Hadamard gates and a controlled-phase gate. On a linear cluster the two-qubit operations `CNOT` or `CPHASE` (controlled-PHASE) can be implemented by changing the order of measurements. Given a perfect cluster state and feed-forward control based on measurements made on ancilla qubits could ensure a deterministic photonic quantum computation.

With active feed-forward control, whereby a later measurement depends on an earlier measurement and its fed forward result, Prevedel *et al.* [346] have demonstrated 'error-free' single-qubit and two-qubit gate operations in an implementation of Grover's search algorithm. The resources required by a linear optics implementation can be reduced by several orders of magnitude if `CNOT` gates are replaced with other techniques.

Recent developments show that a hybrid approach, such as storing quantum information in the single-photon sources which are themselves inherently quantum mechanical, may have

advantages; the photonic qubits are easy to transport and are robust against decoherence, but the execution of gate operations required to prepare photon entanglement encounters difficulties. Even the more improved methods of preparing photon entanglement based on measurement induced nonlinearities have to contend with probabilistic interactions and most often require photon detection based on number discrimination. For large entangled multi-photon states the result is long preparation times and a large number of resources for photon storage which limits the eventual use of on-chip devices

A photonic module proposed by [120] would rapidly prepare a large class of entangled photon states for quantum computation and communication applications in a completely deterministic way without the need for number-discriminated photon detection. The internal construction of such a module is independent of the state being prepared, no single photon detection is necessary and the coherence time required for the atomic qubit is limited by the time it takes to measure specific stabilizers (of the Gottesman formalism) describing the state. The module is a generic resource for preparation of any geometric graph state, including states for optical cluster state computation, Bell states for quantum cryptographic protocols for quantum dense coding, for purification protocols and quantum repeaters, $N$ photon GHZ states for loss protection scheme for optical quantum computation.

The operation of the module is based on atom/cavity interaction, where the atomic system used in the cavity can be rubidium or nitrogen-vacancy (NV) diamond; they have coherence times long enough to perform two-photon parity measurements, or to prepare highly entangled states very quickly (NV diamond). Single photons of known polarity from a single-photon source pass as a two-photon train through the module where they interact with an atom/cavity system and must induce a non-demolition bit-flip on the two-level atomic qubit. The atom/cavity qubit is initially prepared in state $\mid 0\rangle$. After the two photons have passed through the module, the state of the atom/cavity qubit is measured and the two output photons are projected into an entangled Bell state of either even parity, if the atom qubit is in state $\mid 0\rangle$, or odd parity if the atom qubit is in state $\mid 1\rangle$.

The measurement of the atom/cavity system never collapses the two photons to an unentangled state. The number of entangled photons that are prepared by the module depends only on the number of photons sent through the module, no change in the internal structure of the module is necessary. The stabilizer formalism used for describing large entangled states is closely liked to the concept of parity measurement. Any $N$ photon stabilized state can be prepared by performing a parity check on the $N$ stabilizers describing the state.

Assuming that the individual photons within a train can be selectively routed (using the time tag of each pulse) and local operations are applied to them, the parity measurement performed by the module prepares a stabilizer state. For example, a cluster state intended for quantum computation has a well known stabilizer structure [350] with a maximum parity-weight of five. The optical module has to maintain a decoherence time of the atom/cavity system long enough to permit the passage of five photons between initialization and measurement. In the end multiple module combined with an appropriate single-photon source, optical wave plates and classical routing can be used to build a static on-chip system for fast preparation of specific entangled states either large cluster states for quantum computation or multiple Bell pairs in sequence for quantum communications and cryptography.

The most important challenge to optical quantum computing, by far more serious than photon sources, detectors and nonlinearities, and similar to decoherence in matter-based

qubits, is the photon loss. However, the advances in photonic quantum computing support photonic qubits and benefits photonic quantum communication between matter qubits such as trapped ions, quantum dots, solid state impurities/dopants.

## 6.18 Summary and Further Readings

Any physical implementation of a quantum information processing system must be based on a system insulated from unwanted interactions with the environment, must be controlled through some required interactions implementing logical quantum gates, and must be measured. The physical implementation must satisfy a set of generic requirements formulated by DiVincenzo in 2000 and recently reformulated and generalized for systems with low level of decoherence.

The first technology for quantum information processing devices proposed I. Cirac and P. Zoller [97] and demonstrated experimentally by C. Monroe *et al.* [298] was the cold ion trap. The ion traps use laser-induced coupling between electron spin states and the vibrational motion (phonon) of relatively heavy charged particles, ions, trapped in high vacuum at very low temperature.

The main advantages of ion traps are: some of the techniques have already been developed; the decoherence rates associated with the decay of excited ionic state and the damping of the ionic motion can be made small; quantum jumps, can be used for reading out the result of a computation with high probability. The major problems are: spontaneous transitions in the vibrational motion due to heating; internal transitions in the ions driven by thermal radiation; experimental instabilities in the laser beam power, RF voltages; mechanical vibrations; and fluctuating external magnetic fields.

The first experimental demonstration of a `CNOT` quantum gate was reported by a group at the National Institute of Standards and Technology (NIST) in Boulder, Colorado, [298]. The characteristic states of the two qubits are two internal and two external states, respectively, of a single $^9Be^+$ ion stored in a coaxial-resonator RF ion trap.

The very long coherence times associated with the nuclear spins stimulated David Cory [104, 105] and Neil Gershenfeld and Isaac Chuang [171] to propose the use of Nuclear Magnetic Resonance (NMR) as an alternative implementation of a quantum computer. Single-qubit gates, spin state vector rotations in Hilbert space, can be directly implemented in the rotating frame by applying RF pulses to individual nuclear spins. The liquid-state NMR has been used to demonstrate Shor's factoring algorithm by using a seven-qubit custom-synthesized molecule [94, 432] to factor the number 15.

Solid-state is extremely versatile and could offer solutions for a scalable quantum computer technology as most phenomena possible in physics can be embodied in a solid-state system. Solid state physics has enabled the creation of a wide variety of artificial structures and devices thus, it is no wonder that in 2000 the number of proposals for solid-state implementations of quantum computers was almost equal to all the other proposals put together.

Quantum dots are atom-like compound semiconductors structures where charge carriers are confined in all three spatial dimensions. A semiconductor quantum dot may contain from a few thousands to tens of thousands of atoms arranged in a nearly defect-free three-dimensional crystal lattice. The carriers in the quantum dot interact strongly with lattice

vibrations and could be strongly influenced by defects, surfaces, or interfaces. The size of a quantum dot is of the order of the Fermi wavelength. Quantum dots can be lithographically defined or self-assembled.

Quantum computing implementation based on semiconductor quantum dot electron spins has two significant advantages: the quantum dot arrays could be scalable to $\geq 100$ coupled qubits; and the spin decoherence times for conduction band electrons in III-IV and II-VI semiconductors are "very" long.

In 1982, Daniel Tsui, Horst Störmer, and Arthur Gossard [425] discovered experimental evidence of a new phenomena called the fractional quantum Hall effect, FQHE. Robert Laughlin [268] explained the new experimental facts as the result of the condensation of the two-dimensional electron gas in the heterojunction into a new type of collective ground state, a kind of quantum fluid incompressible due to long-range Coulomb interaction between electrons.

The $\nu = \frac{1}{3}$ step observed experimentally was explained by assuming that each electron in the two-dimensional electron gas (2DEG) captured three flux quanta and thus formed a sort of composite particles which behaved like bosons and could be condensed at very low temperatures and very strong magnetic fields. The ground state of this electron gas forms a quantum fluid and its elementary excitations are fractionally charged quasiparticles and quasiholes. When only one electron is added to the ground state of the 2DEG, it becomes excited and a number of quasiparticles is created; the quasiparticles are not true particles but manifestations of the interplay of the electrons in the quantum fluid.

Topological quantum computing introduced by Kitaev in 1997 and further developed by Friedman, Kitaev, Larsen, and Wang has the potential to revolutionize fault-tolerance and justifies the interest in alternative physical realizations of such systems. Universal topological computation requires both a topological phase and quasiparticles with non-Abelian braiding statistics and with topological properties in a special class. The system's ground state should have a stable degeneracy resistant to weak local perturbations in the presence of multiple quasiparticles or in a non-trivial topology. So far, the incompressible fractional quantum Hall states, such as $\nu = 5/2$ and $12/5$ are the only experimentally established topological phases. The topological quantum computation tries to use the emergent properties of many-particle systems to encode and manipulate quantum information in a way which is fault-tolerant. There are other physical systems, besides the fractional quantum Hall states, which are assumed to contain topological phases, such as transition metal oxides and ultra-cold atoms in optical traps.

After almost two decades of quantum information science we are prepared to make an objective examination of different complete architectures of quantum computers, of the resources required to implement algorithms of useful sizes with negligible error. In practice, we may discover that imperfections in the coherent control of the qubits may hinder the good performance of the quantum computer more than the decoherence itself. A large scale, fault tolerant quantum computer belongs to a new form of technological reality.

New research papers are published at an overwhelming rate in major science journals such as *Physical Review A, Physical Review B, Physical Review Letters, Nature, Applied Physics Letters, New Journal of Physics, Journal of the Optical Society of America B, Proceedings of the Royal Society, London.* The *Virtual Journal of Quantum Information* distributed online by the American Institute of Physics (AIP) and the American Physical Society (APS)

contains citations of all recent publication in the field. The site *http://arxiv.org* owned and operated by Cornell University gives free access to e-prints of research papers in several fields of science; of special interest physical implementations are those published under identifiers such as quantum physics, condensed matter.

# 7 Appendix. Observable Algebras and Channels

In this appendix we present a formal framework for some of the concepts we are familiar with from Quantum Mechanics and Quantum Information Theory, while avoiding to introduce more intricate concepts from Operator Theory. Our presentation follows closely the notations and conventions in [231].

A physical system, be it a classical or a quantum system, is characterized by its *state* $\rho$ and the *effect* $A$ of a measurement of a physical property, or an observable; an effect is associated with an operator and $A$ denotes both the effect and the operator. We assume that the experiment to observe the system produces statistical results; the effect $A$ may be observed or not in state $\rho$. The formalism is based upon two sets, $\mathcal{S}$, the set of states, and $\mathcal{E}$, the set of effects/operators, and a map:

$$(\rho, A) \in \mathcal{S} \times \mathcal{E} \mapsto \rho(A).$$

The functional[72] $\rho(A) \in [0,1]$ is the probability of observing effect $A$ (or applying the operator A) to a system in state $\rho$.

Let $\mathcal{H}$ be a Hilbert space and $\mathbb{B}_{\mathcal{H}}$ be the algebra of bounded operators[73] in $\mathcal{H}$ and $\mathbb{I}$ the identity operator in this algebra. If $\psi$ is an element of $\mathbb{B}_{\mathcal{H}}$ and $A$ is an operator then the expression $\langle \psi, A\psi \rangle$ is the scalar product of the elements before and after the transformation $A$.

An observable algebra, $\mathcal{A}$, is a linear subspace of $\mathbb{B}_{\mathcal{H}}$ with the following properties:

- $\mathcal{A}$ contains the identity, $\mathbb{I} \in \mathcal{A}$,
- $\mathcal{A}$ is closed under multiplication of operators:

$$(A, B) \in \mathcal{A} \Rightarrow AB \in \mathcal{A}$$

- $A^*$, the adjoint of operator $A$ is also in $\mathcal{A}$.

$$A \in \mathcal{A} \Rightarrow A^* \in \mathcal{A}.$$

The *norm* of operator $A \in \mathcal{A}$, inherited from $\mathbb{B}_{\mathcal{H}}$, is defined as

$$|| A ||= \sup_{||\psi||-1} =|| A\psi ||, \quad \forall \psi \in \mathcal{H}.$$

The *operator ordering*, inherited from $\mathbb{B}_{\mathcal{H}}$, is defined as

$$A \geq B \Leftrightarrow \langle \psi, A\psi \rangle \geq \langle \psi, B\psi \rangle, \quad \forall \psi \in \mathcal{H}.$$

[72] A functional is a real-valued function on a vector space of functions.

[73] A bounded operator is a linear transformation $L$ between vector spaces $X$ and $Y$ with the property that the ratio of the norms of $L(x)$ and $x \in X$ is bounded by the same number $k$:

$$|| L(x) ||_Y \leq k || x ||_X \quad \text{with} \quad k > 0, \quad \forall x \in X, x \neq 0.$$

The *dual* space of $\mathcal{A}$, $\mathcal{A}^*$, is the set of all linear forms on $\mathcal{A}$. The set of the states of the system described by the algebra $\mathcal{A}$, $\mathcal{S}(\mathcal{A})$ is:

$$\mathcal{S}(\mathcal{A}) = \{\rho \in \mathcal{A}^* \mid \rho \geq 0, \rho(\mathbb{I}) = 1\}.$$

In this expression the positivity of the functional $\rho$ means: $\rho \geq 0 \Leftrightarrow \rho(A) \geq 0, \forall A \in \mathcal{A}$. The set of the effects of the system described by the algebra $\mathcal{A}$, $\mathcal{E}(\mathcal{A})$, is:

$$\mathcal{E}(\mathcal{A}) = \{A \in \mathcal{A} \mid A \geq 0, A \leq \mathbb{I}\}.$$

Both spaces are *convex*; if $0 \leq \lambda \leq 1$ then:

$$\forall \rho, \sigma \in \mathcal{S}(\mathcal{A}) \Rightarrow \lambda\rho + (1-\lambda)\sigma \in \mathcal{S}(\mathcal{A}) \quad \text{and} \quad \forall A, B \in \mathcal{E}(\mathcal{A}) \Rightarrow \lambda A + (1-\lambda)B \in \mathcal{S}(\mathcal{A}).$$

The extreme points[74] of the two sets have special properties:

- the extreme points of $\mathcal{S}(\mathcal{A})$ are called *pure states*;
- the extreme points of $\mathcal{E}(\mathcal{A})$ are called *propositions*; they are effects that occur with probability one.

When $\mathcal{A} = \mathbb{B}_{\mathcal{H}}$, the observable algebra $\mathcal{A}$ coincides with $\mathbb{B}_{\mathcal{H}}$ and then the correspondence between the concepts from the algebra of observables and the concepts we are familiar with from quantum mechanics is straightforward:

$\boxed{\mathcal{A} \mapsto \mathcal{H}_d, \text{Hilbert space of dimension } d.}$

$\boxed{\text{Vectors } \psi \in \mathcal{A} \mapsto \texttt{ket} \text{ vectors } \mid \psi\rangle \in \mathcal{H}_d.}$

$\boxed{\text{Vectors } \psi^* \in \mathcal{A}^* \mapsto \texttt{bra} \text{ vectors } \langle\psi \mid \in \mathcal{H}_d.}$

$\boxed{\text{Scalar product } \langle A, B\rangle \mapsto \mathrm{tr}(A^*B).}$

$\boxed{\text{Effect } A \in \mathcal{E}(\mathcal{A}) \mapsto \text{Positive operator } A \text{ bounded by } \mathbb{I}.}$

$\boxed{\text{Proposition } P \in \mathcal{E}(\mathcal{A}) \mapsto \text{Projector } P.}$

$\boxed{\text{States} \mapsto \text{density matrices, positive and normalized self-adjoint operators.}}$

$\boxed{\text{State space} \mapsto \text{the set of density matrices.}}$

$\boxed{\text{When } d = 2 \text{ the state space } \mathcal{S}(\mathcal{H}_2) \mapsto \text{the Bloch ball.}}$

$\boxed{\text{Extreme points of } \mathcal{S}(\mathcal{H}_2) \mapsto \text{pure states, points on the Bloch sphere.}}$

The Hilbert space of dimension $d$, is a bounded algebra of $d \times d$ complex matrices, $\mathbb{B}_{\mathcal{H}_d}$. The positivity of the operator $\rho$ is a consequence of the positivity of the functional $\rho$:

$$\rho(\mid \psi\rangle\langle\psi \mid) \geq 0 \Leftrightarrow \langle\psi, \rho(\psi)\rangle \geq 0.$$

The normalization condition $\rho(\mathbb{I}) = 1$ translates into $\mathrm{tr}(\rho) = 1$

[74] The extreme point of a convex set $S$ is a point in the set $S$ which does not lie in any open line segment joining two points of $S$.

$$\rho(\mathbb{I}) = 1 \Leftrightarrow \mathrm{tr}(\rho) = 1.$$

Consider now the dual space $\mathbb{B}^*_{\mathcal{H}_d}$ and call $\mathbb{I}^\perp \in \mathbb{B}^*_{\mathcal{H}_d}$ the dual of the unit operator $\mathbb{I}$; let $\sigma_1, \sigma_2, \ldots, \sigma_{d^2-1}$ be an orthonormal basis in $\mathbb{I}^\perp$. Then each self-adjoint operator $\rho$, $\mathrm{tr}(\rho) = 1$, can be expressed as:

$$\rho = \frac{\mathbb{I}}{d} + \frac{1}{2}\sum_{i=1}^{d^2-1} x_i\sigma_i =: \frac{\mathbb{I}}{d} + \frac{1}{2}\bar{x}\bar{\sigma}, \quad \text{with} \quad \bar{x} \in \mathbb{R}^{d^2-1}.$$

This expression has a geometric interpretation. When $d = 2$ we chose the Pauli matrices as a basis, $\bar{\sigma} = (\sigma_I, \sigma_x, \sigma_y, \sigma_z)$, and the state space $\mathcal{S}(\mathcal{H}_2)$ coincides with the Bloch ball $\bar{x} \in \mathbb{R}^3, \mid \bar{x} \mid \leq 1$ because the positivity of $\rho$ implies $\mid \bar{x} \mid \leq 1$, in fact

$$\rho \geq 0 \Leftrightarrow \mid \bar{x} \mid \leq 1.$$

The pure states, the extremes of $\mathcal{S}(\mathcal{H}_2)$, are the states on the Bloch sphere $\bar{x} = 1$.

We now consider a finite set of values $x \in X$ and an observable $E$; the effect $E_x \in \mathcal{E}(\mathcal{A}$ means that $x$ is the outcome of observing $E$ in algebra $\mathcal{A}$, it is true if we observe the value $x$ and false otherwise. If the system is in state $\rho$ then $p_x = \rho(E_x)$ is the probability to obtain $x$ as the outcome of a measurement.

Given the observable algebra $\mathcal{A}$ a <u>positive operator valued measure</u> (POV) on the finite set $X$ is a family of effects

$$E = E_x : x \in X,\ 0 \leq E_x \leq \mathbb{I}$$

if and only if $\sum_{x \in X} E_x = \mathbb{I}$ holds.

We are now concerned with composite systems and their representation in an algebra of observables. Consider two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ and vectors $\mid \psi_1\rangle, \mid \varphi_1\rangle \in \mathcal{H}$ and $\mid \psi_2\rangle, \mid \varphi_2\rangle \in \mathcal{K}$. The bilinear form $\mid \psi_1\rangle \otimes \mid \psi_2\rangle$ called the tensor product is defined as: $\mid \psi_1\rangle \otimes \mid \psi_2\rangle(\varphi_1, \varphi_2) = \langle \psi_1 \mid \varphi_1)\rangle\langle \psi_2) \mid \varphi_2\rangle$. To describe composite systems we introduce the tensor product of two bounded operators $A_1 \in \mathbb{B}_{\mathcal{H}}$ and $A_2 \in \mathbb{B}_{\mathcal{K}}$ applied to the tensor product $\mid \psi_1\rangle \otimes \mid \psi_2\rangle$:

$$A_1 \otimes A_2(\mid \psi_1\rangle \otimes \mid \psi_2\rangle) = (A_1 \mid \psi_1\rangle) \otimes (A_2 \mid \psi_2\rangle).$$

The space $\mathbb{B}_{\mathcal{H}\otimes\mathcal{K}}$ is the span of all tensor products $A_1 \otimes A_2$. The tensor product can be extended to more than two systems.

The <u>partial trace</u> of $\rho \in \mathbb{B}_{\mathcal{H}\otimes\mathcal{K}}$ is:

$$\mathrm{tr}_{\mathcal{H}}\left[\mathrm{tr}_{\mathcal{K}}(\rho)A\right] = \mathrm{tr}_{\mathcal{H}\otimes\mathcal{K}}\left(\rho A \otimes \mathbb{I}\right), \quad \forall A \in \mathbb{B}_{\mathcal{H}}.$$

**Proposition.** *For each element* $\mid \Psi\rangle$ *of the tensor product* $\mathcal{H}\otimes\mathcal{K}$ *there exist two orthonormal sets of vectors* $\mid \psi_j\rangle \in \mathcal{H},\ 1 \leq j \leq n$ *and* $\mid \varphi_k\rangle \in \mathcal{K},\ 1 \leq k \leq n$ *such that*

$$\mid \Psi\rangle = \sum_j \sqrt{\lambda_j} \mid \psi_j\rangle \otimes \mid \varphi_j\rangle$$

*with* $\sqrt{\lambda_j}$ *called the Schmidt coefficients. We do not require either set of vectors to be a basis, in other words* $n \leq \dim(\mathcal{H})$ *and* $n \leq \dim(\mathcal{K})$.

The following proposition is a consequence of the Schmidt decomposition.

**Proposition.** *Each state* $\rho \in \mathbb{B}^*_{\mathcal{H}}$ *can be extended to a pure state* $\mid \Psi\rangle$ *on a larger system with the Hilbert space* $\mathcal{H}\otimes\mathcal{K}$ *such that* $\rho = \mid \Psi\rangle\langle\Psi \mid$.

In Section 1.15 we observed that quantum information processing requires several types of procedures including: free-time quantum evolution, controlled-time evolution, preparation, and measurement; the process of encoding classical information into a quantum system involves a preparation step and the process of retrieving classical information from a quantum system requires a measurement. The mathematical description of each of these steps requires the concept of *channel* or map.

A <u>channel</u> converts input systems characterized by the input observable algebra $\mathcal{A}$ into output systems characterized by an output observable algebra $\mathcal{B}$; sometimes $\mathcal{A}$ and $\mathcal{B}$ coincide. A channel $T$ is thus a mapping: $T : \mathcal{A} \mapsto \mathcal{B}$ with several properties:

- $T$ maps effects to effects thus it must be positive, $T(A) \geq 0,\ \forall A \geq 0$ and bounded, $T(\mathbb{I}) \leq \mathbb{I}$.
- Given two channels $T_1 : \mathcal{A}_1 \mapsto \mathcal{B}_1$ and $T_2 : \mathcal{A}_2 \mapsto \mathcal{B}_2$ then the channel $T = T_1 \otimes T_2$ maps composite systems of the type $\mathcal{A}_1 \otimes \mathcal{A}_2$ into $\mathcal{B}_1 \otimes \mathcal{B}_2$. The channel $T$ must be positive as well.

The *dual channel/map* $T^*$ is defined for all states $\rho \in \mathcal{B}^*$ and all operators $A \in \mathcal{A}$:

$$T^* : \mathcal{B}^* \mapsto \mathcal{A}^* \Rightarrow T^*\rho(A) = \rho(TA).$$

There are two ways to describe a channel:

- *The Heisenberg picture:* we invoke first the channel $T : \mathcal{A} \mapsto \mathcal{B}$ and then measure the effect $A$ on the output systems. Thus $T$ measures the set of effects on the input systems:

$$T : \mathcal{E}(\mathcal{B}) \mapsto \mathcal{E}(\mathcal{A}).$$

- *The Schrödinger picture:* we consider the states and interpret the channel as a transformation of $\mathcal{A}$ systems in state $\rho \in \mathcal{S}(\mathcal{A})$ into $\mathcal{B}$ systems in state $T^*(\rho)$:

$$T^* : \mathcal{S}(\mathcal{A}) \mapsto \mathcal{S}(\mathcal{B}).$$

The dual map $T^*$ is called the Scrödinger representation of the channel $T$.

Let $T$ be a liner map (channel) from an observable algebra $\mathcal{A}$ to another $\mathcal{B}$.

- $T : \mathcal{A} \mapsto \mathcal{B}$ is positive if $T(A) \geq 0,\ \forall A \in \mathcal{A},\ A \geq 0$.
- $T : \mathcal{A} \mapsto \mathcal{B}$ is completely positive if:

$$T \otimes \mathbb{I}_{\mathbb{B}} : \mathcal{A} \otimes \mathbb{B}(\mathbb{C}^n) \mapsto \mathbb{B}_{\mathcal{H}} \otimes \mathbb{B}(\mathbb{C}^n),\ \forall n \in \mathbb{N}$$

  with $\mathbb{I}_{\mathbb{B}}$ the identity element in $\mathbb{B}(\mathbb{C}^n)$. The dual map $T^*$ is complectly positive if $T$ is.
- $T : \mathcal{A} \mapsto \mathcal{B}$ is unital if $T(\mathbb{I}) = \mathbb{I}$

**Theorem.** *(Stinespring Dilation Theorem) Every completely positive map:*

$$T : \mathcal{A} \mapsto \mathcal{B} : \quad A \in \mathcal{A}, \mathcal{A} = \mathbb{B}_{H_1}, \dim(\mathcal{H}_1) = d_1, \quad B \in \mathcal{B}, \mathcal{B} = \mathbb{B}_{H_2}, \dim(\mathcal{H}_2) = d_2$$

*has a unique decomposition of the form:*

$$T(A) = V^* \left(A \otimes \mathbb{I}_{\mathcal{K}}\right) V$$

*where $\mathcal{K}$ is an additional Hilbert space and $V$ is an operator:*

$$V : \mathcal{H}_2 \mapsto \mathcal{H}_1 \otimes \mathcal{K}.$$

*The Hilbert space $\mathcal{K}$ and the $V$ are chosen such that $\forall \mid \varphi\rangle \in \mathcal{H}_2$ the span of $\left(A \otimes \mathbb{I}_{\mathcal{K}}\right) V \mid \varphi\rangle$ is dense in $\mathcal{H}_1 \otimes \mathcal{K}.T$ The minimal $\mathcal{K}$ satisfies the condition:*

$$\dim(\mathcal{K}) \leq d_1^2 d_2.$$

The Stinespring Dilation Theorem is a critical result for Quantum Information Theory and Quantum Error Correction as we have seen when we discussed the role of measurements carried out on ancilla qubits in Chapters 3 and 5.

We conclude this presentation with a theorem regarding the purification of mixed states; the result is based upon the fact that a $d \times d$ matrix represents a bilinear form as well as an operator on a vector space $V$. The proof of the theorem can be found in [231].

**Theorem.** *Given a bipartite system in $\mathcal{H} \otimes \mathcal{H}_1$ for every state $\rho \in \mathcal{H} \otimes \mathcal{H}_1$ there exists:*

- *A Hilbert space $\mathcal{K}$;*
- *A pure state $\sigma \in \mathcal{H} \otimes \mathcal{K}$; and*
- *A channel $T : \mathbb{B}_{H_1} \mapsto \mathbb{B}_{\mathcal{K}}$ such that:*

$$\rho = \left(\mathbb{I}_{\mathcal{B}^*} \otimes T^*\right) \sigma$$

  *with $\mathbb{I}_{\mathcal{B}^*}$ the identity element in $\mathcal{B}^*_{\mathcal{H}}$.*

The pure state $\sigma$ is a purification of the state $\rho$ and can be chosen so that $\mathrm{tr}_{\mathcal{K}}(\sigma)$ does not have zero eigenvalues and in this case $T$ and $\sigma$ are unique (up to a unitary equivalence).

# References


[1] S. Aaronson and D. Gottesman, "Improved Simulation of Stabilizer Codes.", *Phys. Rev. A*, 70:052328 - 052342, 2004.

[2] E. S. Abers. *Quantum Mechanics.* Prentice Hall, Upper Saddle River, NJ, 2003.

[3] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Loyd, and O. Regev. "Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation." Preprint, arxiv.org/quant-ph/0405098 v2, 2005.

[4] Y. Aharonov and D. Bohm. "Significance of Electromagnetic Potentials in the Quantum Theory." *Phys. Rev.* 115(3):485 - 490, 1959.

[5] D. Aharonov and M. Nen-Or. "Fault-Tolerant Quantum Computation with Constant Error Rate." Preprint, arxiv.org/quant-ph/9906129 v1, 1999.

[6] N. I. Akhiezer abd I. M. Glazman. *Theory of Linear Operators in Hilbert Spaces, vol 2.* Ungar, New York, N.Y., 1963.

[7] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki. "Dynamical Description of Quantum Computing: Generic Nonlocality of Quantum Noise." *Phys. Rev. A*, 65:062101, 2002.

[8] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Accuracy Threshold for Concatenated Distance-3 Codes. *Quantum Information and Computation* 6:97 - 165, 2006.

[9] P. Aliferis and A. W. Cross. "Subsystem Fault Tolerance with the Bacon-Shor Code." *Phys. Rev. Lett.*, 98:220502, 2007.

[10] D. Z. Albert. *Quantum Mechanics and Experience.* Harvard University Press, Cambridge, MA, 1992.

[11] A. G. Anderson, R. Garwin, E. L. Hahn, J. W. Horton, and G. L. Tucker. "Spin Echo Serial Storage Memory." *J. of Applied Physics*, 26:1324, 1955.

[12] A. G. Anderson and E. L. Hahn. "Spin Echo Storage Technique." *U. S. Patent # 2,174,714, 1955.*

[13] D. Arovas, J. R. Schrieffer, and F. Wilczek. "Fractional Statistics and the Quantum Hall Effect." *Phys. Rev. Lett.,* 53(7):722 - 724, 1984.

[14] V. I. Arnold. *Mathematical Methods of Classical Mechanics. Springer Verlag*, ISBN 978-0387968902, 1997.

[15] R. B. Ash. *Information Theory.* Dover Publishing House, New York, NY, 1965.

[16] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter. "Elementary Gates for Quantum Computation." *Phys. Rev. A,* 52:3457 - 3570, 1995.

[17] D. Bacon. "Operator Quantum Error Correcting Subsystems for Self-Correcting Quantum Memories." *Phys. Rev. A*, 73:012340, 2006.

[18] D. Bacon and A. Casaccino. "Quantum Error Subsystem Codes from Two Classical Linear Codes." Preprint, arxiv.org/quant-ph/0610088 v2, 2006.

[19] A. Barenco, D. Deutsch, and A. Ekert. "Conditional Quantum Dynamics and Logic Gates." *Phys. Rev. Lett.*, 74(20):4083 - 4086, 1995.

[20] M. D. Barrett, J. Chiaverini, T. Schaetz, J. Britton, W. M.Itano, J. D. Jost, E. Knill, C. Langer, D. Leibfried, R. Ozeri, and D. J. Wineland. "Deterministic Quantum Teleportation of Atomic Qubits" *Nature*, 429:737 - 739, 2004.

[21] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. "Noncommuting States Cannot Be Broadcast." *Phys. Rev. Lett.*, 76(15):2818 - 2821, 1996.

[22] H. Barnum, M. A. Nielsen, and B. Schumacher. "Information Transmission Through a Noisy Quantum Channel." *Phys. Rev. A*, 57(6):4153 - 4175, 1998.

[23] J. D. Barrow, P. C. W. Davis, C. L. Harper (Eds). *Science and Ultimate Reality. Quantum Theory, Cosmology and Complexity.* Cambridge University Press, 2004.

[24] S. D. Bartlett and H. M. Wiseman. "Entanglement Constrained by Superselection Rules." Preprint, arxiv.org/quant-ph/0303140, 2003.

[25] J. S. Bell. "On the Einstein-Podolsky-Rosen Paradox." *Physics*, 1:195 - 200, 1964.

[26] J. S. Bell. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy.* Cambridge University Press, Cambridge, 1987.

[27] P. Benioff. "The Computer as a Physical System: A Microscopic Quantum Mechanical Hamiltonian Model of Computers as Represented by Turing Machines." *J. of Statistical Physics*, 22:563 - 591, 1980.

[28] P. Benioff. "Quantum Mechanical Models of Turing Machines that Dissipate no Energy." *Phys. Rev. Lett.*, 48:1581 - 1584, 1982.

[29] P. Benioff. "Quantum Mechanical Models of Turing Machines." *J. of Statistical Physics*, 29:515 - 546, 1982.

[30] C. H. Bennett. "Logical Reversibility of Computation." *IBM J. of Research and Development*, 17:525 - 535, 1973.

[31] C. H. Bennett. "The Thermodynamics of Computation - A Review." *Int. J. of Theoretical Physics*, 21:905 - 928, 1982.

[32] C. H. Bennett and G. Brassard. "Quantum Cryptography: Public Key Distribution and Coin Tossing." *Proc. IEEE Conf. on Computers, Systems, and Signal Processing, IEEE Press,* 175 - 179, 1984.

[33] C. H. Bennett and S. J. Wiesner. "Communication via One- and Two-particle Operators on Einstein-Podolsky-Rosen States." *Phys. Rev. Let.,* 69:2881 - 2884, 1992.

[34] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. "Teleporting an Unknown State via Dual Classical and Einstein-Podolsky-Rosen Channels." *Phys. Rev. Lett.*, 70(13):1895 - 1899, 1993.

[35] C. H. Bennett, G. Brassard, R. Jozsa, D. Mayers, A. Peres, B. Schumacher, and W. K. Wootters. "Reduction of Quantum Information Entropy by Reversible Extraction of Classical Inforamtion." *J. of Modern Optics*, 41(12):2307 - 2314, 1994.

[36] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. "Concentrating Partial Entanglement by Local Operations." Preprint, arxiv.org/quant-ph/9511030, 1994.

[37] C. H. Bennett. "Quantum Information and Computation." *Physics Today*, 24 - 30, October 1995.

[38] C. H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W. K. Wootters. "Mixed State Entanglement and Quantum Error Correction." *Phys. Rev. Lett. A*, 54:3824 - 3851, 1996.

[39] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels." *Phys. Rev. Lett.*, 76:722 - 725, 1996.

[40] C. H. Bennett, D. DiVincenzo, J. A. Smolin, and W. K. Wootters. "Mixed States Entanglement and Quantum Error Correction." *Phys. Rev. A*, 54:3824 - 3851, 1996.

[41] C. H. Bennett, C. A. Fuchs, and J. A. Smolin. "Entanglement-enhanced Classical Communication on a Noisy Quantum Channel." Preprint, arxiv.org/quant-ph/9611006, 1996.

[42] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. "Strengths and Weaknesses of Quantum Computation." *SIAM Journal on Computing,* 26:1510 - 1523, 1997.

[43] C. H. Bennett, D. DiVincenzo, and J. A. Smolin. "Capacities of Quantum Erasure Channels." *Phys. Rev. Lett.* 78(16):3217 - 3220, 1997.

[44] C. H. Bennett and P.W. Shor. "Quantum Information Theory." *IEEE Trans. on Information Theory,* 44(6):2724 - 2742, 1998.

[45] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. "Entanglement-Assisted Classical Capacity of Noisy Quantum Channels." Preprint, http://arxiv.org/quant-ph/9904023, August, 1999.

[46] C. H. Bennett and D. P. DiVicenzo. "Quantum Information and Computation." *Science.* 404:247 - 255, 2000.

[47] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal. "Entanglement-Assisted Capacity of a Quantum Channel and the Reverse Shannon Theorem." *IEEE Trans. on Information Theory*, 48(10):2637 - 2655, 2002.

[48] C. H. Bennett, T. Mor, and J. A. Smolin. "The Parity Bit in Quantum Cryptography." Preprint, http://arxiv.org/quant-ph/9604040, 2002.

[49] C. H. Bennett. "Quantum Information Processing." *Computer Science: Reflections on the Field, Reflections from the Field.* National Research Council. Committee on the Fundamentals of Computer Scienc: Challenges and Opportunities, pp. 51 - 56. National Academies Press, 2004.

[50] E. Berlekamp. *Algebraic Coding Theory.* McGraw-Hill, New York, N.Y., 1968.

[51] E. Bernstein and U. Vazirani. "Quantum Complexity Theory." *SIAM J. of Computing*, 26:1411 - 1473, 1997.

[52] C. Berrou, A. Glavieux, and P. Thitimajshima. "Near Shannon Limit Error-correcting Codes and Decoding: Turbocodes." *Proc. IEEE Int. Conf. on Communication*, 1064-1070, 1993.

[53] M. V. Berry. "Quantal Phase Factors Accompanying Adiabatic Changes." *Proc. Royal Society London,* A392:45 - 57, 1984.

[54] A. Berthiaume and G. Brassard. "The Quantum Challenge to Structural Complexity Theory." *Proc. 7-th Annual Conf. on Structure in Complexity Theory*, 132 - 137, IEEE Press, Los Alamitos, CA, 1992.

[55] A. Berthiaume and G. Brassard. "Oracle Quantum Computing." *Proc. Workshop on Physics of Computation*, 195 - 199, IEEE Press, Los Alamitos, CA, 1992.

[56] I. Bialynicki, J. Mycielski, and J. H. Eberly. "A Reply to the Comment by David Farrelly, Ernestine Lee, and T. Uzer". *Comm. in Math. Phys.*, 44:129, 1975.

[57] G. Birkhoff and J. von Neumann. "The Logic of Quantum Mechanics." *Annals of Mathematics*, 37(4):823 - 843, 1936.

[58] G. Birkhoff and S. Mac Lane. *A Survey of Modern Algebra.* Macmillan Publishing, New York, NY, 1965.

[59] D. Biron, O. Biham, E. Biham, M. Grassl, D. A. Lidar. "Generalized Grover Search Algorithm for Arbitrary Initial Amplitude Distribution." Preprint, arxiv.org/quant-ph9801066, 1998.

[60] E. Biham, O. Biham, D. Biron, M. Grassl, D. A. Lidar, and D. Shapira. "Analyis of Generalized Grover Search Algorithms Using Recursion Equations." *Phys. Rev. A*, 63:012310, 2001.

[61] R. B. Blakestad, C. Ospelkaus, A. P. VanDevender, J. M. Amini, J. Britton, D. Leibfried, and D. J. Wineland. "High-Fidelity Transport of Trapped-Ion Qubits through an X-junction Trap Array." *Preprint, arxiv.org/quant-ph/0901.0533*, 2009.

[62] B. B. Blinov, L. Deslauriers, P. Lee, M. J. Masden, R. Miller, and C. Monroe. "Sympathetic Cooling of $Cd^+$ Isotopes." *Phys. Rev. A*, 65(4):040304, 2002.

[63] F. Bloch. "Nuclear Induction." *Phys. Rev.*, 70:460, 1946.

[64] D. Bohm. *Quantum Theory.* Prentice Hall, Upper Saddle, NJ, 1951.

[65] G. Boole. *An Investigation of the Laws of Thought.* Dover Publications, ISBN-13: 978-0486600284, New Edition, 1958.

[66] M. Born and V. A. Fock. "Beweis des Adiabatensatzes." *Zeitschrift für Physik A, Hadrons and Nuclei,* 51(3-4):165 - 180, 1928.

[67] M. Born. "The Statistical Interpretations of Quantum Mechanics." *Nobel Lectures, Physics 1942 - 1962, 256 - 267. December 11, 1954.* Also http://nobelprize.org/nobel_prizes/physics/ laureates/1954/born-lecture.pdf.

[68] R. C. Bose and R. K. Ray-Chaudhuri. "On a Class of Error-correcting Binary Codes." *Information and Control,* 3:68 - 79, 1960.

[69] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. "Experimental Quantum Teleportation." *Nature,* 390:575 - 579, 1997.

[70] D. Bouwmeester, A. Ekert, and A. Zeilinger, Editors. *The Physics of Quantum Information*, Springer Verlag, Heidelberg, 2003.

[71] H. E. Brandt. "Positive Operator Valued Measure in Quantum Information Processing." *American J. of Physics*, 5:434 - 479, 1998.

[72] H. E. Brandt. "Positive Operator Valued Measure in Quantum Measurement with a Postive Operator-valued Measure." *J. of Optics B: Quantum and Semiclassic Optics.*, 5:S266 - S270, 2003.

[73] G. Brassard. "Searching a Quantum Phone Book." *Science*, 275:627 - 628, 1997.

[74] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp. "Quantum Amplitude Amplification and Estimation." Preprint, arxiv.org/quanth-ph/0005055 v1, 2000.

[75] S. Bravyi. "Universal Quantum Computation with the $\nu = 5/2$ Fractional Quantum Hall State." *Phys. Rev. A,* 73(4):042313, 2006.

[76] G. K. Brennen, C. M. Caves, P. S. Jessen, and I. H. Deutsch. "Quantum Logic Gates in Optical Lattices." *Phys. Rev. Lett.*, 82(5):1060 - 1063, 1999.

[77] L. de Broglie. "The Wave Nature of the Electron." *Nobel Lectures, Physics 1922 - 1941*, 244 - 256. December 12, 1929. Also: http://nobelprize.org/nobel_prizes/physics/laureates/1929/broglie-lecture.pdf.

[78] J. Brown. *The Quest for the Quantum Computer.* Simon and Schuster, New York, NY, 1999.

[79] G. Burkard, H. A. Engel, and D. Loss. "Spintronics and Quantum Dots for Quantum Computing and Quantum Communication." *Fortschritte der Physik*, 48(9-11):965 - 986, 2000.

[80] A. W. Burks, H. H. Goldstine, and J. von Neumann. "Preliminary Discussion of the Logical Design of an Electronic Computer Instrument." *Report to the US Army Ordnance Department*, 1946. Also in: *Papers of John von Neumann.* W. Asprey and A. W. Burks, Editors, 97 - 146, MIT Press, Cambridge, MA, 1987.

[81] A. R. Calderbank and P. W. Shor. "Good Quantum Error-Correcting Codes Exist." *Phys. Rev. A*, 54(42):1098 - 1105, 1996.

[82] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloan. "Quantum Error Correction and Orthogonal Geometry." *Phys. Rev. Lett.*, 78:405-408, 1997.

[83] F. E. Camino, W. Zhou, and V. J. Goldman. "Aharonov-Bohm Electron Interferometer in the Integer Quantum Hall Regime." *Phys. Rev. B,* 72:155313, 2005.

[84] F. E. Camino, W. Zhou, and V. J. Goldman. "Aharonov-Bohm Superperiod in a Laughlin Quasiparticle Interferometer." *Phys. Rev. Lett.,* 95:246802, 2005.

[85] F. E. Camino, W. Zhou, and V. J. Goldman. "Experimental Realization of a Primary-Filling $e/3$ Quasiparticle Interferometer." *http://arxiv.org/ftp/cond-mat/papers/0611/0611443.pdf*, 2006.

[86] F. E. Camino, W. Zhou, and V. J. Goldman. "Quantum Transport in Electron Fabry-Pérot Interferometers." *Phys. Rev. B,* 76:155305, 2007.

[87] F. E. Camino, W. Zhou, and V. J. Goldman. "Experimental Realization of Laughlin Quasiparticle Interferometers." *Physica E,* 40(5):949 - 953, 2008. *Proc. Conf. Electronic Properties of 2D Systems (EP2DS-17)*, Genoa, Italy, 2007.

[88] C. de C. Chamon, D. E. Freed, S. A. Kivelson, S. L Sondhi, and X. G. Wen. "Two Point-Contact Interferometer for Quantum Hall Systems." *Phys. Rev. B*, 55(4):2331 - 2343, 1997.

[89] N. J. Cerf and C. Adami. "Accesible Information in Quantum Mesaurements." Preprint, http://arxiv.org/abs/quant-ph/9611032, 1996.

[90] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. B. Blakestad, J. Britton, W. M.Itano, J. D. Jost, E. Knill, C. Langer, R. Ozeri, and D. J. Wineland. "Realization of Quantum Error Correction." *Nature*, 432:602 - 605, 2004.

[91] A. M. Childs and I. L. Chuang. "Universal Quantum Computation with Two-Level Trapped Ions." *Phys. Rev. A*, 63(1):012306, 2001.

[92] I. L. Chuang and Y. Yamamoto. " Quantum Bit Regeneration." *Phys. Rev. Lett.*, 76:4281 - 4284, 1996.

[93] I. L. Chuang and M. A. Nielsen. "Prescription for experimental Determination of the Dynamics of a Quantum Black Box." *J. of Modern Optics*, 44(11/12):2455 - 2467, 1997.

[94] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd. "Experimental Realization of a Quantum Algorithm." *Nature*, 393(6681):143 - 146, 1998.

[95] I. L. Chuang "IBM's Test-Tube Quantum Computer Makes History." *http://domino.research.ibm.com/comm/pr.nsf/pages/news.20011219_quantum.html*

[96] A. Church. "A Note on the Entscheidungsproblem." *J. of Symbolic Logic,* 1:40-41, 1936

[97] J. I. Cirac and P. Zoller. "Quantum Computation with Cold Trapped Ions." *Phys. Rev. Lett.*, 74(20):4091 - 4094, 1995.

[98] J. I. Cirac, L. Duan, and P. Zoller. "Quantum Optical Implementation of Quantum Information Processing." Preprint, arxiv.org/quanth-ph/0405030 v1,2004. See also *Experimental Quantum Computation and Information*, Ed. P.De Martini and C. Monroe. IOS Press, Amsterdam, 2002.

[99] J. I. Cirac and P. Zoller. "A Scalable Quantum Computer with Ions in an Array of Microtraps." *Nature*, 404:579 - 581, 2000.

[100] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. "Proposed Experiment to Test Local Hidden-Variable Theory." *Phys. Rev. Lett.*, 23:880, 1969.

[101] J. P. Clemens, S. Siddiqui, and J. Gea-Banacloche. "Quantum Error Correction Against Correlated Noise." *Phys. Rev. A*, 69:062313, 2004.

[102] R. Cleve, A. Ekert, L. Henderson, C. Macchiavello, and M. Mosca. "On Quantum Algorithms." Preprint, arxiv.org/quant-ph/9903061 v1, 1999.

[103] G. Collins. "Computing with Quantum Knots." *Scientific American,* 57 -63, April 2005.

[104] D. G. Cory, A. F. Fahmy, and T. F. Havel. "Nuclear Magnetic Resonance Spectroscopy: An Experimentally Accessible Paradigm for Quantum Computing." *Proc. PhysComp96*, T. Toffoli, M. Biafore, and J. Leao, Editors. New England Complex Systems Institute, 87 - 91, 1996.

[105] D. G. Cory, A. F. Fahmy, and T. F. Havel. "Ensemble Quantum Computing by NMR Spectroscopy." *Proc. National Academy of Science*, 94(5):1634 - 1639, 1997.

[106] D. G. Cory, M. D. Price, and T. F. Havel. "Nuclear Magnetic Resonance Spectroscopy: An Experimentally Accessible Paradigm for Quantum Computing." *Physica D*, 120:82 - 101, 1998.

[107] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, New York, NY, 1991.

[108] A. W. Cross, D. P. DiVincenzo, and B. M. Terhal, "A Comparative Study for Quantum Fault Tolerance.", Preprint, arxiv.org/quant-ph/07111556, 2007.

[109] W. van Dam. "A Universal Quantum Cellular Automaton." *Proc., Worshop on Physics and Computation (PhysComp96), T. Toffoli, M. Biafore, and J. Leao, Editors.* New England Complex Systems Institute, 323 - 331, 1996.

[110] K. Davidson. *$C^*$- Algebras by Example.* Fields Institute Monographs, American Math. Society, Providence, 1996.

[111] E. B. Davies. "Information and Quantum Measurement," *IEEE Trans. on Information Theory,* vol. IT-24(5):596 - 599, 1978.

[112] C. M. Dawson and M. A. Nielsen. "The Solovay-Kitaev Algorithm." Preprint, arxiv.org/quanth-ph/0505030, 2005.

[113] B. DeMarco, A. Ben-Kish, D. Leibfried, V. Meyer, M. Rowe, B. M. Jelencović, W. M. Itano, J. Britton, C. Langer, T. Rosemband, and D. J. Wineland. "Experimantal Demonstration of a Controlled-NOT Wave-Packet Gate." *Phys. Rev. Lett.,* 89(26):267901, 2002.

[114] R. de-Picciotto, M. Reznikov, M. Heiblum, V. Umansky, G. Bunin, and D. Mahalu. "Direct Observation of a Fractional Charge." *Nature*, 389:162 - 164, 1997.

[115] D. Deutsch. "Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer." *Proc. of the Royal Society London A*, 400:97 - 117, 1985.

[116] D. Deutsch. "Quantum Computational Networks." *Proc. of the Royal Society London A*, 425:73 - 90, 1989.

[117] D. Deutsch and R. Jozsa. "Rapid Solution of Problems by Quantum Computations." *Proc. of the Royal Society London A*, 439:553 - 558, 1992.

[118] D. Deutsch, A. Barenco, and A. Ekert. "Universality in Quantum Computation." *Proc. of the Royal Society London A*, 449:669 - 677, 1995.

[119] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera. "Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels." *Phys. Rev. Lett.*, 80:2818 - 2821, 1996. "Erratum: Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels." *Phys. Rev. Lett.*, 80:2022, 1998.

[120] S. J. Devitt, A. D. Greentree, R. Ionicioiu, J. L.Q'Brien, W. J. Munro, and L. C. L. Hollenberg. "The Photonic Module: an On-demand Resource fro Photonic Entanglement." *Phys. Rev. A,* 76: 052312, 2007.

[121] R. G. DeVoe. "Elliptical Ion Traps and Trap Arrays for Quantum Computation." *Phys. Rev. A,* 58(2):910 - 914, 1998.

[122] D. Dhar, L. L. Grover, and S. M. Roy. "Preserving Quantum States Using Inverting Pulses: A Super-Zeno Effect." *Phys. Rev. Lett.*, 96(10):100405, 2006.

[123] F. Diedrich, J. C. Bergquist, W. M. Itano, and D. J. Wineland. "Laser Cooling to the Zero-Point Energy of Motion." *Phys. Rev. Lett.,* 62(4):403 - 406, 1989.

[124] I. Dimov, B. I. Halperin, and C. Nayak. "Spin Order in Paired Quantum Hall States." Preprint, http://arxiv.org/PS_cache/arxiv/pdf/0710/0710.1921v1.pdf, 2007.

[125] P. A. M. Dirac. "Theory of Electrons and Positrons." *Nobel Lecture*, December 12, 1933. Also, http://nobelprize.org/nobel_prizes/physics/laureates/1933/dirac-lecture.pdf.

[126] P. A. M. Dirac. *The Principles of Quantum Mechanics.* Fourth Edition. Sec. 2, 4 - 7, Clarendon Press, Oxford, 1967.

[127] D. P. DiVincenzo. "Quantum Computation." *Science*, 270:255 - 261, 1995.

[128] D. P. DiVincenzo. "Two-bit Gates are Universal for Quantum Computation." *Phys. Rev. A*, 51:1015 - 1022, 1995.

[129] D. P. DiVincenzo and P. W. Shor. "Fault-Tolerant Error Correction with Efficient Quantum Codes." *Phys. Rev. Lett.*, 77:3260 - 3263, 1996.

[130] D. P. DiVincenzo. "Quantum Gates and Circuits." *Phil. Trans. Royal Soc. London A*, 454:261 - 276 1998, Also, Proc.: Mathematical, Physical and Engineering Sciences, Vol. 454, No. 1969, Quantum Coherence and Decoherence (January 1998), pp. 261-276 Also, Preprint, arxiv.org/quanth-ph/9705009, May, 1997.

[131] D. P. DiVincenzo. "The Physical Implementation of Quantum Computation." *Fortschritte der Physik*, 48(9-11):771 - 783, 2000.

[132] M. Dolev, M. Heilblum, V. Umansky, A. Stern, and D. Mahalu. "Towards Identification of a Non-Abelian State: Observation of a Quarter of Electron Charge at $\nu = 5/2$ Quantum State." Preprint, http://arxiv.org/ftp/arxiv/papers/0802/0802.0930.pdf, 2008.

[133] A. Einstein. "Elektrodynamik bewegter Körper." *Annalen der Physik*, 891:921, 1905.

[134] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Phys. Rev.*, 47:777, 1935.

[135] A. Einstein. *Autobiographycal Notes* in *Albert Einstein, Philosopher-Scientist* (P.A. Schilpp Editor), Libraray of Living Philosophers, 1949.

[136] J. P. Eisenstein, K. B. Cooper, L. N. Pfeiffer, and K. W. West. "Insulating and Fractional Quantum Hall States in the First Excited Landau Level." *Phys. Rev. Lett.,* 88(7):076801, 2002.

[137] A. K. Ekert, B. Huttner, G. M. Palma, and A. Peres. "Eavesdropping on Quantum-Cryptographical Systems." *Phys. Rev. A*, 50:1047 - 1057, 1996.

[138] A. K. Ekert and C. Machiavlello. "Quantum Error-Correction for Communication." *Phys. Rev. Lett.*, 77:2585 - 2588, 1996.

[139] A. K. Ekert and R. Jozsa. "Quantum Algorithms: Entanglement Enhanced Information Processing." *Proc. Royal Society London A*, 356(1743):1769 - 1782, 1998.

[140] A. K. Ekert, P. Hayden, and H. Inamori. "Basic Concepts in Quantum Computing." Preprint, arxiv.org/quant-ph/0011013 v1, 2000.

[141] P. Elias. "Coding for Noisy Channels." *IRE. Conv. Rec.* 4:37-46, 1955.

[142] J. M. Elzerman, R. Hanson, J. S. Greidanus, L. H. Willems van Beveren, S. De Franceschi, L. M. K. Vandersypen, S. Taucha, and L. P. Kouwenhoven. "Tunable Few-Electron Double Quantum Dots with Integrated Charge Read-Out." *Physica E*, 25:135 - 141, 2004.

[143] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven. "Single-Shot Read-Out of an Individual Electron Spin in a Quantum Dot." *Nature*, 430:431 - 435, 2004.

[144] R. M. Fano. *Transmission of Information: A Statistical Theory of Communication.* Wiley, New York, 1961.

[145] R. M. Fano. "A Heuristic Discussion of Probabilistic Decoding." *IEEE. Trans. Information Theory*, IT-9:64 - 74, 1963.

[146] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. "Quantum Computation by Adiabatic Evolution." Preprint, arxiv.org/quant-ph/0001106v1 v1, 2000.

[147] A. Feinstein. *Foundation of Information Theory.* Mc.Graw-Hill, New York, 1958.

[148] W. Feller. *An Introduction to Probability Theory and Its Applications.* Wiley, New York, 1957.

[149] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Volumes 1,2, and 3.* Addison-Wesley, Reading, MA, 1977.

[150] R. P. Feynman. "Simulating Physics with Computers." *Int. J. of Theoretical Physics* 21:467 - 488, 1982.

[151] R. P. Feynman. "Quantum Mechanical Computers." *Foundations of Physics* 16:507 - 531, 1986.

[152] R. P. Feynman. *QED - The Strange Theory of Light and Matter.* Princeton University Press, Princeton, NJ, 1985.

[153] R. P. Feynman. *Lectures on Computation.* Addison-Wesley, Reading, MA, 1996.

[154] D. L. Forney. *Concatenated Codes.* MIT Press, Cambridge, Ma, 1966.

[155] D. L. Forney. "Performance and Complexity" 1995 Shannon Lecture, Preprint, http://www.itsoc.org/review/06sha.pdf, December 2005.

[156] G. D. Forney, M. Grassl, and S. Guha. "Convolutional and Tail-Bitting Quantum Error-Correcting Codes." Preprint, http://arxiv.org/abs/quant-ph/0511016 v1, 2005.

[157] E. Fredkin. "Digital Machines: An Informational Process Based on Reversible Universal Cellular Automata." *Physica D*, 45:254 - 270, 1990.

[158] M. H. Freedman, M. Larsen, and Z. Wang. "A Modular Functor Which Is Universal for Quantum Computation," Preprint, http://arxiv.org/abs/quant-ph/0001108 v2, 2000.

[159] M. H. Freedman, A. Kitaev, and Z. Wang. "Simulation of Topological Field Theories by Quantum Computers," *Comm. Math. Phys.* 227:587 - 603, 2000.

[160] M. Freedman, C. Nayak, and K. Walker. "Towards Universal Topological Quantum Computation in the $\nu = \frac{5}{2}$ Fractional Quantum Hall State." *Phys. Rev. B,* 73(24):245307, 2006.

[161] H. Fritzsch. *An Equation that Changed the World.* University of Chicago Press, Chicago, IL, 2004.

[162] C. A. Fuchs. "Distinguishability and Accessible Information in Quantum Theory." Preprint, http://arxiv.org/abs/quant-ph/9601020 v1, 1996.

[163] C. A. Fuchs. "Nonorthogonal Quantum States Maximize Classical Channel Capacity." Preprint, http://arxiv.org/abs/quant-ph/9703043 v1, 1997.

[164] W. Fulton and J. Harris. *Representation Theory - A First Course.* Third Edition. Springer-Verlag, Heidelberg, 1997.

[165] B. Furrow. "A Panoply of Quantum Algorithms." Preprint, arxiv.org/quant-ph/0606127, 2006

[166] P. Galison. *Einstein's Clocks, Poincare's Maps.* W. W. Norton & Co. Publishing House, London, 2003.

[167] R. G. Gallagher. "A Simple Derivation of the Coding Theorem and Some Applications." *IEEE Trans. Information Theory*, IT-11:3 - 18, 1965.

[168] R. G. Gallagher. *Information Theory and Reliable Communication.* Wiley, New York, N.Y., 1968.

[169] S. Gasparoni, J.-W. Pan, P. Walter, T. Rudolph, and A. Zeilinger. "Realization of a Photonic Controlled-NOT Gate Sufficient for Quantum Computation." *Phys. Rev. Lett.,* 93(2): 020504, 2004.

[170] I. M. Gelfand. *Lectures on Linear Algebra.* Dover Publications, New York, NY, 1989.

[171] N. A. Gershenfeld and I. L. Chuang. "Bulk Spin-Resonance Quantum Computation." *Science*, 275(5298):350 - 356, 1997.

[172] A. M. Gleason. "Measures on a Close Subspace of a Hilbert Space." *J. of Mathematics and Mechanics*, 6:885-893, 1957.

[173] D. Gorenstein and N. Zierler. "A Class of Error-correcting Codes in $p^m$ Symbols." *J. of the Society for Industrial and Applied Mathematics (SIAM)*, 9:207-214, 1961.

[174] D. Gottesman. "Stabilizer Codes and Quantum Error Correction." *Ph.D. Thesis, California Institute of Technology*, Preprint, arxiv.org/quant-ph/9705052 v1, May 1997.

[175] D. Gottesman. "Theory of Fault-Tolerant Computation." *Phys. Rev. A*, 57: 127, 1998.

[176] D. Gottesman and I. L. Chuang. "Demonstrating the Viability of Universal Quantum Computation Using Teleportaion and Single-Qubit Operations." *Nature (London)*, 402:390 - 393, 1999.

[177] D. Gottesman. "An Introduction to Quantum Error Correction." *Proc. Symp. Applied Mathematics*, Preprint, arxiv.org/quant-ph/00040072 v1, 2000.

[178] R. M. Gray. *Entropy and Information Theory.* Springer Verlag, Heidelberg, 1990.

[179] M. Grassl, W. Gieselman, and T. Beths. "Quantum Reed-Solomon Codes." *Lecture Notes in Computer Science,* Springer Verlag, Heidelberg, Vol. 1719, pp. 231 - 244, 1999.

[180] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, Orlando, FL, 1980.

[181] P. Grangier, G. Roger and A. Aspect. "Experimental Evidence for a Photon Anticorrelation Effect on a Beam Splitter: A New Light on Single-Photon Interferences." *Europhysics Lett.,* 1(4):173 - 179, 1986.

[182] R. B. Griffiths and C.-S. Niu. "Semiclassical Fourier Transform for Quantum Computation." *Phys. Rev.*, 76:3228 - 3231, 1996.

[183] L. K. Grover. "A Fast Quantum Algorithm for Database Search." *Proc., ACM Symp. on Theory of Computing*, ACM Press, NY, 212 - 219, 1996. Also updated version: Preprint, arxiv.org/quanth-ph/9605043, 1996.

[184] L. K. Grover. "Quantum Mechanics Helps in Searching for a Needle in a Haystack." *Phys. Rev. Lett.*, 78:325 - 328, 1997.

[185] L. K. Grover. "A Framework for Fast Quantum Mechanical Algorithms." *Proc., Symp. on Theory of Computing*, ACM Press, NY, 53 - 62, 1998.

[186] L. K. Grover. "Quantum Computers Can Search Rapidely Using Almost Any Transformation." *Proc., Symp. on Theory of Computing*, ACM Press, NY, 53 - 62, 1998.

[187] Ł. K. Grover. "Searching with Quantum Computers." *Phys. Rev. Lett.,* 80(19):4329 - 4332, 1998.

[188] L. K. Grover. "From Schrödinger's Equation to the Quantum Search Algorithm." *American J. of Physics*, 69(7):769 - 777, 2001.

[189] L. K. Grover and A. M. Sengupta. "Classical Analog of Quantum Search." *Phys. Rev. A*, 65(3):032319, 2002.

[190] L. K. Grover. "Fixed-Point Quantum Search." *Phys. Rev. Lett.*, 95(15):150501, 2005.

[191] L. K. Grover. "A Different Kind of Quantum Search." Preprint, arxiv.org/quant-ph/0503205 v1, 2005.

[192] S. Guiasu. *Information Theory and Applications.* McGraw-Hill, New York, 1976.

[193] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U. D. Rapoli, M. Rieba, P. O.Schmidt, C. Becher, O. Güne, W. Dür, and R. Blatt. "Scalable Multiparticle Entanglement of Trapped Ions." *Nature*, 438:643 - 646, 2005.

[194] F. D. M. Haldane. "Fractional Quantization of the Hall Effect: A Heirarcy of Incompressible Quantum Fluid States." *Phys. Rev. Lett.,* 51(7):605 - 608, 1983.

[195] P. C. Haljan, P. J. Lee, K-A. Brickman, M. Acton, L. Deslauriers, and C. Monroe. "Entanglement of Trapped-Ion Clock States." *Phys. Rev. A*, 72:062316, 2005.

[196] E. H. Hall. "On the New Action of the Magnet on a Electric Currents." *American J. of Science,* 19(3):200 - 205, 1880. *Phil. Magazine,* 9(5):225 - 230, 1880.

[197] E. H. Hall. "On the New Action of Magnetism on a Permanent Electric Current." *American J. of Science,* 20(3):161 - 186, 1880. *Phil. Magazine,* 10(5):301 - 328, 1880.

[198] B. I. Halperin. "Quantized Hall Conductance, Current-Carrying Edge States, and the Existence of Extended States in a Two-Dimensional Disordered Potential." *Phys. Rev. B*, 25(4):2185 - 2190, 1982.

[199] B. I. Halperin. "Statistics of Quasiparticles and the Hierarchy of Fractional Quantized Hall States." *Phys. Rev. Lett.*, 52(18):1583 - 1586, 1984.

[200] R. W. Hamming. "Error Detecting and Error Correcting Codes." *Bell Systems Tech. Journal,* 29:147 - 160, 1950.

[201] L. Hardy. "Why is Nature Described by Quantum Theory" in J. D. Barrow, P. C. W. Davis, C. L. Hrper (Eds). *Science and Ultimate Reality. Quantum Theory, Cosmology and Complexity.* Cambridge University Press, 2004, also "Quantum Theory from Five Reasonable Axioms." Preprint, http://arxiv.org/abs/quant-ph/0101012 v4, 2001.

[202] Y. Hardy and W.-H. Steeb. *Classical and Quantum Computing.* Birkhäuser, Boston, MA, 2001.

[203] A. W. Harrow, B. Recht, I. L. Chuang. "Efficient Discrete Approximations of Quantum Gates." *J. Math. Phys.* 43:4445 - 4450, 2002.

[204] M. B. Hastings. "Superadditivity of Communication Capacity Using Entangled Inputs." *Nature Physics*, 5:255-257, 2009.

[205] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmorland, and W. K. Wooters. "Classical Information Capacity of a Quantum Channel." *Phys. Rev. A*, 54(1):1869 - 1876, 1996.

[206] M. Hayashi and H. Nagaoka. "General Formulas for Capacity of Classical-quantum Channels." *IEEE Trans. on Information Theory,* 49(7):1753-1768, 2003.

[207] W. Heisenberg. "The Development of Quantum Mechanics." *Nobel Lectures, Physics 1922- 1942*, 290 - 301, December 11, 1933. also http://nobelprize.org/nobel_prizes/physics/laureates/1932/heisenberg-lecture.pdf.

[208] K. -E. Hellwig and K. Kraus. "Pure Operations and Measurements." *Comm. Math. Phys.*, 11:214 - 220, 1969. See also "Operations and Measurements II." *Comm. Math. Phys.*, 16:142 - 147, 1970.

[209] A. Hocquenghem. "Codes Correcteurs d'Erreurs." *Chiffres (Paris)*, 2: 147-156, 1959.

[210] H. F. Hofmann and S. Takeuchi. "Quantum Phase Gate for Photonic Qubits Using only Beam Splitters and Postselection." *Phys. Rev. A,* 66(2): 024308, 2001.

[211] A. S. Holevo. "Statistical Problems in Quantum Physics." *Proc. Second Japan-URSS Symp. on Probability Theory, (G. Marayama and J. V. Prohorov, Eds.)*, Lecture Notes in Mathematics, Vol. 330, Springer Verlag, 104-119, 1973.

[212] A. S. Holevo. "The Capacity of Quantum Channel with General Signal States." *IEEE Trans. on Inform. Theory,* 44:269 - 273, 1998.

[213] J. P. Home, M. J. McDonnell, D. M. Lucas, G. Imreh, B. C. Keitch, D. J. Szwer, N. R. Thomas, S. C. Webster, D. N. Stacey, and A. M. Steane. "Deterministic Entanglement and Tomography of Ion-Spin Qubits." *New J. of Physics*, 8:188, 2006.

[214] C. K. Hong and L. Mandel. "Experimental Realization of a Localized One-Photon State." *Phys. Rev. Lett.,* 56(1):58 - 60, 1986.

[215] P. Hoyer. "Arbitrary Phases in Quantum Amplitude Purification." *Phys. Rev. A,* 62:052304, 2000.

[216] R. J. Hughes, G. I. Morgan, and C. G. Peterson. "Quantum Key Distribution over a 48 km Long Optical Fibre Network." *J. of Modern Optics,* 47:533 - 547, 2000.

[217] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small. "Quantum Information Processing Quantum Dot Spins and Cavity QED." *Phys. Rev. Lett.*, 83:42-4, 1999.

[218] J. K. Jain. "Composite-Fermion Approach for the Fractional Quantum Hall Effect." *Phys. Rev. Lett.,* 63(2):199 - 202, 1989.

[219] D. Jaksch, H.-J. Briegel, J. I. Cirac, and P. Zoller. "Entanglement of Atoms via Cold Controlled Collisions." *Phys. Rev. Lett.*, 82(9):1975 - 1978, 1999.

[220] Y. Ji, Y. Chung, D. Sprinzak, M. Heilblum, D. Mahalu, and H. Shtrikman. "An Electronic Mach-Zehnder Interferometer." *Nature,* 422:415 - 418, 2003.

[221] R. Jozsa. "Fidelity for Mixed Quantum States." *J. of Modern Optics*, 41(12):2315 - 2323, 1994.

[222] R. Jozsa and B. Schumacher. "A new Proof of the Quantum Noiseless Coding Theorem." *J. of Modern Optics*, 41(12):2343-2349, 1994.

[223] R. Jozsa. "Entanglement and Quantum Computation." *Geometric Issues in the Foundations of Science.* S. Hugget, L. Mason, K.P. Tod, S.T. Tsou, and N. M. J. Woodhouse, Eds. Oxford University Press, 1997. .

[224] R. Jozsa. *Quantum Algorithms and the Fourier Transform.* Preprint, arxiv.org/quant-ph/9707033 v1, 1997.

[225] R. Jozsa. "Searching in Grover's Algorithm." Preprint, arxiv.org/quant-ph/9901021 v1, 1999.

[226] R. Jozsa. "Quantum Factoring, Discrete Logarithms, and the Hidden Subgroup Problem." Preprint, arxiv.org/quant-ph/0012084 v1, 2000.

[227] R. Jozsa. "Illustrating the Concept of Quantum Information." Preprint arxiv.org/quant-ph/0305114 v1, 2003.

[228] J. Justensen and T. Hoholdt. *A Course in Error Correcting Codes.* European Mathematical Society, Zuerich, Switzerland, 2004.

[229] B. E. Kane. "Silicon-based Nuclear Spin Quantum Computer." *Nature,* 393:133 - 137, 1998.

[230] B. E. Kane. "Silicon-based Quantum Computation." Preprint arxiv.org/quant-ph/0003031 v1, 2000.

[231] M. Keyl. "Fundamentals of Quantum Information." Preprint arxiv.org/quant-ph/0202122 v1, 2002.

[232] A. Ya. Khinchin. *Mathematical Foundations of Information Theory.* Dover Publications, Mineola, NY, 1957.

[233] D. Kielpinski. *Entanglement and Decoherence in a Trapped-Ion Quantum Register.* Ph.D. Thesis, Univ. of Colorado, 2001.

[234] D. Kielpinski, A. Ben-Kish, J. Britton, V. Meyer, M. A. Rowe, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. "Recent Results in Trapped-Ion Quantum Computing." Preprint, arxiv.org/quant-ph/0102086 v1,, 2001.

[235] D. Kielpinski, V. Meyer, M. A. Rowe, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. "A Decoherence-Free Quantum Memory Using Trapped Ions." *Science,* 291(5506): 1013 - 1015, 2001.

[236] D. Kielpinski, C. Monroe, and D. J. Wineland. "Architecture for a Large-Scale Ion-Trap Quantum Computer." *Nature,* 417:709 - 711, 2002.

[237] J. M. Kikkawa and D. D. Awschalom. "Resonant Spin Amplification in n-Type GaAs". *Phys. Rev. Lett.*, 80(19):4313 - 4316, 1998.

[238] B. E. King, C. S. Wood, C. J. Myatt, Q. A. Turchette, D. Leibfried, W. M. Itano, C. Monroe, and D. J. Wineland. "Cooling the Collective Motion of Trapped Ions to Initialize a Quantum Register." *Phys. Rev. Lett.*, 81:1525 - 1528, 1998.

[239] A. Yu. Kitaev. "Quantum Measurements and the Abelian Stabilizer Problem." Preprint, arxiv.org/quant-ph/9511026 v1, 1995.

[240] A. Yu. Kitaev. "Quantum Computations and Error Correction." *Russian Mathematical Surveys,* 52:1191-1249, 1997.

[241] A. Yu. Kitaev. "Fault-Tolerant Quantum Computations by Anions." Annals Phys. 303:2-30, 2003.

[242] R. Klesse and S. Frank. "Quantum Error Correction in Spatially Correlated Quantum Noise." *Phys. Rev. Lett.*, 95:230503, 2005.

[243] K. von Klitzing, G. Dorda, and M. Pepper. "New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance." *Phys. Rev. Lett.,* 45(6):494 - 497, 1980.

[244] E. Knill, R. Laflamme, and W. H. Zurek. "Threshold Accuracy for Quantum Computation." Preprint, arxiv.org/quant-ph/9610011, 1996.

[245] E. Knill and R. Laflamme. "Concatenated Quantum Codes." Preprint, arxiv.org/quant-ph/9608012, 1996.

[246] E. Knill and R. Laflamme. "Theory of Quantum Error-Correcting Codes." *Phys. Rev. A,* 55(2):900 - 911, 1997.

[247] E. Knill, R. Laflamme, and W. H. Zurek. "Resilient Quantum Computation: Error Models and Thresholds." *Proc. Royal Society London A*, 454:365 - 384, 1998.

[248] E. Knill, R. Laflamme, and W. H. Zurek. "Resilient Quantum Computation." *Science*, 279:342, 1998.

[249] E. Knill, I. Chuang, and R. Laflamme. "Effective Pure States for Bulk Quantum Computation." *Phys. Rev. A,* 57(5):3348 - 3363, 1998.

[250] E. Knill, R. Laflamme, and L. Viola. "Theory of Quantum Error Correction for General Noise." *Phys. Rev. Lett.,* 84(11):2525 - 2528, 2000.

[251] E. Knill, R. Laflamme, and G. J. Milburn. "A Scheme for Efficient Quantum Computation with Linear Optics." *Nature*, 409:46 - 52, 2001.

[252] E. Knill, R. Laflamme, H. Barnum, D. Dalvit, J. Dziarmaga, J. Gubernatis, L. Gurvis, G. Ortiz, L. Viola, and W. H. Zurek. "Quantum Information Processing." *Kluwer Encyclopedia of Mathematics, Supplement III*, 2002.

[253] M. Koashi, T. Yamamoto, and N. Imoto. "Probabilistic Manipulation of Entangled Photons." *Phys. Rev. A,* 63(3):030301, 2001.

[254] A. N. Kolmogorov. "Logical Basis for Information Theory and Probability Theory." *IEEE Trans. on Information Theory,* IT-14: 662 - 664, 1968.

[255] A. N. Kolmogorov and S. V. Fomin. *Elements of the Theory of Functions and Functional Analysis.* Dover Publications, Mineola, NY, 1999.

[256] L. P. Kouwenhoven, D. G. Austing, and S. Tarucha. "Few-Electron Quantum Dots." *Reports on Progress in Physics*, 64(6):701 - 736, 2001.

[257] K. Kraus. "States, Effects, and Operators: Fundamental Notions of Quantum Theory." Springer Verlag, Berlin, 1983.

[258] D. Ktretschmann, D. Schlingemann, and R. F. Werner. "The Information–Disturbance Tradeoff and the Continuity of Stinespring's Representation." Preprint, arxiv.org/quant-ph/0605009, 2006.

[259] E. Kreyszig. *Advanced Engineering Mathematics.* Wiley, New York, NY, 1998.

[260] D. W. Kribs, R. Laflamme, and D. Poulin. "A Unified and Generalized Approach to Quantum Error Correction." *Phys. Rev. Lett.,* 94:180501, 2005.

[261] D. W. Kribs, R. Laflamme, D. Poulin, and M. Lesosky. "Operator Quantum Error Correction." Preprint, arxiv.org/quant-ph/0504189 v3, 2006.

[262] B. O. Küppers. *Information and the Origin of Life.* MIT Press, Cambridge, Mass., ISBN 0-262-11141-X, 1990.

[263] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. "Quantum Computers." *Nature,* 464:4552, 2010.

[264] R. Laflamme, C. Miquel, J.-P. Paz, and W. H. Zurek. "Perfect Quantum-Error Correcting Code." *Phys. Rev. Lett.*, 77: 198 - 201, 1996.

[265] R. Landauer. "Irreversibility and Heat Generation in the Computing Process." *IBM J. of Research and Development*, 5:182 - 192, 1961.

[266] R. B. Laughlin. "Quantized Hall Conductivity in Two Dimensions." *Phys. Rev. B,* 23(10):5632 - 5633, 1981.

[267] R. L. Laughlin. "Quantized Motion of Three Two-Dimensional Electrons in a Strong Magnetic Field." *Phys. Rev. B,* 27(6):3383 - 3389, 1983.

[268] R. L. Laughlin. "Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations." *Phys. Rev. Lett.,* 50(18): 1395 - 1398, 1983.

[269] J. M. Lee. "Introduction to Smooth Manifolds." *Springer Verlag,* ISBN 978-0387954486, 2002.

[270] D. Leibfried, B. DelMarco, V. Meyer, D. Lucas, M. Barrett, J. Britton, W. M. Itano, B. Jelenkoviĉ, C. Langer, T. Rosenband, and D. J. Wineland. "Experimental Demonstration of a Robust, High-Fidelity, Geometric Two Ion-Qubit Phase Gate." *Nature*, 422:412 - 415, 2003.

[271] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Post, C. Langer, and D. J. Wineland. "Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States." *Science,* 304:1476 - 1478, 2004.

[272] D. Leibfried, E. Knill, S. Seidlin, J. Britton, R. B. Blakestad, J. Chiaverini, D. B. Hume, W. M. Itano, J. D.Jost, C. Langer, R. Ozeri, R. Reichle, and D. J. Wineland. "Creation of a Six-Atom 'Schrödinger Cat' State." *Nature*, 438:639 - 642, 2005.

[273] C. Livermore, C. H. Crouch, R. M. Westervelt, K. L.Campman,and A. C. Grossard. "The Coulomb Blockade in Coupled Quantum Dots." *Science*, 274(5291):1332 - 1335, 1996.

[274] S. Lloyd. "A Potentially Realizable Quantum Computer." *Science*, 261:1569 - 1571, 1993.

[275] S. Lloyd "Almost any Quantum Logic Gate is Universal." *Phys. Rev. Lett.*, 75:346 - 349, 1995.

[276] S. Lloyd. "Capacity of a Noisy Communication Channel." *Phys. Rev. A*, 55:1613 - 1622, 1997.

[277] D. Loss and D. P. DiVincenzo "Quantum Computation with Quantum Dots." *Phys. Rev. A*, 57(1): 120 - 126, 1998.

[278] F. Lu and D. C. Marinescu "An $R \mid\mid C_{max}$ Quantum Scheduling Algorithm." *Quantum Information Processing* 6(3):159-178, 2007.

[279] F. Lu and D. C. Marinescu "Quantum Error Correction of Time-Correlated Errors." *Quantum Information Processing* 6(4):273 - 293, 2007.

[280] D. J. C. MacKay. *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, 2003.

[281] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error Correcting Codes.* North Holland, Amsterdam, 1983.

[282] O. Mandel, M. Greiner, A. Widera, T. Rom, T. W. Hänsch, and I. Bloch. "Controlled Collisions for Multi-Particle Entanglement of Optically Trapped Atoms." *Nature,* 425(6961):937 - 940, 2003.

[283] Y. Manin. "Classical Computing, Quantum Computing, and Shor's Algorithm." Talk at the Bourbaki Seminar, June 1999. Preprint, arxiv.org/quant-ph/9903008 v1, 1999.

[284] D. C. Marinescu and G. M. Marinescu *Approaching Quantum Computing* Prentice Hall, Upper Saddle River, New Jersey, 2004.

[285] D. C. Marinescu and G. M. Marinescu. "Quantum Information: a Glimpse at the Strange and Intriguing Future of Information." *The Computer Journal,* 50(5):505 - 521, 2007.

[286] S. McCartney. *ENIAC; The Triumphs and Tragedies of the World's First Computer.* Walker and Company Publishing House, New York, NY, 1999.

[287] B. Mc.Millan. "The Basic Theorems of Information Theory." *Ann. Math. Stat,* 24:196 - 212, 1953.

[288] Z. Meglicki, *Introduction to Quantum Computing.* http://beige.ucs.indiana.edu/M743, 2005.

[289] N. D. Mermin. " The Topologival Theory of Defects in Ordered Media." *Reviwes of Modern Physics,* 51(3):591 - 648, 1979

[290] N. D. Mermin. "What's Wrong with these Elements of Reality?" *Physics Today,* 43(6):9 - 11, 1990.

[291] N. D. Mermin. *Quantum Computer Science; An Introduction.* Cambridge University Press, Cambridge, England, 2007.

[292] E. Mertzbacher. *Quantum Mechanics.* Third Edition. Wiley, New York, NY, 1998.

[293] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. "Experimental Demonstration of Entanglement-Enhanced Rotation Angle Estimation Using Trapped Ions." *Phys. Rev. Lett.,* 86(26):5870 - 5873, 2001.

[294] G. J. Milburn. "Quantum Optical Fredkin Gate." *Phys. Rev. Lett.,* 62:2124 - 2127, 1988.

[295] A. Mizel, D. A. Lidar, and M. Mitchell. "Simple Proof of Equivalence between Adiabatic Quantum Computation and the Circuit Model." *Phys. Rev. Lett.,* 99(7):070502, 2007.

[296] C. Mochon. "Anyons from Non-solvable Finite Groups are Sufficient for Universal Quantum Computation." Preprint, arxiv.org/quant-ph/0206128 2003.

[297] K. Mølmer and A. Sørensen. "Multiparticle Entanglement of Hot Trapped Ions." *Phys. Rev. Lett.*, 82(9):1835 - 1838, 1999.

[298] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland. "Demonstration of a Fundamental Quantum Logic Gate." *Phys. Rev. Lett.*, 74(25):4714 - 4718, 1995.

[299] C. Monroe, D. Leibfried, B. E. King, D. M. Meekhof, W. M. Itano, and D. J. Wineland. "Simplified Quantum Logic with Trapped Ions." *Phys. Rev. A*, 55:2489 - 2491, 1997.

[300] G. Moore and N. Read. "Nonabelions in the Fractional Quantum Hall Effect." *Nuclear Physics B,* 360(2-3):362 - 369, 1991.

[301] R. H. Morf. "Transition from Quantum Hall to Compressible States in the Second Landau Level." *Phys. Rev. Lett.,* 80(7):1505 - 1508, 1998.

[302] M. Mosca and A. Ekert. "The Hidden Subgroup Problem and Eigenvalue Estimation on a Quantum Computer." Preprint, arxiv.org/quant-ph/9903071 v1, 1999.

[303] D. E. Muller. "Application of Boolean algebra to Switching Circuit Design and to Error Detection." *IEEE Trans. on Computers,* 3:612, 1954.

[304] C. Nayak and F. Wilczek. "2n-Quasihole States Realize $2^{n-1}$-Dimensional Spinor Braiding Statistics in Paired Quantum Hall States." *Nuclear Physics B,* 479: 529 - 553, 1996.

[305] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. "Non-Abelian Anyons and Topological Quantum Computation." *Reviews of Modern Physics,* 80:1083 - 1159, 2008.

[306] M. A. Nielsen and I. L. Chuang. "Programmable Quantum Gate Arrays." *Phys. Rev. Lett.*, 79:321, 1997.

[307] M. A. Nielsen. "Quantum Information Theory." *Ph.D. Thesis, University of New Mexico*, Preprint, arxiv.org/quant-ph/0011103, v6, 2000.

[308] M. A. Nielsen and I. L. Chuang. *Quantum Computing and Quantum Information.* Cambridge University Press, 2000.

[309] M. A. Nielsen. "Quantum Computation by Measurement and Quantum Memory." *Physics Letters A*, 308(2-3):96 - 100, 2003.

[310] M. A. Nielsen. "Optical Quantum Computation Using Cluster States." Preprint, arxiv.org/quant-ph/0402005, v1, 2004.

[311] E. Novais and H. U. Baranger, "Decoherence by Correlated Noise and Quantum Error Correction." *Phys. Rev. Lett.*, 97: 040501, 2006.

[312] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White. "Quantum Process Tomography of a Controlled-NOT Gate." *Phys. Rev. Lett.,* 93(8): 080502, 2004.

[313] J. L. O'Brien, G. J. Pryde, A. G. White, and T. C. Ralph. "High-Fidelity Z-Measurement Error Encodeing of Optical Qubits." *Phys. Rev. A,* 71(6): 060303, 2005.

[314] B. W. Ogburn and J. Preskill. "Topological Quantum Computation." *Lecture Notes in Computer Science*, 1509:341 - 359, Springer-Verlag, Heidelberg, 1999.

[315] H. Ollivier and J-P. Tillich. "Description of a Quantum Convolutional Code." Preprint, arxiv.org/quant-ph/0304189 v2, 2003.

[316] S. Olmschenk, D. N. Matsukevich, P. Maunz, D. Hayes, L. -M. Duan, C. Monroe. "Quantum Teleportation between Distant Matter Qubits." *Science,* 323(5913):486 - 489, 2009.

[317] R. Omnès. *The Interpretation of Quantum Mechanics.* Princeton University Press, Princeton, NJ, 1994.

[318] G. M. Palma, K.-A. Suominen, and A. Ekert. "Quantum Computers and Dissipation." Proc. Royal Soc. A, 452:567, 1996.

[319] J. W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger. "Experimental Test of Quantum Nonlocality in Three-Photon Greenberger-Horne-Zeilinger Entanglement." *Nature,* 403(6769):515 - 519, 2000.

[320] C. M. Papadimitriou. *Computational Complexity.* Addison-Wesley, Reading, MA, 1994.

[321] W. Pauli. "Exclusion Principle and Quantum Mechanics." *Nobel Lectures, Physics 1942 - 1962*, 27 - 43, December 13, 1946. Also: http://nobelprize.org/nobel_prizes/physics/laureates/1945/pauli-lecture.pdf.

[322] W. Peterson and E. J. Weldon. *Error Correcting Codes.* Second Edition. MIT Press, Cambride, MA, 1972.

[323] T. Pellizzari, S. A. Gardiner, J. I. Cirac, and P. Zoller. "Decoherence, Continuous Oservation, and Quantum Computing: A Cavity QED Model." *Phys. Rev. Lett.*, 75(21):3788 - 3791, 1995.

[324] R. Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe.* Vintage Books, ISBN 978-0-679-77631-4, 2007.

[325] A. Peres and W. K. Wootters. "Optimal Detection of Quantum Information." *Phys. Rev. Letters* 66(9):1119 - 1122, 1991.

[326] A. Peres. *Quantum Theory: Concepts and Methods.* Kluwer Academic Press, Dordrecht, Boston, London, 1995.

[327] M. R. Peterson, Th. Jolicoeur, and S. Das Sarma. "Finite Layer Thickness Stabilizes the Pfaffian State for the 5/2 Fractional Quantum Hall Effect: Wavefunction Overlap and Topological Degeneracy." Preprint: http://arxiv.org/PS_cache/arxiv/pdf/0803/0803.0737, 2008.

[328] C. A. Petri. "Nets, Time, and Space." *Theoretical Computer Science,* 153:3-48, 1996.

[329] P. M. Petroff, A. Lorke, and A. Imamoglu. "Epitaxially Self-Assembled Quantum Dots." *Physics Today*, 45(5):46 - 52, 2001.

[330] D. Petz. "Entropy, von Neumann and the von Neumann Entropy" Preprint, arxiv.org/math-ph/0102013 v1, 2001.

[331] T. B. Pittman, B. C. Jacobs, and J. D. Franson. "Probabilistic Quantum Logic Operations Using Polarizing Beam Splitters." *Phys. Rev. A,* 64(6): 062311, 2001.

[332] T. B. Pittman, M. J. Fitch, B. C. Jacobs, and J. D. Franson. "Experimental Controlled-NOT Logic Gate for Single Photons in the Coincidence Basis." *Phys. Rev. A,* 68(3): 032316, 2003.

[333] T. B. Pittman, B. C. Jacobs, and J. D. Franson. "Demonstration of Quantum Error Correction Using Linear Optics." *Phys. Rev. A,* 71(5): 052332, 2005.

[334] M. K. E. L. Planck. "The Genesis and Present State of Development of the Quantum Theory." *Nobel Lectures, Physics 1901 - 1922.* June 2, 1920. Also: http://nobelprize.org/nobel_prizes/physics/laureates/1918/planck-lecture.pdf.

[335] M. B. Plenio and V. Vitelli. "The Physics of Forgetting: Landauer's Erasure Principle and Information Theory." Preprint, arxiv.org/quant-ph/0103108 v1, 2001.

[336] S. Popescu. "Bell's Inequalities and Density Matrices. Revealing "Hidden" Nonlocality." Preprint, arxiv.org/quant-ph/9502005 v1, 1995.

[337] A. O. Pittinger. *An Introduction to Quantum Algorithms.* Birkhäuser, Boston, MA, 1999.

[338] D. Poulin. "Stabilizer Formalism for Operator Quantum Error Correction." *Phys. Rev. Lett.*, 95:230504, 2005.

[339] J. F. Poyatos, J. I. Cirac, and P. Zoller. "Quantum Gates with "Hot" Trapped Ions." *Phys. Rev. Lett.*, 81(6): 1322 - 1325, 1998.

[340] J. F. Poyatos, J. I. Cirac, and P. Zoller. "Complete Characterization of a Quantum Process: The Two-Bit Quantum Gate." *Phys. Rev. Lett.*, 78(2):390 - 393, 1997.

[341] J. F. Poyatos, J. I. Cirac, and P. Zoller. "Schemes of Quantum Computations with Trapped Ions." *Fortschritte der Physik*, 48(9-11):785 - 799, 2000.

[342] J. Preskill. "Fault Tolerant Quantum Computation." Preprint, arxiv.org/quant-ph/9712048 v1, 1997.

[343] J. Preskill. *Lecture Notes for Physics 229: Quantum Information and Computing.* California Institute of Technology, 1998.

[344] J. Preskill. "Quantum Clock Synchronization and Quantum Error Correction." Preprint, arxiv.org/quant-ph/0010098 v1, 2000.

[345] J. Preskill. "Topological Quantum Computation," Chapter 9 of Class Notes on Quantum Computing. http://www.theory.caltech.edu/ preskill/ph219/topological.pdf

[346] R. Prevedel, P. Walter, F. Tiefenbach, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger. "High-Speed Linear Optics quantum Computing Using Active Feed-Forward." *Nature,* 445:65 - 69, 2007.

[347] E. M. Purcell, H. C. Torrey, and R. V. Pound. "Resonance Absorption by Nuclear Magnetic Moments in a Solid." *Phys. Rev.*, 69:37, 1946.

[348] I. P. Radu, J. B. Miller, C. M. Marcus, M. A. Kastner, L. N. Pfeiffer, and K. W. West. "Quasiparticle Tunneling in the Fractional Quantum Hall State at $\nu = 5/$." *http://arxiv.org/PS_cache/arxiv/pdf/0803/0803.3530v1.pdf*, 2008.

[349] B. Rahn, A.C. Doherty, H. Mabuchi. "Exact Performance of Concatenated Quantum Codes." Preprint, arxiv.org/quant-ph/0206061 v1, 2002.

[350] R. Raussendorf and H. J. Briegel. "One-Way Quantum Computer." *Phys. Rev. Lett.*, 86(22):5188 - 5191, 2001.

[351] R. Raussendorf, D. E. Browne, and H. J. Briegel. "Measurement-based Quantum Computation on Cluster States." *Phys. Rev. A,* 68(2):022312, 2003.

[352] N. Read and E. Rezayi. "Quasiholes and Fermionic Zero Modes of Paired Fractional Quantum Hall States: the Mechanism for Non-Abelian Statistics." *Phys. Rev. B,* 54(23):16864, 1999.

[353] I. S. Reed. "A Class of Multiple-error-correcting Codes and the Decoding Scheme." *IEEE Trans. on Information Theory,* 4:3849, 1954.

[354] I. S. Reed and G. Solomon. "Polynomial Codes over Certain Finite Fields" *SIAM J. of Applied Mathematics*, 8:300 - 304, 1960.

[355] M. Reed and B. Simon. "Methods of Modern Mathematical Physics. Vol. 1: Functional Analysis." Academic Press, New York, 1972.

[356] B. W. Reichardt and L. K. Grover. "Quantum Error Correction of Systematic Errors Using a Quantum Search Framework." *Phys. Rev. A*, 72(4):042326, 2005.

[357] E. H. Rezayi and F. D. M. Haldane. "Incompressible Paired Hall State, Stripe Order, and the Composite Fermion Liquid Phase in Half-filled Landau Levels." *Phys. Rev. Lett.,* 84(20):4685 - 4688, 2000.

[358] M. Riebe, H. Häffner, C. F. Roos, W. Hänsel, J. Benhelm, G. P. T. Lancaster, T. K. Körber, C. Becher, F. Schmidt-Kaler, D. F. V. James, and R. Blatt. "Deterministic Quantum Teleportation with Atoms." *Nature*, 429:734 - 737, 2004.

[359] M. Riebe, K. Kim, P. Schindler, T. Monz, P. O. Schmidt, T. K. Kärber, W. Hänsel, H. H"affner, C. F. Roos, and R. Blatt. "Process Tomography of Ion Trap Quantum Gates." *Phys. Rev. Lett.*, 97:220407, 2006.

[360] E. Rieffel and W. Polak. "An Introduction to Quantum Computing for Non-Physicists." *ACM Computing Surveys*, 32(3):300 - 335, 2000.

[361] C. Roos, T. Zeigler, H. Rohde, H. C. Nägerl, J. Eschner, D. Leibfried, F. Schmidt-Kaler, and R. Blatt. "Quantum State Engineering on an Optical Transition and Decoherence in a Paul Trap." *Phys. Rev. Lett.,* 83(23):4713 - 4716, 1999.

[362] P. Roulleau, F. Portier, P. Roche, A. Cavanna, G. Faini, U. Gennser, and D. Mailly. "Direct Measurement of the Coherence Length of Edge States in the Integer Quantum Hall Regime." *Phys. Rev. Lett.,* 100:126802, 2008.

[363] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. "Experimental Violation of a Bell's Inequality with Efficient Detection." *Nature,* 409(6822):791 - 794, 2001.

[364] C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, D. J. Wineland, and C. Monroe. "Experimental Entanglement of Four Particles." *Nature*, 404:256 - 259, 2000.

[365] F. Schmidt-Kaler, C. Roos, H. C. Nägerl, H. Rohde, S. Gulde, A. Mundt, M. Lederbauer, G. Thalhammer, T. Zeiger, P. Barton, L. Hornekaer, G. Reymond, D. Leibfried, J. Eschner, and R. Blatt. "Ground State Cooling, Quantum State Engineering and Study of Decoherence of Ions in Paul Traps." *J. of Modern Optics,* 47(14/15):2573 - 2582, 2000.

[366] F. Schmidt-Kaler, H. Häffner, M. Riebe, S. Guide, G. P. T. Lancaster, T. Deuschle, C. Becher, C. F. Roos, J. Eschner, and R. Blatt. "Realization of the Cirac-Zoller Controlled-NOT Quantum Gate." *Nature*, 422:408 - 411, 2003.

[367] E. Schrödinger. "The Fundamental Idea of Wave Mechanics." *Nobel Lectures, Physics 1922- 1941*, 305 - 314. December 12, 1933. Also: http://nobelprize.org/nobel_prizes/physics/laureates/1933/schrodinger-lecture.pdf.

[368] E. Schrödinger. "The Present Situation in Quantum Mechanics." *Die Naturwissenschaften*, 23: 807 - 812; 823- 828; 944-849, 1935. Also *Proc. of the Cambridge Philosophical Society*, 31:555 - 563, 1935 and 32:446 - 452, 1936.

[369] B. W. Schumacher. "Quantum Coding." *Phys. Rev. A*, 51(4):2738 - 2747, 1995.

[370] B. W. Schumacher, M. D. Westmorland and W. K. Wootters. "Limitations on the Amount of Accessible Information in a Quantum Channel." *Phys. Rev. Lett.,* 76:3452 - 3455, 1996.

[371] B. W. Schumacher. "Entropy Exchange and Coherent Quantum Information.", *Proc. Fourth Workshop on Physics and Computation,* (T. Toffoli, M. Biafore and J. Leao, Eds), New England Complex Systems Institute, Cambridge, MA, 292 - 296 1996.

[372] B. W. Schumacher and M. A. Nielsen. "Quantum Data Processing and Error Correction." *Phys. Rev. A,* 54(4): 2629 - 2636, 1996.

[373] B. W. Schumacher. "Sending Quantum Entanglement through Noisy Quantum Channels." *Phys. Rev. A,* 54(4):2614 - 2628, 1996.

[374] B. W. Schumacher and M. D. Westmorland. "Sending Quantum Information via Noisy Quantum Channels." *Phys. Rev. A,* 56(1):131 - 138, 1997.

[375] C. E. Shannon. "A Mathematical Theory of Communication." *Bell Sys. Tech. Journal*, 27:379 - 423 & 623 - 656, 1948.

[376] C. E. Shannon. "Communication in the Presence of Noise." *Proc. of the IRE*, 37:10 - 21, 1949.

[377] C. E. Shannon. "Certain Results in Coding Theory for Noisy Channels." *Information and Control*, 1(1):6 - 25, 1957.

[378] C. E. Shannon and W. Weaver. *A Mathematical Theory of Communication.* University of Illinois Press, Urbana, IL, 1963.

[379] M. S. Sherwin, A. Imamoglu, and T. Montroy. "Quantum Computation with Quantum Dots and Terahertz Cavity Quantum Electrdynamics." *Phys. Rev. A*, 60(5):3508 - 3514, 1999.

[380] P. W. Shor. "Algorithms for Quantum Computation: Discrete Log and Factoring." *Proc., 35 Annual Symp. on Foundations of Computer Science*, 124 - 134, IEEE Press, Piscataway, NJ, 1994.

[381] P. W. Shor. "Scheme for Reducing Decoherence in Quantum Computer Memory." *Phys. Rev. A*, 52(4):2493 - 2496, 1995.

[382] P. W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer." Preprint, arxiv.org/quant-ph/9508027 v2, 1996.

[383] P. W. Shor. "Fault-Tolerant Quantum Computation." *37th Annual Symp. on Foundations of Computer Science*, IEEE Press, Piscataway, NJ, 56 - 65, 1996.

[384] P. W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer." *SIAM J. of Computing,* 26:1484 - 1509, 1997.

[385] P. W. Shor and J. Preskill. "Simple Proof of Security of the BB84 Quantum Key Distribution Protocol." Preprint, arxiv.org/quant-ph/0003004, 2000.

[386] P. W. Shor. "Introduction to Quantum Algorithms." Preprint, arxiv.org/quant-ph/0005003 2001.

[387] P. W. Shor. "Why Haven't More Quantum Algorithms Been Found?" *J. of the ACM*, 50(1):87 - 90, 2003.

[388] P. W. Shor. "Capacities of Quantum Channels and How to Find Them" Preprint, arxiv.org/quant-ph/0304102 v1, 2003.

[389] P. W. Shor. "Quantum Channels." http://www-math.mit.edu/ shor/NewDirections.

[390] P. W. Shor. "Equivalence of Additive Questions in Quantum Information Theory." *Comm. Math. Phys.*, 246:453 - 472, 2004.

[391] P. W. Shor. "The Bits Don't Add Up." *Nature Physics*, 5:247 - 248, 2009.

[392] D. R. Simon. "On the Power of Quantum Computation." *SIAM Journal on Computing,* 26:1474 - 1483, 1997.

[393] L. Sklar. *Philosophy of Physics.* Westview Press, Boulder, CO, 1992.

[394] B. Sklar. "A Primer on Turbo Code Concepts." *IEEE Comm. Magazine*, 35(4):94 - 102, 1997.

[395] B. Sklar. *Digital Communication: Fundamentals and Applications,* Second Edition, Prentice Hall, Upper Saddle River, N.J., 2001.

[396] T. Sleator and H. Weinfurter. "Realizable Universal Quantum Logic Gates." *Phys. Rev. Lett.*, 74:4087 - 4090, 1995.

[397] D. Slepian and J. K. Wolf. " Noiseless Coding of Correlated Information Sources." *IEEE Trans. on Information Theory,* IT-19:471 - 480, 1973.

[398] G. Smith, J. Smolin and A. Winter. "The Quantum Capacity with Symmetric Side-channels." *IEEE Trans. Information Theory,* IT-54(9):4208-4217, 2008.

[399] G. Smith and J. Yard. "Quantum Communication with Zero-Capacity Channels." *Science*, 321:1812-1815, 2008.

[400] A. Sørensen and K. Mølmer. "Quantum Computation with Ions in Thermal Motion." *Phys. Rev. Lett.*, 82(9):1971 - 1974, 1999.

[401] A. Sørensen and K. Mølmer. "Entanglement and Quantum Computation with Ions in Thermal Motion." *Phys. Rev. A*, 62:022311, 2000.

[402] A. M. Steane. "Multiple Particle Interference and Quantum Error Correction." *Proc. Roy. Soc. London A* 452:2551, 1996.

[403] A. M. Steane. "The Ion Trap Quantum Information Processor." Preprint, arxiv.org/quant-ph/9608011 v2, 1996.

[404] A. M. Steane. "Quantum Computing." *Reports on Progress in Physics*, 61: 117, 1998.

[405] A. M. Steane, "Error Correcting Codes in Quantum Theory." *Phys. Rev. Lett.,* 77:793 - 797, 1996.

[406] A. M. Steane. "Active Stabilization, Quantum Computation, and Quantum State Synthesis." *Phys. Rev. Lett.*, 78:2252, 1997.

[407] A. Steane. "The Ion Trap Quantum Information Processor." *Applied Physics B,* 64: 623 - 642, 1997.

[408] A. M. Steane. "Efficient Fault-Tolerant Quantum Computing." *Nature (London)*, 399:124, 1999.

[409] A. M. Steane and D. M. Lucas. "Quantum Computing with Trapped Ions, Atoms and Light." *Fortschritte der Physik,* 48:839 - 858, 2000.

[410] A. M. Steane. "Overhead and Noise Threshold of Fault-Tolerant Quantum Error Correction." *Phys. Rev. A*, 68:042322, 2003.

[411] O. Stern. "The Method of Molecular Rays." *Nobel Lectures, Physics 1942 - 1961*, 8 - 16. Also: http://nobelprize.org/nobel_prizes/physics/laureates/1943/stern-lecture.pdf.

[412] W. F. Stinespring. "Positive Functions on $C^*$-Algebras." *Proc. of the American Mathematical Society,* 211-216, 1955.

[413] M. Sudan. "Lecture Notes for Essential Coding Theory", http://theory.lcs.mit.edu/madhu, 2005.

[414] N. S. Szabo and R. I. Tanaka. *Residue Arithmetic and Its Applications to Computer Technology.* McGraw-Hill, New York, NY, 1967.

[415] L. Szilárd. "Über die Entropieverminderung in einem Thermodynamichen System bei Eingriffen Intelligenter Wesen." *Zeitschrifft für Physik,* 53:840 - 856, 1929.

[416] M. S. Tame, R. Prevedel, M. Paternostro, P. Böhi, M. S. Kim, and A. Zeilinger. "Experimental Realization of Deutsch's Algorithm in a One-Way Wuantum Computer." Preprint, arxiv.org/quant-ph/0611186, v2, 2007.

[417] T. Tanamoto. "Quantum Gates by Coupled Quantum Dots and Measurement Procedure in Field-Effect-Transistor Structure." *Fortschritte der Physik*, 48(9-11):1005 - 1021, 2000.

[418] S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, and L. P. Kouwenhoven. "Shell Filling and Spin Effects in a Few Electron Quantum Dots." *Phys. Rev. Lett.,* 77(17):3613 - 3616, 1996.

[419] B. M. Terhal. "Is Entanglement Monogamous?" *IBM J. of Research and Development*, 48(1):71 - 78, 2004.

[420] B. M. Terhal, A. C. Doherty, and D. Schwab. "Symmetric Extensions of Quantum States and Local Hidden Variable Theories." *Phys. Rev. Lett.,* 90(15):157903, 2003.

[421] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé, and K. Chan. "A Silicon Nanocrystals Based Memory." *Applied Physics Letters,* 68(10):1377 - 1379, 1996.

[422] T. Toffoli. "Reversible Computing." In *Automata, Languages, and Programming*, Eds. J. W. de Bakker and J. van Leeuwen, pp. 632 - 644, Spinger Verlag, New Yoprk, NY, 1980.

[423] A. C. de la Tore, A. Daleo, and I. Garcia-Mata. *The Photon-Box Bohr-Einstein Debate Demythologized.* Preprint, http://arxiv.org/abs/quant-ph/9910040 v1 1999.

[424] L. A. Tracy, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. "Spin Transition in the Half-Filled Landau Level." *Phys. Rev. Lett.,* 98(8):086801, 2007.

[425] D. C. Tsui, H. L. Störmer, and A. C. Gossard. "Two-Dimensional Magnetotransport in the Extreme Quantum Limit." *Phys. Rev. Lett.,* 48(22):1559 - 1562, 1982.

[426] T. Tulsi, L. K. Grover, and A. D. Patel. "A New Algorithm for Directed Quantum Search." Preprint, arxiv.org/quanth-ph/0505007 v2, 2005.

[427] Q. A. Turchette, C. S. Wood, B. E. King, C. J. Myatt, D. Leibfried, W. M. Itano, C. Monroe, and D. J. Wineland. "Deterministic Entanglement of Two Ions." *Phys. Rev. Lett.,* 81:3631 - 3634, 1998.

[428] Q. A. Turchette, D. Kielpinski, B. E. King, D. Leibfried, D. M. Meekhof, C. J. Myatt, M. A. Rowe, C. A. Sackett, C. S. Wood, W. M. Itano, C. Monroe, and D. J. Wineland. "Heating of Trapped Ions from the Quantum Ground State." *Phys. Rev. A*, 61(6):063418, 2000.

[429] A. M. Turing. "On Computable Numbers with an Application to the Entscheidungsproblem." *Proc. London Mathematical Society 2*, 42:230, 1936.

[430] A. Uhlmann. "The Transition Probability in the State Space of $C^*$-Algebra." *Rep. Math. Phys.*, 9:273 - 279, 1976.

[431] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, R. Cleve, and I. L. Chuang. "Experimental Realization of an Order-Finding Algorithm with an NMR Quantum Computer." *Phys. Rev. Lett.,* 85(25):5452 - 5455, 2000.

[432] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. S. Chuang. "Experimental Realization of Shor's Quantum Factoring Algorithm Using Nuclear Magnetic Resonance." *Nature*, 414:883 - 887, 2001.

[433] L. Vandersypen and I. Chuang. "NMR Computing - Lessons for the Furture." *Quantum Information and Computation,* 1:134 - 142, 2001.

[434] L. M. K. Vandersypen and I. S. Chuang. "NMR Techniques for Quantum Control and Computation." *Reviews of Modern Physics*, 76: 1037 - 1069, 2004.

[435] S. J. Van Enk, J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi. "Quantum State Transfer in a Quantum Network: a Quantum-Optical Implementation." *J. of Modern Optics*, 44(10):1727 - 1736, 1997.

[436] S. A. Vanstone and P. C. van Oorschot. *An Introduction to Error Correcting Codes with Applications.* Kluwer, Boston, MA, 1989.

[437] V. Vedral. "The Role of Entropy in Quantum Information Theory." Preprint, arxiv.org/quant-ph/0102094 v1, 2001.

[438] S. Verdù and T. S. Han. "A General Formula for Classical Channel Capacity." *IEEE Trans. on Information Theory,* 40:1147 - 1157, 1994.

[439] G. Vidal. "Efficient Classical Simulation of Slightly Entangled Quantum Computations." *Phys. Rev. Lett.* 91:14792 - 14976, 2003.

[440] A. J. Viterbi. "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm." *IEEE Trans. on Information Theory,* IT-13:260 - 269, 1967.

[441] J. von Neumann. "Thermodynamik quantummechanischer Gesamtheiten (The Thermodynamics of Quantum Mechanical Ensembles)." *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Matematisch-Physikalische Klasse,* 1:273 - 291, 1927. Also, http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi?PPN252457811

[442] J. von Neumann. *Mathematical Foundations of Quantum Mechanics.* Trans. R. T. Bayer. Princeton University Press, Princeton, NJ, 1955.

[443] J. von Neumann. "Probabilistic Logic and Synthesis of Reliable Organisms from Unreliable Components." In *Automata Studies*, C. E. Shannon and J. McCarthy, Editors. Princeton University Press, Princeton, NJ, 1956.

[444] J. von Neumann. "Fourth University of Illinois Lecture." *Theory of Self-Reproduced Automata*, A. W. Burks, Editor. 66, University of Illinois Press, Urbana, IL, 1966.

[445] J. Watrous. "Lecture Notes: Theory of Quantum Information." University of Waterloo, http://www.cs.uwaterloo.ca/ watrous/lecture-notes.html, 2007.

[446] A. Wehrl "General Properties of Entropy." *Reviews of Modern Physics*, 50(2):221 - 260, 1978.

[447] C. F. von Weizsäcker. *Die Einheit der Natur* (The Unity of Nature). Ed. F. Zucker. Farrar, Straus, and Giroux, 1971

[448] C. F. von Weizsäcker and E. von Weizsäcker. "Wideraufname der begrifflichen Frage: Was ist Information?" (Revisting the Fundamental Question: What is Information?). *Nova Acta Leopoldina*, 206, 535, 1972.

[449] X. G. Wen and Q. Niu. "Ground-State Degeneracy of the Fractional Quantum Hall States in the Presence of a Random Potential and on High-Genus Riemann Surfaces." *Phys. Rev. B,* 41(13):9377 - 9396, 1990.

[450] X. G. Wen. "Chiral Luttinger Liquid and the Edge Excitations in the Fractional Quantum Hall Effect." *Phys. Rev. B,* 41(14):12838 - 12844, 1990.

[451] X. G. Wen. "Edge Transport Properties of the FQH States and Impurity Scattering of a One-Dimensional Charge-Density Wave." *Phys. Rev. B,* 44:5708 - 5719, 1991.

[452] X. G. Wen. "Non-Abelian Statistics in the Fractional Quantum Hall States." *Phys. Rev. Lett.,* 66(6):802 - 805, 1991.

[453] X. G. Wen. "Topological Orders and Edge Excitations in FQH States." *Advances in Physics* 44:405 - 473, 1995.

[454] R. F. Werner. "Quantum States with Einstein-Podolsky-Rosen Correlations Admitting a Hidden-Variable Model." *Phys. Rev. A*, 40:4277 - 4281, 1989.

[455] E. P. Wigner. *Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren,* Vieweg Verlag, Braunschweig (1931). Translated into English: J. J. Griffin, Group Theory and its Application to the Quantum Mechanics of Atomic Spectra Academic Press, New York (1959).

[456] F. Wilczek. "Quantum Mechanics of Fractional-Spin Particles." *Phys. Rev. Lett.*, 49(14):957 - 959, 1982.

[457] R. Willett, J. P. Eisenstein, H. L. Störmer, D. C. Tsui, A. C. Gossard, and J. H. English. "Observation of an Even-Denominator Quantum Number in the Fractional Quantum Hall Effect." *Phys. Rev. Lett.,* 59(15):1776-1779, 1987.

[458] R. I. Willett, M. J. Manfra, I. N. Pfeiffer, and K. W. West. "Interferometric Measurement of Filling Factor 5/2 Quasiparticle Charge." *http://arxiv.org/ftp/arxiv/papers/0807/0807.0221.pdf*, 2008.

[459] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof. "Experimental Issues in Coherent Quantum State Manipulation of Trapped Atomic Ions." *J. of Research NIST,* 103:259 - 328, 1998.

[460] D. J. Wineland, M. Barrett, J. Britton, J. Chiaverini, B. DeMarco, W. M. Itano, B. Jelencovic, C. Langer, D. Leibfried, V. Meyer, T. Rosenband, and T. Schätz. "Quantum Information Processing with Trapped Ions." Preprint, arxiv.org/quant-ph/0212079 v2, 2003.

[461] W. K. Wootters and W. H. Zurek. "A Single Quantum Cannot Be Cloned." *Nature*, 299:802 - 803, 1982.

[462] W. K. Wootters. "Entanglement of Formation of an Arbitrary State of Two Qubits." *Phys. Rev. Lett.,* 80(10): 2245 - 2248, 1998.

[463] J. M. Wozencraft and B. Reiffen. *Sequential Decoding.* MIT Press, Cambridge, MA, 1961.

[464] J. S. Xia, W. Pan, C. L. Vincete, E. D. Adams, N. S. Sullivan, H. L. Stormer, D. C. Tsui, L. N. Pfeiffer, K. W. Baldwin, and K. W. West. "Electron Correlation in the Second Landau Level: A Competition Between Many Nearly Degenerate Quantum Phases." *Phys. Rev. Lett.,* 93(17):176809, 2004.

[465] A. Yao. "Quantum Circuit Complexity." *Proc. Symp. on Foundations of Computer Science*, 352 - 361, 1993.

[466] C. Zalka. "Grover's Quantum Searching Algorithm is Optimal." *Phys. Rev. A* 60:2746 - 2751, 1999.

[467] P. Zanardi and M. Rasatti. "Noiseless Quantum Codes." *Phys. Rev. Lett.,* 79:3306-3308, 1998.

[468] P. Zanardi and F. Rossi. "Quantum Information in Semiconductors: Niseless Encoding in a Quantum-dot Array." *Phys. Rev. Lett.,* 81:4752-4755, 1998.

[469] P. Zanardi, D. A Lidar, and S. Lloyd. "Quantum Tensor Product Structures are Observable Induced." *Phys. Rev. Lett.,* 92:060402, 2004.

[470] O. Zilberberg, B. Braunecker, and D. Loss. "Controlled-NOT for Multiparticle Qubits and Topological Quantum Computation Based on Parity Measurements." Preprint, http://arxiv.org/PS_cache/arxiv/pdf/0708/0708.1062v2.pdf, 2007.

[471] W. H. Zurek. "Decoherence and the Transition from Quantum to Classical." *Physics Today*, 44(10):36 - 44, 1991.

[472] W. H. Zurek. "Decoherence, Einselection, and the Quantum Origins of the Classical." Preprint, arxiv.org/quant-ph/0105127v3, 2003.

[473] W. H. Zurek. "Probabilities from Envariance." Preprint, arxiv.org/quant-ph/0405161, 2004.

[474] "Quantum Information Science and Technology Roadmap." *ARDA - Quantum Information Science and Technology Project*, Version 2.0, 2004. Preprint, http://qist.lanl.gov/qcomp_map.shtml

[475] Conference on Lasers and Electron Optics. http://optics.org/articles/news/9/6/3/1.

[476] *Semiconductor Industry Association Roadmap 2000-2001.* http://public.itrs.net.

[477] *Stanford Encyclopedia of Philosophy.* http://plato.stanford.edu/entries/qm-copengahen.

# 8 Glossary

**adiabatic** Term frequently used in thermodynamics meaning without transfer of heat (without change in temperature). For example, if we open the valve of a gas canister, the gas rushes out and expands without having the time to equalize its temperature with the environment. The rushing gas feels cool.

**Aharonov - Bohm effect** Given a magnetic field $\mathbf{B}$ consisting of a single field line with flux $\Phi_B$; at positions $R$ not on the flux line, the field is $\mathbf{B} = 0$, but there must be a vector potential $\mathcal{B}(R)$ satisfying the expression

$$\oint_C \mathcal{B}(R) \cdot dR = \Phi,$$

for closed contours $C$ threaded by the flux line. Aharonov and Bohm had shown [4] that in quantum mechanics such vector potentials, though corresponding to zero-value field, have physical significance. Berry proved [53] that the effect of such vector potentials on a system of particles with charge $q$ clustered inside a box at a zero-field position $R$ and moved around the magnetic field line can be interpreted as a geometrical phase factor *Berry phase* acquired by the system's single-valued state vector after completion of the closed circuit $C$. The geometrical phase factor for any state $\mid n(R)\rangle$ of energy $E_n$ independent of position $R$ and unaffected by the vector potential is given as

$$\gamma_n(C) = \frac{q}{\hbar} \oint_C \mathcal{B}(R) \cdot dR = \frac{q}{\hbar}\Phi.$$

Notice that the geometrical phase factor is independent of the state $n$ and even of the contour $C$ if the system makes a complete rotation about the flux line. The geometrical phase factor can be observed by interference between the particles inside the rotated box and particles in a box which was not moved round the closed contour.

**amplitude, probability amplitude** Given a set of basis or logical states $\mid i\rangle$, a quantum system may be in a superposition of these states, $\mid \psi\rangle = \sum_i \alpha_i \mid i\rangle$. The complex numbers $\alpha_i$ are called amplitudes or probability amplitudes.

**ancillas** Auxiliary qubits used to assist in a computation.

**anti-commutator of two operators** Given two operators $\mathbf{A}$ and $\mathbf{B}$, the *anti-commutator* of the two operators is $\{\mathbf{A}, \mathbf{B}\} = \mathbf{AB} + \mathbf{BA}$. We say $\mathbf{A}$ anti-commutes with $\mathbf{B}$ if $\{\mathbf{A}, \mathbf{B}\} = 0$.

**anyons** Indistinguishable particles in two dimensions that are neither bosons nor fermions. Consider a gas of electrons squeezed between two slabs of semiconductor materials such that, the movement of electrons is restricted to two dimensions only. At very low temperature and in a strong magnetic field the two-dimensional electron gas has a strongly entangled ground (lowest energy) state separated from all other states by an energy gap. This lowest-energy state carry an electric charge that is not an integer multiple of the electron charge and does not have the quantum numbers associated with electrons [345]. The properties of the anyons manifest themselves as the Fractional Quantum Hall (FQH) effect. Anyons are suggested as candidates to explain high temperature superconductors.

**atomic radiation cascade (ARC)** A system which produces two polarization entangled photons which move in opposite directions. An atom, e.g., $Ca^{40}$, laser pumped to an excited

$S$, the upper level of the cascade state, emits two photons that appear to have the same polarization.

**basis states of a quantum system** Linearly independent, complete subset of orthonormal state vectors in the vector space of the quantum system. Two orthonormal states such as $\mid 0\rangle$ and $\mid 1\rangle$ could be the basis states of a qubit.

**Bayes rule** Probability rule. Assume that it is known that event $\mathcal{A}$ occurred, but it is not known which one of the set of mutually exclusive and collectively exhaustive events $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_n$ has subsequently occurred. Then the conditional probability that one of these events, $\mathcal{B}_j$, occurs, given that $\mathcal{A}$ occurs is

$$P(\mathcal{B}_j \mid \mathcal{A}) = \frac{P(\mathcal{A} \mid \mathcal{B}_j)P(\mathcal{B}_j)}{\sum_i P(\mathcal{A} \mid \mathcal{B}_i)P(\mathcal{B}_i)}$$

with $P(\mathcal{B}_j \mid \mathcal{A})$ as the *a posteriori probability.*

**BB84** A quantum key distribution protocol proposed by Charles Bennett and Gilles Brassard in 1984.

**BCH (Bose-Chaudhuri-Hocquenghem) bound** Let $g(x)$ be the generator polynomial of the cyclic $[n, k]$ code $\mathcal{C}$ over $GF(q)$. Let $\beta \in GF(q)$ be a primitive element of order $n$. If $g(x)$ has among its zeros $\beta^{\alpha}, \beta^{\alpha+1}, \ldots, \beta^{\alpha+d-2}$ then the minimum distance of the cyclic code is $d_{min}(\mathcal{C}) \geq d$.

**BCH codes** Cyclic codes independently discovered by A. Hocquenghem [209] and R.C. Bose and R.K. Ray-Chaudhuri [68]. Given the prime power $q$, and integers $m$ and $d$, the $BCH_{q;m;d}$ is obtained as follows: let $n = q^m$ and $GF(q^m)$ be an extension of $GF(q)$ and let $C'$ be the (extended) $[n; n-(d-1); d]_{q^m}$ Reed-Solomon code obtained by evaluating polynomials of degree at most $(n-d)$ over $GF(q^m)$ at all the points of $GF(q^m)$. Then the code $BCH_{q;m;d}$ is the $GF(q)$-subfield subcode of $C'$. A BCH code has dimension at least $n - m(d-1)$ [413].

**Bell, John Stewart (1928 - 1990).** British physicist whose work led to the possibility of exploring seemingly philosophical questions in quantum mechanics, such as the nature of reality, directly through experiments. Bell started from locality and argued for the existence of deterministic hidden variables. He was able to show that the behavior predicted by quantum theory could not be duplicated by a hidden variable theory if the hidden variables acted locally. Bell's Inequality is a constraint on the sum of averages of measured observables, based on assumptions of *local realism* (*i.e.*, of locality and of existence of hidden variables).

**Bell states** Four distinct quantum states of a two-particle system with a very strong coupling of the individual states of the particles. These states form an orthonormal basis:

$$\mid \beta_{00}\rangle = \frac{\mid 00\rangle + \mid 11\rangle}{\sqrt{2}}, \quad \mid \beta_{01}\rangle = \frac{\mid 01\rangle + \mid 10\rangle}{\sqrt{2}},$$

$$\mid \beta_{10}\rangle = \frac{\mid 00\rangle - \mid 11\rangle}{\sqrt{2}}, \quad \mid \beta_{11}\rangle = \frac{\mid 01\rangle - \mid 10\rangle}{\sqrt{2}}.$$

**Bennett, Charles** Distinguished American scientist working at IBM Research who has made outstanding contributions to quantum information theory. He built on the work of Rolf Landauer to show that general-purpose computation can be performed by a logically

and thermodynamically reversible apparatus; he proposed a re-interpretation of Maxwell's demon, attributing its inability to break the second law to the thermodynamic cost of destroying, rather than acquiring, information; together with several collaborators he discovered quantum teleportation.

**Berry phase** See the entry for the *Aharonov - Bohm effect*.

**black box** A computational device whose inner structure is unknown and which implements a set of operations. An experiment with a black box, called a *query*, consists of supplying an input and observing the output. The smallest number of queries necessary to determine the operation is called the *query complexity*.

**Bloch sphere** A geometrical representation of the pure state space of a two-level quantum mechanical system. We represent impure states as points inside a Bloch sphere, thus sometimes we talk about a Bloch ball.

**block code** A code where a group of information symbols is encoded into a fixed-length code word by adding a set of parity check or redundancy symbols.

**boson** Quantum particles whose spin quantum number can be $s = 1$, $s = -1$, or $s = 0$ (*e.g.*, photons and mesons). The other type of quantum particles are fermions, particles whose spin quantum number can be $s = 1/2$ or $s = -1/2$ (*e.g.*, electrons, protons, and neutrons). The spin is the quantum number characterizing the intrinsic angular momentum of the quantum particle.

**bounded distance decoding** The strategy when the $[n, k, d]$ code $\mathcal{C}$ corrects patterns of at most $e$ and no other errors. When

$$e = \lfloor \frac{d-1}{2} \rfloor$$

the bounded distance decoding is identical to minimum distance decoding. See also minimum distance decoding.

**Bounded-error Probabilistic Polynomial Time (BPP)** Term used in computational complexity theory to describe a class of probabilistic algorithms. BPP consists of polynomial time algorithms which produce the correct result with a probability at least 2/3 (or other values strictly between 1/2 and 1). We may get the wrong answer when we run such an algorithm once, but by repeatedly running the algorithm and then taking a majority vote, the probability of obtaining the correct answer can be made arbitrarily close to 1, while maintaining the polynomial running time. See also *polynomial algorithms*.

**Bounded-error Quantum Polynomial (BQP)** Term used in computational complexity theory to describe a class of probabilistic algorithms. BQP functions are all functions with the domain being the set of binary strings computable by uniform quantum circuits whose number of gates is polynomial in the number of input qubits, and which give the correct answer at least 2/3 of the time.

**bounded operator** Linear transformation $L$ between vector spaces $A$ and $B$ with the property that the ratio of the norms of $L(a)$ and $a \in A$ is bounded by the same number $k$ is $|| L(a) ||_B \leq k || a ||_A$ with $k > 0$, $\forall a \in A, a \neq 0$.

**braid group** $B_n$ the braid group on $n$ strands is an infinite group with applications in knot theory and topological quantum computing. $B_n$ has an intuitive geometric representation, it shows how $n$ strands can be laid out to connect two groups of $n$ objects.

**capacity of a classical memoryless communication channel** A discrete memoryless channel is characterized by the triplet $(X, Y, \{p(y|x), x \in \{X, y \in Y\})$ with $X$ the input channel alphabet, $Y$ the output channel alphabet, and $\{p(y|x), x \in X, y \in Y\})$ the probability of observing output symbol $y$ when the input symbol is $x$. The capacity of the discrete memoryless channel $(X, Y, \{p(y|x), x \in X, y \in Y\})$ is $C = \max_{p(x)} I(X;Y)$ with $p(x)$ the input distribution.

**capacity of a quantum memoryless communication channel** A quantum communication channel has four distinct capacities [47]: (i) $C$, the ordinary capacity for transmitting classical information; (ii) $Q$, the ordinary capacity for transmitting quantum states, typically $Q < C$; (iii) $Q_2$, the classically-assisted quantum capacity for transmitting quantum states with the assistance of a two-way classical channel, $Q \leq Q_2 \leq C$ and (iv) $C_E$, the entanglement-assisted classical capacity, $Q \leq C \leq C_E$.

**cardinality of a finite set** Let $G$ be a finite set. The cardinality of $G$, denoted as $\mid G \mid$ represents the number of elements in the set.

**cavity quantum electrodynamics (CQED)** In CQED systems individual photons circulating in a high-finesse resonator can interact strongly via their mutual coupling to a single atom positioned inside the cavity. The atom is strongly coupled to the cavity mode in a manner that allows for efficient transfer of electromagnetic fields from input to output (photon) channels. The atom-cavity system may be viewed as a quantum-optical device.

**Cauchy sequence** A sequence $\{a_n\}$ is Cauchy if for any $\epsilon > 0$ there exists $N \in \mathbb{R}$, such that $\mid\mid a_k - a_r \mid\mid \; < \epsilon$ for $k, r > N$.

**cavity** Cavities are optical resonators with dimensions much larger than the optical wavelength. They are arrangements of optical components in an "open" setup where the light is allowed to circulate either by bouncing back and forth between two end mirrors (linear cavities), or by doing round trips in two opposite directions (ring cavities). A laser cavity contains a gain medium which can compensate for the cavity losses in each round trip of the light.

**cavity modes** Cavity modes are self-consistent electric field distributions inside a cavity which are self-reproducing after each cavity round trip of the light. A cavity is *stable* if it has such modes (the modes do not always exist). The stability depends on the cavity parameters, such as curvature of reflecting surfaces, distance between optical components. Laser frequencies correspond precisely to certain mode frequencies. A laser can operate on a single mode of its cavity and then is called a single-frequency laser. Single mode operation is required, among others, to drive resonant cavities, for high resolution spectroscopy, for coherent beam combining of laser outputs. While the modes of an empty cavity are mutually orthogonal, the modes of a laser cavity are no longer so due to transversely varying gain and loss. The optical frequencies for which the optical phase is self-consistently reproduced after each round trip of the cavity (i.e., the optical phase shift is an integer multiple of $2\pi$) are called *mode frequencies* or *resonance frequencies*. The *Q factor* is the ratio of resonance frequency and bandwidth. Bandwidth is defined as the width of the resonance frequency line.

**Church-Turing principle** Every function which can be regarded as computable can be computed by a by a mechanical device equivalent to a Turing machine (Turing's Thesis) or by use of Church's $\lambda$-calculus. David Hilbert formulated in 1930s the Entscheidungsprob-

lem, which asked if there was a mechanical procedure for separating mathematical truths from mathematical falsehoods. While studying this problem Alonzo Church and Stephen Kleene introduced the $\lambda$-definable functions. Church proved that "every effectively calculable function (effectively decidable predicate) is general recursive." Turing proved that every function which could be regarded as computable is computable by a Turing machine.

**closed quantum system** An idealization of a system of quantum particles. The system is assumed to be isolated and the interaction of the particles with the environment is non-existent. In reality, we can only construct quantum systems with a very weak interaction with the environment.

**`CNOT`, `controlled-NOT` gate** A two-qubit gate. It has as input a control qubit and a target qubit. The control input is transferred directly to the control output of the gate. The target output is equal to the target input if the control input is $\mid 0\rangle$, and it is flipped if the control input is not $\mid 0\rangle$.

**coherent states** Minimum uncertainty states. For such states, $\Delta(X)\Delta(P_X) = \hbar/2$, rather than $\Delta(X)\Delta(P_X) \geq \hbar/2$. Let $\mid e_i\rangle$ be a set of basis sets. A quantum system in state $\mid \varphi\rangle = \sum_i \alpha_i \mid e_i\rangle \in \mathcal{H}_n$ is in a *coherent superposition* of basis states $\mid e_i\rangle$ if its density matrix $\rho$ is not diagonal, $\rho \neq 1/nI_n$; if, in addition, the system is in a pure state then it is said to be *completely coherent*. If $\rho$ is diagonal, $\rho = 1/nI_n$ the system is in an incoherent superposition of states; in particular, a qubit is in an incoherent superposition of states if

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In case of an incoherent superposition, the basis states are prepared independently and there is no definite phase relationship between the bases states.

**commutator of two operators** Given two operators $\mathbf{A}$ and $\mathbf{B}$, the *commutator* of the two operators is $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\ \mathbf{B}\ -\ \mathbf{B}\ \mathbf{A}$. We say $\mathbf{A}$ commutes with $\mathbf{B}$ if $[\mathbf{A}, \mathbf{B}] = 0$.

**complementarity** A principle enounced in 1928 by Niels Bohr. This basic tenant of quantum mechanics says that quantum systems exhibit a wave-particle duality. The behavior of such phenomena as light and electrons is sometimes wavelike and sometimes particle-like and it is impossible to observe simultaneously both aspects.

**complete positivity** The linear transformation $\mathcal{E}(\rho)$ enjoys *complete positivity* if there exist operators $\mathbf{A}_i$ such that

$$\mathcal{E}(\rho) = \sum_i \mathbf{A}_i \rho \mathbf{A}_i^\dagger \quad \text{with} \quad \sum_i \mathbf{A}_i^\dagger \mathbf{A}_i \leq \mathbf{I}.$$

When $\sum_i \mathbf{A}_i^\dagger \mathbf{A}_i = \mathbf{I}$ then $\mathrm{tr}\,(\mathcal{E}(\rho)) = 1$.

**complete vector space** A vector space is complete if any Cauchy sequence is convergent. See also the entry for the *Cauchy sequence*.

**concatenated codes** Codes introduced by David Forney based upon serial concatenation of two or more codes. For example the *outer code* (applied first and removed last) could be a Reed-Solomon code followed by a convolutional *inner code* (applied last, removed first). Concatenated codes encode first a block of $k$ information symbols into $n$ symbols using the outer code and then encode each one of the $n$ symbols using the inner code. Turbo codes

are a refinement of concatenated codes; they use the encoding structure of concatenated codes and an iterative algorithm for decoding. See also convolutional codes.

**convex subspace of a vector space** A subset that contains all the points of a straight line connecting any two points in the subset. The set of density operators in a Hilbert space $\mathcal{H}_n$ is a convex linear subset of $n \times n$ Hermitian matrices.

**convolutional codes** Concatenated codes do not encode blocks of input symbols but a continuous input steam. They are primarily used to maximize the transmission rate in case of satellite communication and remote sensing.

**conservative logic gate** A logic gate that conserves the number of 1s at its input. For example, the `Fredkin` gate is a conservative gate.

**convex set** A convex set $S$ in a vector space $V$ over the field $\mathbb{R}$ is one when the line segment joining any two points in $S$ lies entirely in $S$.

**Copenhagen interpretation of quantum mechanics** The Copenhagen school is the name given to the group of theoreticians who shared the views of Niels Bohr regarding quantum mechanics interpretations.

**correspondence rule in Copenhagen interpretation** Heuristic principle requiring that in such cases when the influence of Plank's constant could be neglected the numerical values predicted by quantum mechanics should be the same as if they were predicted by classical radiation theory.

**coset of a group** For a subgroup $H$ of a group $G$, $H \subset G$, and an element $x \in G$; define $xH$ to be the set $\{xh : \ h \in H\}$ and $Hx$ to be the set $\{hx : \ h \in H\}$. A subset of $G$ of the form $xH$ for some $x \in G$ is said to be a *left coset of* $H$ and a subset of the form $Hx$ is said to be a *right coset* of H. For any subgroup H, we can define an equivalence relation ($\equiv$) by $x \equiv y$ if $x = yh$ for some $h \in H$. The equivalence classes of this equivalence relation are exactly the left cosets of $H$, and an element $x \in G$ is in the *equivalence class* $xH$.

**Coulomb blockade** A result of the discreteness of the electron charge, it manifests itself as the increased resistance at small bias voltage of a low-capacitance tunnel junction. The transfer by tunnelling of one electron through the junction barrier increases the tunnel junction capacitor with one elementary charge; that causes a voltage buildup $V = e/C$ and an increase in electrostatic energy of the system by an amount $E = e^2/(2C)$, where $e$ is the electron charge of $1.6 \times 10^{-19}$ coulomb and $C$ is the capacitance of the junction. For a small capacitance, the voltage buildup can be large enough to prevent tunnelling by another electron. The electric current is suppressed and the resistance of the junction is no longer constant for as long as the bias voltage is raised but does not reach the buildup value $V$. The Coulomb blockade can be observed experimentally only at low temperature where the charging energy $E$ is larger than the thermal energy $kT$ of the charge carriers, that is, $e^2/2C \gg kT$, where $k$ is Boltzman's constant and $T$ is the absolute temperature. For typical capacitances of $C \leq 10^{-15}$ farad this regime can be reach at temperatures of the order of milliKelvin and the discreteness of the electron energy spectrum has to be taken into account.

**CSS, Calderbank-Shore-Steane, codes** Quantum codes formed from two classical codes.

**cyclic code** Code invariant under the cyclic permutation of bits or qubits.

**cyclic redundancy check (CRT)** Error-detecting code; the parity check symbols are computed over the characters of the message and are then appended to the packet by the networking hardware.

**cyclic Discrete Fourier Transform,** $DFT$ If $n$ divides $2^k - 1$ and $\beta$ is an element of order $n$ in $GF(2^k)$, then the cyclic $DFT$ over $GF(p^k)$ is described by the matrix

$$DFT = (\beta^{ij})_{0 \leq i,j \leq n-1} \quad \text{with} \quad \beta^n = 1.$$

The inverse cyclic Discrete Fourier Transform is denoted as $DFT^{-1}$.

**data processing inequality** Informally, the data processing inequality states that one cannot get more information by processing a set of data than the information provided by the data to begin with. If the random variables $X \mapsto Y \mapsto Z$ form a Markov chain then

$$I(X;Z) \geq I(X;Y).$$

Random variables $X$, $Y$, and $Z$ form a Markov chain if the conditional probability distribution of $Z$ depends only upon $Y$ and it is independent of $X$:

$$p_{XYZ}(x,y,z) = p_X(x)p_{Y|X}(y|x)p_{Z|Y}(z|y) \quad \Longrightarrow \quad X \mapsto Y \mapsto Z.$$

**de Broglie, Louis Victor Pierre Raymond, duke (1892 - 1987).** Distinguished French physicist who speculated that nature did not single out light as being the only matter which exhibits a wave-particle duality. He proposed that ordinary "particles" such as electrons, protons, or bowling balls could also exhibit wave characteristics in certain circumstances. Quantitatively, he associated a wavelength $\lambda$ to a particle of mass $m$ moving at speed $v$ the two being related by *de Broglie's equation*

$$p = \frac{h}{\lambda}$$

with $p$ the momentum of the particle and $h$ Planck constant

**decoherence** The most significant cause of errors in quantum systems. The destruction of the superposition of pure quantum states due to the interaction of the quantum system with the environment. At large scale the world obeys the laws of classical Physics due to decoherence.

**degenerate code** A quantum code for which linearly independent correctable errors acting on the coding space produce linearly dependent states. Degenerate codes have the potential to be more efficient than non-degenerate codes because they are better than many known bounds on efficiency suggest.

**density matrix of a quantum system** The density matrix or operator is a simplifying notation to represent pure or mixed states. For a pure state $\mid \varphi\rangle$, the density operator is $\rho = \mid \varphi\rangle\langle\varphi \mid$; for mixed states, it is a probabilistic expression $\rho = \sum_i \mu_i \mid \varphi_i\rangle\langle\varphi_i \mid$ with $\sum_i \mu_i = 1$, and with more than one $\mu_i \neq 0$.

**depolarizing channel model** The quantum error process is memoryless and the probability of error is $p$, thus $\text{Prob}(\sigma_I) = 1 - p$ and $\text{Prob}(\sigma_x) = \text{Prob}(\sigma_y) = \text{Prob}(\sigma_z) = p/3$.

**Descartes, Rene (1556 - 1650).** Very influential French mathematician and philosopher of the XVII century. He founded analytical geometry and invented the Cartesian coordinate

system. Descartes is also a major promoter of rationalism, a philosophical method later embraced by Baruch Spinoza and Gottfried Leibnitz. Famous also for the statement "Cogito ergo sum," Latin for "I think, therefore I exist."

**Deutsch, David** Distinguished physicist born in Haifa, Israel, and educated at Cambridge and Oxford Universities in the UK. Member of the Quantum Computation and Cryptography group at Clarendon Laboratory at Oxford University, author of the theory of parallel universes.

**Deutsch's principle** Extends the Church-Turing principle. Every finitely realizable physical system can be perfectly simulated by a universal computing machine operating by finite means.

**diameter of the electron** According to high-energy electron-electron scattering, the diameter of the electron is $< 10^{-18}$ m.

**dilution refrigerator** The operation of the dilution refrigerator is based on the phase transition a liquid mixture of two helium isotopes, $^3$He and $^4$He undergoes at temperatures below 0.86 K, the triple critical point, the point in the phase diagram of liquid He where the three boundary lines between "normal liquid", "superfluid liquid", and "unstable composition" meet. When cooled below this critical point, the He mixture undergoes a spontaneous phase separation and forms a $^3$He-rich phase and a $^3$He-poor phase. The transport of $^3$He atoms from the $^3$He-rich phase into the $^3$He-poor phase requires energy. If atoms are made to continuously cross this boundary, they cool the mixture. Temperatures as low as 2 miliKelvin (mK) can be reached with the best dilution systems.

**Dirac, Paul Adrien Maurice (1902 - 1984).** Distinguished British mathematician. His work has been concerned with the mathematical and theoretical aspects of quantum mechanics. He used a noncommutative algebra for calculating atomic properties leading to the relativistic theory of the electron (1928) and the theory of holes (1930). This latter theory required the existence of a positive particle having the same mass and charge as the known (negative) electron. This particle, the positron, was discovered experimentally at a later date (1932) by C. D. Anderson. The importance of Dirac's work lies essentially in his famous wave equation, which introduced special relativity into Schrödinger's equation. In 1932, he became Lucasian Professor of Mathematics at Cambridge. In 1933, Dirac shared the Nobel Prize for Physics with Schrödinger.

**Dirac's `ket` and `bra` notation** Notation used in quantum mechanics for state vectors.

**discrete memoryless classical communication channel** Channel over a finite alphabet and with a finite number of states. The memoryless property of the communication channel implies that the output of the channel is a Markov process, it is affected only by the current input and not by the history of the channel states.

**discrete memoryless quantum communication channel** The quantum systems are described as vectors in finite-dimensional Hilbert spaces. The memoryless property of the communication channel implies that the output of the channel is affected only by the current input and not by the history of the channel states.

**distinguishable states of a quantum system** Two states of a quantum system are distinguishable if they are orthogonal. If two states are distinguishable, then a measurement exists that guarantees to determine which one of the two states the system is in.

**dual operator/matrix** $A^\dagger$, the dual, or the adjoint of a matrix $A$, is obtained by trans-

posing the matrix and then taking the complex conjugate of each element. The order of the two operations can be reversed. See also the entry for *adjoint matrix*.

**efficient computation** A computation is efficient if it requires, at most, polynomial resources as a function of the size of the input. For example, if a computation returns the value of the function $f(x)$, with $x$ as a bit string of length $n$, then the computation of $f(x)$ is efficient if the number of steps it requires is bounded by $n^k$ for some $k$.

**eigenvalue** A scalar, $\lambda_i$, associated with an *eigenvector*, $\mid \psi_i\rangle$, of a linear operator (observable), $\mathbf{O}$, as

$$\mathbf{O} \mid \psi_i\rangle = \lambda_i \mid \psi_i\rangle.$$

In quantum mechanics, the eigenvalues of an operator represent those values of the corresponding observable that have non-zero probability of occurring. The set of all the eigenvalues is called the operator (matrix) *spectrum*.

**eigenvector** A state vector, $\mid \psi_i\rangle$, is an eigenvector of a linear operator (observable), $\mathbf{O}$, if when operated on by the operator the result is a scalar multiple of itself as

$$\mathbf{O} \mid \psi_i\rangle = \lambda_i \mid \psi_i\rangle.$$

The scalar $\lambda_i$ is called the *eigenvalue* associated with the eigenvector. If $\mid \psi_i\rangle$ is an eigenvector with eigenvalue $\lambda_i$, then any non-zero multiple of $\mid \psi_i\rangle$ is also an eigenvector with eigenvalue $\lambda_i$. If the set of state vectors $\mid \psi_0\rangle, \mid \psi_1\rangle, \ldots, \mid \psi_{n-1}\rangle$ are eigenvectors to *different* eigenvalues $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$, then the state vectors $\mid \psi_0\rangle, \mid \psi_1\rangle, \ldots, \mid \psi_{n-1}\rangle$ are necessarily linear independent.

**empiricism** Theory of knowledge whose main tenet is that scientific knowledge is related to experience and that experiments are essential for the development of scientific concepts, all hypotheses and theories must be tested against observations of the natural world. The Sophists in the 5th century BC, Aristotle (384 BC - 322 BC), the Stoics and the Epicureans, a generation later, followed by Thomas Aquinas and Roger Bacon in the XIII century, and John Locke, George Berkeley, David Hume, in the XVII and XVIII century, attribute sensory perception and a posteriori observation a major role in the formation of human knowledge. They believed that human knowledge of the natural world is grounded in sense experience. (see also rationalism)

**encoding of a logical qubit into multiple physical qubits.** We encode the state of a qubit $\mid \psi\rangle = \alpha_0 \mid 0\rangle + \alpha_1 \mid 1\rangle$ as $\mid \varphi\rangle = \alpha_0 \mid 0_L\rangle + \alpha_1 \mid 1_L\rangle$ with $0_L$ and $1_L$ resilient to errors. For example, we encode $\mid 0\rangle$ as $\mid 000\rangle$ and $\mid 1\rangle$ as $\mid 111\rangle$. In this case, a random error can cause departures from the subspace spanned by $\mid 000\rangle$ and $\mid 111\rangle$. We shall be able to correct small errors because the component which was $\mid 000\rangle$ is likely to remain in a sub-space spanned by the four vectors, $\mid 000\rangle, \mid 001\rangle, \mid 010\rangle$, and $\mid 100\rangle$, while the component which was $\mid 111\rangle$ is likely to remain in a sub-space spanned by the four vectors, $\mid 111\rangle, \mid 110\rangle, \mid 101\rangle$, and $\mid 011\rangle$.

**entanglement** Nonlocal, nonclassical correlation between two quantum system which gives quantum computers and quantum communication systems their power. The translation of the German term *Verschränkung* used by Schrödinger, who was the first to recognize this quantum effect. It means that a two-particle quantum system is in a state that cannot be written as a tensor product of the states of the individual particles. The two quantum particles share a joint state, and it is not possible to describe one of the particles in isolation.

**entanglement fidelity** A measure of the overlap between the initial purification and the joint state of the quantum system $Q$ and the reference system $R$ after the transmission of $Q$ through a quantum communication channel

$$F_e(\rho, \mathcal{E}) = \langle \varphi^{(Q,R)} \mid (\mathcal{I}_R \otimes \mathcal{E}) \left( \mid \varphi^{(Q,R)} \rangle \langle \varphi^{(Q,R)} \mid \right) \mid \varphi^{(Q,R)} \rangle$$

with $\mid \varphi^{(Q,R)} \rangle$ the joint state of the quantum system $Q$ and the reference system $R$ before the transmission and $\mathcal{I}_R$ the identity transformation of the reference system $R$.

**entropy** Measure of the uncertainty, or the degree of disorder of a system.

**entropy exchange** If $\mathcal{E}$ is the quantum transformation of the input state $\rho$, carried out during transport through a quantum communication channel then the entropy exchange is defined as the von Neumann entropy of the environment after the operation [276, 371]:

$$S_e(\rho, \mathcal{E}) = S(\rho_{final}^{(E)}).$$

Initially the environment $E$ is in a pure state as a result of a purification operation; $\rho_{final}^{(E)}$ is the density matrix of $E$ after the operation.

**Entscheidungsproblem** German term for the "decision problem" posed by Hilbert: "Could there exist, at least in principle, a definite method or process, by which it could be decided whether any mathematical assertion was provable?"

**EPR paradox** Experiment proposed by Einstein, Podolsky, and Rosen in 1935 to show that quantum mechanics is not a complete theory.

**ergodic process** Stochastic process for which time averages and set averages are equal to one another.

**ergodicity** An attribute of stochastic systems; generally, a system that tends in probability to a limiting form that is independent of the initial conditions.

**Euler, Leonhard (1707 - 1783).** Eminent mathematician born in Basel, Switzerland who introduced new concepts in analysis and revised almost all of the branches of pure mathematics known at that time. In 1748, he wrote *Introductio in Analysin Infinitorum*, an introduction to pure analytical mathematics, and in 1755, the *Institutiones Calculi Differentialis*. Euler's formulae:

$$\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta}), \qquad \cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}).$$

**extreme point of a convex set** The extreme point of a convex set $S$ is a point in the set $S$ which does not lie in any open line segment joining two points of $S$.

**fanout** A measure of the ability of a logic gate output to drive a number of inputs of other logic gates of the same type to form more complex circuits.

**Fast Fourier Transform (FFT)** FFT is an algorithm proposed by J. W. Cooley and J. W. Tukey in 1965, which reduces the number of operations to compute the Fourier Transform from $2n^2$ to $2n \log_2 n$. A similar idea is used for the Quantum Fourier Transform. The algorithm decomposes the transformation for $n = 2^m$, recursively, into two transforms

of length $n/2$ using the identity

$$\sum_{j=0}^{n-1} a_j \; e^{-i2\pi jk/n} = \sum_{j=0}^{n/2-1} a_{2j} \; e^{-i2\pi(2j)k/n} + \sum_{j=0}^{n/2-1} a_{2j+1} \; e^{-i2\pi(2j+1)k/n}$$
$$= \sum_{j=0}^{n/2-1} a_j^{even} e^{-i2\pi jk/(n/2)} + e^{-i2\pi k/n} \sum_{j=0}^{n/2-1} a_j^{odd} e^{-i2\pi jk/(n/2)}.$$

**Fermi, Enrico (1901 - 1954).** Distinguished Italian-born physicist who discovered in 1926 the statistical laws, nowadays known as the "Fermi statistics", governing the particles subject to Pauli's exclusion principle (now referred to as *fermions*, in contrast with *bosons* which obey the Bose-Einstein statistics). Fermi directed a classical series of experiments which ultimately led to the atomic pile and the first controlled nuclear chain reaction. He played an important part in solving the problems connected with the development of the first atomic bomb.

**Fermi level, Fermi energy** The Fermi level is the last energy level occupied by electrons at the zero absolute temperature, $T = 0$ K. According to the Pauli exclusion principle, at $T = 0$ K the electrons of an atom fill all available levels starting with the lowest one. The top of this sea of electrons, also called the "Fermi sea", defines the *Fermi level*, or the *Fermi energy*. The Fermi energy, $E_F$, is the maximum energy occupied by an electron at $T = 0$ K. The Fermi energy is very large in comparison with the energy that an electron could gain by ordinary physical interactions with the environment. For example, the Fermi energy for copper is $E_F = 7$ eV, while the thermal energy acquired by an electron at $T = 300$ K, the room temperature, is $kT = 0.026$ eV, where $k$ is Boltzmann's constant.

**fermions** Quantum particles with a spin multiple of one-half (*e.g.*, electrons, protons, and neutrons). The spin quantum number of fermions can be $s = 1/2$, $s = -1/2$. The spin of a complex system of fermions is a multiple of $\pm 1/2$. See also *bosons*.

**Feynman, Richard Phyllips (1918 - 1988).** Distinguished American physicist who received his doctorate from Princeton in 1942 under J. A. Wheeler (he was also advised by E. Wigner) and worked on the atomic bomb project at Princeton University (1941 - 1942) and then at Los Alamos (1943 - 1945). Feynman's main contribution was to quantum mechanics; he introduced diagrams (now called Feynman diagrams) that are graphic analogues of the mathematical expressions needed to describe the behavior of systems of interacting particles. He was awarded the Nobel Prize in 1965, jointly with J. Schwinger and S-I. Tomonoga, for fundamental work in quantum electrodynamics and physics of elementary particles; his later work led to the current theory of quarks, fundamental in pushing forward an understanding of particle physics. He made significant contributions to the field of quantum computing as well. According to his obituary published in the Boston Globe, "He was widely known for his insatiable curiosity, gentle wit, brilliant mind and playful temperament."

**fidelity of a pure quantum state** Let $\mid \varphi\rangle$ be a pure state of a quantum system. The fidelity of a possibly mixed state with density matrix $\rho$ and $\mid \varphi\rangle$ is defined as

$$F = \langle \varphi \mid \rho \mid \varphi \rangle.$$

**fidelity of two operators** The fidelity of positive operators **A** and **B** is defined as $F(A, B) = \mathrm{tr}\sqrt{A^{1/2}BA^{1/2}}$.

**fidelity of two probability density functions** The fidelity of two probability density functions, $p_X(x)$ and $p_Y(x)$ is defined as $F(p_X(x), p_Y(x)) = \sum_x \sqrt{p_X(x)p_Y(x)}$.

**forward classical communication costs** FCCC measures the capacity of the classical channel needed to simulate a quantum channel.

**Fredkin gate** A three-qubit gate with two target inputs, $a$, $b$, and one control input, $c$, and three outputs, $a'$, $b'$, and $c' = c$. When the control input is not set, $c = 0$, and the target inputs appear unchanged at the output, $a' = a$ and $b' = b$. When the control is set, $c = 1$, and the target inputs are swapped, $a' = b$ and $b' = a$. The control input is always transferred to the output unchanged. There are both classical and quantum versions of the `Fredkin` gate.

**functional** A real-valued function on a vector space $V$.

**fundamental theorem of algebra** Any polynomial equation of degree $n$ has $n$ solutions.

**fundamental theorem of arithmetic** Any positive integer $n > 1$ can be expressed as a product of prime numbers with unique non-negative exponents as

$$n = 2^{n_2} \cdot 3^{n_3} \cdot 5^{n_5} \cdot \ldots = \prod_{p \text{ prime}} p^{n_p}.$$

**Galois, Évariste (1811 - 1832)** French mathematician who developed the Galois theory and articulated first the concept of finite field; he also and introduced the term *group* for the well-known algebraic structure.

**Galois field** A finite field with $q$ elements is also called a *Galois field*, and it is denoted as $GF(q)$. In particular, $\mathbb{Z}_p$ is denoted as $GF(p)$. It is easy to prove that a finite field has $p^n$ elements with $p$ a prime number.

**Gauss-Markov process** A stochastic process $X(t)$ with three properties:

1. If $h(t)$ is a non-zero scalar function of t, then $Z(t) = h(t)X(t)$ is also a Gauss-Markov process.

2. If $f(t)$ is a non-decreasing scalar function of t, then $Z(t) = X(f(t))$ is also a Gauss-Markov process.

3. There exists a non-zero scalar function $h(t)$ and a non-decreasing scalar function $f(t)$ such that $X(t) = h(t)W(f(t))$, where $W(t)$ is the standard Wiener process. Every Gauss-Markov process can be synthesized from the standard Wiener process.

See also the entry for *standard Wiener process*.

**gate (quantum gate)** Building block of a quantum circuit. A physical system capable of transforming one or more qubits.

**Gedanken experiment** "Gedanken" is the German word for "thought." A thought experiment enables us to prove or disprove a conjecture when the experiment enabling us to study the physical phenomena is not feasible. We construct the result of the thought

experiment according to a set of assumptions and a model of the system: all based upon universally accepted laws of Physics.

**generators of a group** A set of $\mu$ elements $\langle g_1, g_2, \ldots, g_\mu \rangle$ is called the generator of the group $G$ if the element $g_i$ can be written as a product of (possibly repeated) elements from the list $\langle g_1, g_2, \ldots, g_\mu \rangle$, $\forall g_i \in G$. We write $G = \langle g_1, g_2, \ldots, g_\mu \rangle$. The number $\mu$ of generators of group $G$ satisfies the inequality: $\mu \leq \log(\mid G \mid)$.

**General Linear Group** Multiplicative group of non-singular $n \times n$ matrices. Denoted as $GL(n)$ or as $GL(n, \mathbb{R})$, $GL(n, \mathbb{C})$, $GL(n, F)$ if matrices have real, complex, or elements from a field $F$ $a_{ij} \in F$, respectively.

**generators of a stabilizer group** Let $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_{n-k}$ be the generators of the stabilizer group $S$ of a quantum code $\mathcal{Q}$. The eigenvectors of the generators $\{\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_q\}$ have special properties: those corresponding to eigenvalues of $+1$ are the codewords of $\mathcal{Q}$ and those corresponding to eigenvalues of $-1$ are codewords affected by errors. If $\mathbf{M}_j \in M$ and if $\mathbf{M}_j \mid \psi_i \rangle = (+1) \mid \psi_i \rangle$ then $\mid \psi_i \rangle$ is a codeword, $\mid \psi_i \rangle \in \mathcal{Q}$. This justifies the name given to the set $S$, any operator in $S$ leaves the state of a word unchanged. On the other hand, if $\mathbf{M}_j \mid \varphi_k \rangle = (-1) \mid \varphi_k \rangle$ then $\mid \varphi_k \rangle = \mathbf{E}_\alpha \mid \psi_k \rangle$, the state $\mid \varphi_k \rangle$ is the result of an error affecting the codeword $\mid \psi_k \rangle \in \mathcal{Q}$.

**generating function** A formal power series whose coefficients encode information about a sequence $a_n$ indexed by natural numbers $G(a_n; x) = a_n x^n$. For example

$$G(n^2; x) = \sum_{n=0}^{\infty} n^2 x^2 = \frac{x(x+1)}{(1-x)^3}.$$

Informally, a generating function is a "clothline which we hang up a sequence of numbers to display."

**Gilbert-Varshamov bound** If $\mathcal{C}$ is an $[n, k, d]$ linear code over $GF(q)$ and if $M(n, d) = \mid \mathcal{C} \mid$ is the largest possible number of codewords for a given $n$ and $d$ values, then the code satisfies the Gilbert-Varshamov bound

$$M(n, d) \geq \frac{q^n}{\sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k}.$$

**Gleason theorem** Given a Hilbert space $\mathcal{H}_n$ of dimension $n \geq 3$, the only possible measure of the probability of the state associated with a particular linear subspace $\mathcal{H}_m,\ m < n$, of the Hilbert space $\mathcal{H}_n$ is the trace of the operator product of the projection operator $\mathbf{E}$ that projects and the density matrix $\rho$ of the system, $\mathrm{tr}\,[\rho \mathbf{E}]$.

**Hadamard, Jacques (1865 - 1963).** Distinguished French mathematician who proved the prime number theorem, developed Hadamard matrices, and worked on the calculus of variations.

**Hadamard gate** The $H$ gate describes a unitary quantum "fair coin flip" performed upon a single qubit. For example, it transforms an input qubit in state $\mid 0 \rangle$ into a superposition state $(\mid 0 \rangle + \mid 1 \rangle)/\sqrt{2}$, or $(\mid 0 \rangle - \mid 1 \rangle)/\sqrt{2}$. The transfer matrix of a `Hadamard` gate is

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

**Hall, Edwin (1855 - 1938)** American physicist who discovered the Hall effect in 1879.

**Hall effect** The classical effect, discovered by Edwin Hall in 1879, manifests itself as a potential difference, the *Hall voltage*, on the opposite sides of a conductor when an electric current is flowing through the conductor and a magnetic field is applied perpendicular to the current. Electrons moving along a direction perpendicular to an applied magnetic field are acted upon by the Lorentz force oriented normal to both the magnetic field direction and the electron motion direction. In the absence of the magnetic field, the electrons flow in a straight line; when the magnetic field is applied the electrons are deflected, their new path is curved so that the moving charges accumulate on one face of the conductor. Equal charges of opposite sign accumulate on the other face. The separation of charges results in the build up of an electric potential drop (Hall voltage) across the sample.

The Hall voltage for a metal which is characterized by only one type of charge carrier, the electrons, is given by

$$V_{\text{Hall}} = -\frac{IB}{dne},$$

where $I$ is the current flowing along the length of the sample, $B$ is the magnetic flux density, $d$ is the sample thickness, $n$ is the density of the charge carriers (electrons), and $e$ is the electron charge ($1.602 \times 10^{-19} Coulomb$). Other Hall effect characteristics are the Hall coefficient and the Hall resistance defined, respectively, as

$$R_{\text{Hall}} = \frac{V_{\text{Hall}} d}{IB} = -\frac{1}{ne} \quad \text{and} \quad \rho_{\text{Hall}} = \frac{V_{\text{Hall}}}{I}.$$

The Hall effect has the ability to differentiate between positive and negative charges which move in opposite directions and at one time it was the first proof that the electric current in metals is carried by electrons and not by protons. The classical Hall effect can be observed at room temperature and at moderate magnetic fields of less than one tesla (T). The Hall resistance, $\rho_{\text{Hall}}$, varies linearly with the strength of the magnetic field, $B$.

The Hall effect is also present in semiconductors where it differentiates between the two types of carriers, electrons and holes. The Hall voltage has opposite signs for these two carriers, such as negative for $n$-type semiconductors and positive for $p$-type semiconductors. The Hall coefficient for moderate magnetic fields takes into account the contributions of both carriers, electrons and holes, which have different mobilities and are present in different concentrations; it is given by

$$R_{\text{Hall}} = \frac{-n\mu_e^2 + p\mu_h^2}{e(n\mu_e + p\mu_h)^2}$$

with $n$ - the electron concentration, $\mu_e$ - the electron mobility, $p$ - the hole concentration, $\mu_p$ - the hole mobility, and $e$ - the electron charge. The *Hall mobility* can be determined from the equation

$$\mu = \frac{\mid V_{\text{Hall}} \mid}{R_{\text{sheet}} IB}$$

where all the quantities involved can be measured experimentally. The Hall coefficient for large applied magnetic fields takes a simpler form, similar to that for a single type of carrier

$$R_{\text{Hall}} = \frac{1}{(n-p)e}.$$

**Hamiltonian** The name *Hamiltonian* honors the Irish mathematician and astronomer William R. Hamilton who, personally, had nothing to do with quantum mechanics, but in the 1830s happened to restate Lagrange's equations of motion with emphasis on momenta instead of forces. Hamilton's equations are first order differential equations with respect to time and involve the Hamiltonian function **H** which is the total energy expressed as a function of the generalized coordinates for position, $q_i$, and momenta, $p_i$,

$$\frac{d}{dq_i} = \frac{\partial \mathbf{H}}{\partial p_i}, \qquad \frac{d}{dp_i} = -\frac{\partial \mathbf{H}}{\partial q_i}.$$

**Hamming, Richard Wesley (1915 - 1998).** American mathematician best known for his work on error-detecting and error-correcting codes. His fundamental paper on this topic appeared in 1950.

**Hamming bound** Given an $(n, k, d)$ linear code with $n$ the length of the codeword, $k$ the number of information symbols, $r = n - k$ the number of parity check symbols, and $d$ the distance of the code, the Hamming bound gives the minimum number of parity check bits necessary to construct a code with a certain error-correction capability (*e.g.*, a code capable to correct all single-bit errors)

$$2^r \geq \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i}.$$

**Hamming distance** The number of positions where two code words differ.

**Hamming sphere** All $n$-tuples of a block code which are not codewords and are at distance $d \leq e$ from a codeword.

**Heisenberg, Werner (1901 - 1976).** Eminent German physicist; the founder of quantum mechanics. Heisenberg and his fellow student Wolfgang Pauli started their study of theoretical physics under Arnold Sommerfeld in 1920. In $1924 - 1925$, he worked with Niels Bohr at the University of Copenhagen. In 1925, Heisenberg formulated matrix mechanics, the first coherent mathematical version of quantum mechanics. Matrix mechanics was further developed in 1926 in a paper co-authored with M. Born and P. Jordan. He is perhaps best known for the *Uncertainty Principle*, discovered in 1927. In 1928, Heisenberg published *The Physical Principles of Quantum Theory.* In 1932, he was awarded the Nobel Prize in Physics for the creation of quantum mechanics.

**Heisenberg uncertainty principle** Intrinsic property of the quantum systems: the precise knowledge of some basic physical properties such as position and momentum is simply forbidden. The *Uncertainty Principle* states that the uncertainty in determining the position $\Delta X$ and the uncertainty in determining the momentum $\Delta P_X$ at position $X$ are constrained by the inequality

$$\Delta X \; \Delta P_X \geq \frac{\hbar}{2}$$

where $\hbar = h/2\pi$ is a modified Planck constant.

**Hertz, Heinrich (1857 - 1894).** German physicist, the first to demonstrate experimentally the production and detection of Maxwell electromagnetic waves. The photoelectric effect was first discovered accidentally in 1887 by Hertz, while carrying on investigations on the electromagnetic waves.

**hidden variable theory** A theory based on the assumption that there is a variable or variables that determine the real properties of quantum particles. These variables have definite values from the moment the particle is created and they determine the result of the measurement performed on that property of the quantum particle.

**Hilbert, David (1862 - 1943).** Eminent German mathematician; in 1899 he published *Grundlagen der Geometrie*, putting geometry in a formal axiomatic setting. He delivered the speech "The Problems of Mathematics" at the Second International Congress of Mathematicians in Paris, challenging mathematicians to solve fundamental questions such as: continuum hypothesis, the well ordering of the reals, Goldbach conjecture, the transcendence of powers of algebraic numbers, the Riemann hypothesis, and the extension of the Dirichlet principle. Hilbert's work in integral equations led to the research in functional analysis and established the basis for his work on infinite-dimensional space, later called Hilbert space.

**Hilbert space, $n$-dimensional** An $n$-dimensional complex vector space denoted as $\mathcal{H}_n$ with an inner product and thus, with a norm. An $n$-dimensional Hilbert space is isomorphic with $\mathbb{C}^n$.

**Holevo bound** A source generates letters $X$ from an alphabet with probabilities $p_i$; Alice encodes the random variable $X$ and transmits it through a quantum channel; Bob performs a measurement of the quantum system and the result of this measurement is a random variable $Y$. The mutual information, $I(X;Y)$, measures the amount of information about $X$ that Bob can infer from the measurement. Regardless of the measurement performed by Bob, the mutual information is limited by the Holevo bound

$$I(X;Y) \leq S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i).$$

**Holevo-Schumacher-Westmorland (HSW) noisy quantum channel encoding theorem** The HSW theorem was proven independently by Holevo [212] and by Schumacher and Westmorland [374].

**impure states of a single qubit** Impure, or mixed states, are represented by points inside the Bloch sphere. This implies that the trace of the square of their density matrix is less than one; $\mathrm{tr}(\rho^2) < 1$. See also *pure states of a single qubit*.

**index of a subgroup** Let $G$ be a group with cardinality $\mid G \mid$, and $H \subset G$ be a subgroup of $G$ with cardinality $\mid H \mid$. Then the index of the subgroup $H$ is

$$I_{G/H} = \frac{\mid G \mid}{\mid H \mid}.$$

**indistinguishability in quantum mechanics** Principle of quantum mechanics - all quantum particles of the same type are alike; for example, we cannot distinguish an electron from another. Therefore the operation of swapping the position of two electrons in a systems with many electrons leaves the system's state unchanged, or, in other words, the operation is symmetric and it is represented by a unitary transformation acting upon the wave function. In three dimensions an exchange of two bosons is represented by an identity operator; the wave function is invariant and we say that the particle obey the Bose statistics. The

exchange of two fermions in three dimensions changes the sign of the wave function (it is represented by multiplication with $(-1)$); the particles are said to obey Fermi statistics.

**interference** The interaction of optic, acoustic, or electromagnetic waves that are correlated, or coherent with each other; interference occurs if the waves come from the same source, or if the waves have about the same frequency. In quantum mechanics consider a transition between two states in the Hilbert space $\mathcal{H}_n$ with an orthonormal basis $\{\mid e_i\rangle\}$

$$\mid \varphi\rangle = \sum_i \alpha_i \mid e_i\rangle \quad \mapsto \quad \mid \psi\rangle = \sum_i \beta_i \mid e_i\rangle$$

The probability of this transition is the square of the modulus of the inner product of the two states

$$\text{Prob}\,(\mid \varphi\rangle \mapsto \mid \psi\rangle) = \mid \langle \varphi \mid \psi\rangle \mid^2 = \mid \sum_i \alpha_i^* \beta_i \mid^2 = \sum_{i,j} \alpha_i^* \alpha_j \beta_j^* \beta_i$$

The last sum can be separated into two terms thus the expression becomes

$$\text{Prob}\,(\mid \varphi\rangle \mapsto \mid \psi\rangle) = \sum_i \mid \alpha_i \mid^2 \mid \beta_i \mid^2 + \sum_{i,j,i\neq j} \alpha_i^* \alpha_j \beta_j^* \beta_i.$$

The first term corresponds to a classical interference; the second term reflects the quantum interference between $i \neq j$ paths of quantum evolution and the quantum probability rule.

**irreversible/non-invertible gate** A gate characterized by the fact that, knowing the output, we cannot determine the input for all possible combinations of input values. The irreversibility of classical gates, other than `NOT`, means that there is an irretrievable loss of information, and this has very serious consequences regarding the energy consumption of classical gates.

**isomorphism** The word derives from the Greek *iso*, meaning "equal" and *morphosis*, meaning "to form" or "to shape." Formally, an isomorphism is bijective morphism. Informally, an isomorphism is a map which preserves sets and relations among elements. A space isomorphism is a vector space in which addition and scalar multiplication are preserved. Two groups, $G_1$ and $G_2$ with binary operators "+" and "×", are isomorphic if there exists a map, $f : G_1 \mapsto G_2$, which satisfies the relation $f(x + y) = f(x) \times f(y)$. An isomorphism preserves the identities and inverses of a group. See also *automorphism and homomorphism.*

**Kant, Immanuel (1724 - 1804).** German philosopher, the last major exponent of the *Enlightenment* and author of the "Critique of Pure Reason." His work bridges the gap between Rationalism and Empiricism. He defined Enlightenment as an age shaped by the motto "Sapere aude," in Latin, meaning "Dare to know." He expressed the belief that we are never able to transcend the bounds of our own mind, we cannot access the "Ding an sich," in German, the "thing-in-itself."

**kernel of a linear transformation** $\text{Ker}(\mathcal{L})$, the kernel of a linear transformation $\mathcal{L} : \mathcal{H}_a \mapsto \mathcal{H}_b$ is the set of all vectors $\mid \zeta\rangle \in \mathcal{H}_a$ such that $\mathcal{L}(\mid \zeta\rangle) = 0$.

**key distribution** Mechanism for distribution of cryptographic keys; in particular, of public keys.

**Lagrange theorem** If $H$ is a subgroup of a finite group $G$, then $\mid H \mid$ divides $\mid G \mid$.

**Landauer principle** Landauer principle traces the energy consumption in a computation to the act of erasing information. When a computer erases one bit of information, the amount of energy dissipated into the environment is at least $k_B T \ln(2)$, with $k_B$ (Boltzmann's constant) and $T$ (the temperature of the environment). An equivalent formulation: the entropy of the environment increases by at least $k_B \ln(2)$ when one bit of information is erased.

**light** A visible form of electromagnetic radiation. An electromagnetic field consists of an electric and a magnetic field perpendicular to each other and oscillating in a plane perpendicular to the direction of propagation of the electromagnetic wave.

**linear code** A subspace of a vector space; a selected subset of $n$-tuples satisfying some distance properties.

**Laurent series** A complex function $f(z)$ can be represented as a power series with terms of negative degree as

$$f(z) = \sum_{n=-\infty}^{n=+\infty} a_n (z-c)^n.$$

The power series is defined with respect to point $c$ and a path of integration $\gamma$. The path of integration is counterclockwise and must lie in an annulus inside of which f(z) is holomorphic. The constant $a_n$ is

$$a_n = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)dz}{\{(z-c)^{n+1)}}$$

The Laurent series expansion is used when the Taylor series expansion cannot be applied.

**manifold, smooth manifold** A smooth manifold is a geometrical object which is locally the same as the Euclidian space $\mathbb{R}^n$ [269]. Although manifolds resemble Euclidean spaces near each point, the global structure of a manifold may be more complicated. A manifold is typically endowed with a differentiable structure that allows calculus and a Riemannian metric that allows to measure distances and angles. For example, a line and a circle are one-dimensional manifolds, a plane and sphere (the surface of a ball) are two-dimensional manifolds; every point of an $n$-dimensional manifold has a neighborhood homeomorphic to the n-dimensional space $\mathbb{R}^n$.

**matrix exponentiation** If $\beta$ is a real number and if matrix $A$ is such that $A^2 = I$, then

$$e^{i\beta A} = \cos(\beta) I + i \sin(\beta) A.$$

**matrix in a Hilbert space** An $n \times m$ matrix $A$ is regarded as a linear operator from an $n$-dimensional Hilbert space, $\mathcal{H}_n$, to an $m$-dimensional Hilbert space, $\mathcal{H}_m$, namely

$$A : \mathcal{H}_n \longrightarrow \mathcal{H}_m.$$

**maximal quantum test** Consider a quantum system and assume that the maximum number of different outcomes of obtainable in a test of the system is $N$; any test that has exactly $N$ different outcomes is a maximal quantum test. For example, when atoms with spin $s$ are used in a Stern-Gerlach experiment $2s+1$ different outcomes are observed, regardless of the orientation of the magnetic field.

**maximum likelihood decoding** Decoding strategy when a received $n$-tuple is decoded into the code word to minimize the probability of errors.

**Maxwell, James Clerk (1831 - 1879)** Eminent Scottish mathematician and theoretical physicist. His most significant achievement was aggregating a set of equations in electricity, magnetism and inductance Maxwell's equations. Maxwell demonstrated that electric and magnetic fields travel through space, in the form of waves, and at the constant speed of light. He also developed the Maxwell distribution, a statistical means to describe aspects of the kinetic theory of gases.

**Maxwell demon** In 1871, Maxwell proposed the following "thought" experiment. Imagine the molecules of a gas in a cylinder divided in two by a wall which has a slit; the slit is covered with a door controlled by a little demon. The demon examines every molecule of gas and determines its velocity; those of high velocity on the left side are let to migrate to the right side and those with low velocity on the right side are allowed to migrate to the left side. As a result of these measurements, the demon separates hot from cold in blatant violation of the second law of thermodynamics.

**measurement of a quantum system** The process which makes a connection between the quantum and the classical worlds; generally considered as an irreversible operation which destroys quantum information about an observable (property) of a quantum system and replaces it with classical information. In quantum mechanics, the measurement of an observable of a quantum system (such as momentum, energy, or spin) is associated with a Hermitian operator, $\mathbf{A}$, on the Hilbert space of state vectors of the system. If $v$ is an eigenvector of $\mathbf{A}$ with eigenvalue $\lambda$, then measuring the system in a state described by state vector $v$ will always give the result $\lambda$. If the state vector is not an eigenvector of $\mathbf{A}$, the measurement process forces the system to jump (collapse) randomly to a state corresponding to a state vector $v_i$, an eigenvector of $\mathbf{A}$. The result of the measurement is the corresponding eigenvalue of $\mathbf{A}$, $\lambda_i$. See also *projective measurement*.

**memoryless channel** A *discrete memoryless classical channel* is the triplet $(X, Y, \{p(y|x), x \in \{X, y \in Y\})$ with:

- $X$ the input channel alphabet,
- $Y$ the output channel alphabet, and
- $\{p(y|x), x \in X, y \in Y\})$ the probability of observing output symbol $y$ when the input symbol is $x$. The channel is memoryless if the probability distribution of the output depends only upon the input at that time and not upon the past history.

A *discrete memoryless quantum channel* means that the state of the quantum system is a vector in a finite-dimensional Hilbert space; the memoryless property has the same meaning as for the classical information, the mean of the output of the quantum communication channel is only determined by the current input.

**minimum distance decoder** Minimum distance (also called nearest neighbor) decoder for a code $\mathcal{C}$ with distance $d$, is a decoder that, given a received word $c_r$, selects the code word $c \in \mathcal{C}$ that satisfies $d(c_r, c) < d/2$ if such a codeword exists, otherwise declares failure.

**mixed state** A quantum system whose state $\mid \psi\rangle$ is not known precisely is said to be in a mixed state. A mixed state is a superposition of different pure states; the system is in a

state $\mid \varphi_i\rangle$ with probability $p_i$. The density operator of a quantum system in a mixed state is $\rho = \sum_i p_i \mid \varphi_i\rangle\langle\varphi_i \mid$. The trace of the density operator is $\mathrm{tr}(\rho^2) < 1$. A mixed state is also called an *impure state*. See also *pure state of a single qubit*.

**monic polynomial** A polynomial $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + 1$ in which the coefficient of the highest order term is 1.

**mutual information** The mutual information measures the reduction in uncertainty of a random variable $X$ due to another random variable $Y$ is:

$$I(X;Y) = \sum_x \sum_y p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}.$$

**Newmark theorem** The theorem [6] states that one can extend the Hilbert space $\mathcal{H}_n$ in which the POVM operators $A_i$ are defined to a larger Hilbert space, $\mathcal{H}_N$, with $N > n$ such that in $\mathcal{H}_N$ there exist a set of orthogonal projectors $P_i$ with two properties:

- $\sum_i P_i = I$ and
- $A_i$ is the result of projecting $P_i \in \mathcal{H}_N$ into $\mathcal{H}_n$.

**non-degenerate code** A quantum code for which linearly independent correctable errors acting on the coding space produce linearly independent states. See also, *degenerate codes*.

**non-degenerate stabilizer code** A non-degenerate stabilizer code $[n, k, d]$ is one when none of the elements of the stabilizer $S$ (except the identity) has weight less than $d$; in case of a degenerate stabilizer code at least one element of the stabilizer $S$ other than the identity has a weight less than $d$.

**non-demolition measurement** If the density matrix of a qubit is

$$\rho = p \mid 0\rangle\langle 0 \mid + (1-p) \mid 1\rangle\langle 1 \mid + \alpha \mid 0\rangle\langle 1 \mid + \alpha^* \mid 1\rangle\langle 0 \mid,$$

with $\alpha \in \mathbb{C}$, then an "ideal measurement" should provide an outcome of "0" with probability $p$ and an outcome of "1" with probability $(1-p)$, regardless of the value of $\alpha$, regardless of the state of the neighboring qubits, and without changing the state of the quantum computer. A "non-demolition" measurement also leaves the qubit in state $\mid 0\rangle\langle 0 \mid$ after reporting the outcome 0 and leaves it in state $\mid 1\rangle\langle 1 \mid$ after reporting the outcome 1. Quantum error correction requires nondemolition measurements of the error syndrome to preserve the state of the physical qubits.

**norm of a vector in a Hilbert space** A non-negative function which measures the "length" of a vector. If $\mid \psi\rangle \in \mathcal{H}_n$, then the norm, $\| \mid \psi\rangle \|$, can be computed from the inner product of the state vector with itself, $\| \mid \psi\rangle \|^2 = \mid \langle\psi \mid \psi\rangle \mid$.

**normal operator in a Hilbert space** An operator $\mathbf{U} \in \mathcal{H}_n$ with the property that $[\mathbf{U}, \mathbf{U}^\dagger] = \mathbf{U}\mathbf{U}^\dagger - \mathbf{U}^\dagger\mathbf{U} = 0$.

**normal unitary basis of an $n$-dimensional Hilbert space, $\mathcal{H}_n$** A set of $n$ vectors, $\mid \psi_0\rangle, \mid \psi_2\rangle, \ldots, \mid \psi_i\rangle, \ldots, \mid \psi_{n-1}\rangle$, where each vector has the norm (or length) equal to one, $\| \mid \psi_0\rangle \| = \| \mid \psi_1\rangle \| = \ldots = \| \mid \psi_{n-1}\rangle \| = 1$, and any two of them are orthogonal, $\langle\psi_i \mid \psi_j\rangle = 0$, for $(i \neq j)$.

**normalizer of a stabilizer S.** The normalizer $N(S)$ of a quantum code $\mathcal{Q}$ with stabilizer $S \in \mathcal{G}_n$ is the set of elements in $\mathcal{G}_n$ that fix the stabilizer $S$ under conjugation. $\mathcal{G}$ is the $n$-qubit Pauli group. $S$ is a normal subgroup of $N(S)$, $S \subset N(S)$.

**normalization condition in quantum mechanics** Requirement that the square of the modulus of the projections of a state vector on the orthonormal basis sum is equal to 1. This translates into the condition that the sum of probabilities of all possible outcomes of a measurement of a quantum system must be equal to 1.

**observable** A physical property of a quantum system that can be measured by an external observer. Mathematically, each observable $X$ has an associated Hermitian operator $\mathcal{M}_X$ with a complete set of eigenvectors.

**operator** A function used to transform the state of a system.

**operator-sum representation** An operator-sum representation, also called Kraus representation, for transformation $\mathcal{E}(\rho)$ for a system initially in state $\rho$ is $\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger$ with $A_k$ a set of trace preserving operators, in other words $\sum_k A_k^\dagger A_k = I$.

**oracle** The name "oracle" comes form the Latin verb "orare" meaning "to speak". An oracle is a source of wise pronouncements or prophetic opinions with some connection to an infallible spiritual authority. In the ancient world, the Greeks believed that they could read the future with the help of oracles. Pythia, the priestess of the temple of Apollo at Delphi, is probably the most famous oracle, but lesser ones are known. According to Herodotus, when Croesus, the king of Lydia, demanded to know if he should go to war against Cyrus the Great (of Persia), Pythia responded in a truly Delphian style "If you do, you will destroy a great empire." Believing the response favorable, Croesus attacked, but it was his own empire that was ultimately destroyed by the Persians in 547 BC.

**oracle models** An oracle implements a computational model and solves a problem for which there is no efficient solution. Oracle models are used to compare the power of two computational models. For example, in 1994, D. Simon showed that a quantum computer with a specific oracle $\mathcal{O}$ efficiently could solve problems that classical computers having access to the same oracle $\mathcal{O}$ could not solve efficiently. See also *efficient computation.*

**order of an element in a finite field $GF(q)$.** Let $GF(q)$ be a finite field with $q$ elements and $a \in F$. The order of $a$, ord$(a)$ is the smallest integer $s$ such that $a^s = 1 \mod q$.

**orthogonal state vectors** Two vectors, $\mid \psi_a\rangle$ and $\mid \psi_b\rangle$ in $\mathcal{H}_n$, are *orthogonal*, and we write $\mid \psi_a\rangle \perp \mid \psi_b\rangle$ if their inner product is zero $\langle\psi_a \mid \psi_b\rangle = 0 \implies \mid \psi_a\rangle \perp \mid \psi_b\rangle$.

**orthonormal basis in a Hilbert space, $\mathcal{H}_n$** A basis consisting of a set of $n$ orthonormal vectors such as $\{\mid 0\rangle, \mid 1\rangle, \ldots, \mid i\rangle, \ldots, \mid n-1\rangle\}$. Any two vectors from the set are orthogonal, and all have a norm equal to 1. See also *normal unitary basis of an $n$-dimensional Hilbert space, $\mathcal{H}_n$.*

**outer product in a Hilbert space, $\mathcal{H}_n$** *The outer product of a* `ket` *vector and a* `bra` *vector, $\mid \psi_a\rangle\langle\psi_b \mid$, is a linear operator. For example, in $\mathcal{H}_4$, we have*

$$\mid \psi_a\rangle\langle\psi_b \mid = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \begin{pmatrix} \beta_0^* & \beta_1^* & \beta_2^* & \beta_3^* \end{pmatrix} = \begin{pmatrix} \alpha_0\beta_0^* & \alpha_0\beta_1^* & \alpha_0\beta_2^* & \alpha_0\beta_3^* \\ \alpha_1\beta_0^* & \alpha_1\beta_1^* & \alpha_1\beta_2^* & \alpha_1\beta_3^* \\ \alpha_2\beta_0^* & \alpha_2\beta_1^* & \alpha_2\beta_2^* & \alpha_2\beta_3^* \\ \alpha_3\beta_0^* & \alpha_3\beta_1^* & \alpha_3\beta_2^* & \alpha_3\beta_3^* \end{pmatrix}.$$

**parametric down-conversion** *A nonlinear crystal splits incoming photons into pairs of photons of lower energy whose combined energy and momentum are equal to the energy and momentum of the original photon; the energy and momentum must be conserved. The photon pair is entangled in the frequency domain due to phase matching.*

**partial trace** *Let $\mathcal{C} = \mathcal{AB}$ be a composite system consisting of two subsystems $\mathcal{A}$ and $\mathcal{B}$ described by the density operator $\rho_{\mathcal{C}}$. The partial trace of $\rho_{\mathcal{C}}$ over system $\mathcal{B}$ is:*

$$\mathrm{tr}_{\mathcal{B}}\left[\rho_{\mathcal{C}}\right] = \mathrm{tr}_{\mathcal{B}}\left[\ \mid a_1\rangle\langle a_2 \mid \otimes \mid b_1\rangle\langle b_2 \mid\ \right] = \mid a_1\rangle\langle a_2 \mid\ \mathrm{tr}_{\mathcal{B}}\left[\ \mid b_1\rangle\langle b_2 \mid\ \right] = \mid a_1\rangle\langle a_2 \mid \langle b_1 \mid b_2\rangle.$$

with $\mid a_1\rangle, \mid a_2\rangle$ any two vectors in the state space of $\mathcal{A}$ and $\mid b_1\rangle, \mid b_2\rangle$ any two vectors in the state space of $\mathcal{B}$.

**Pauli matrices** Matrices $\sigma_I, \sigma_x, \sigma_y, \sigma_z$ describing transformations (rotations) of a single qubit

$$\sigma_I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = i\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Pauli matrices are Hermitian $\sigma_I^\dagger = \sigma_I,\ \sigma_x^\dagger = \sigma_x,\ \sigma_y^\dagger = \sigma_y,\ \sigma_z^\dagger = \sigma_z$ and unitary: $\sigma_x\sigma_x^\dagger = \sigma_x^2 = \sigma_I,\ \sigma_y\sigma_y^\dagger = \sigma_y^2 = \sigma_I,\ \sigma_z\sigma_z^\dagger = \sigma_z^2 = \sigma_I$; their multiplication table is

| $\times$ | $\sigma_I$ | $\sigma_x$ | $\sigma_y$ | $\sigma_z$ |
|---|---|---|---|---|
| $\sigma_I$ | $\sigma_I$ | $\sigma_x$ | $\sigma_y$ | $\sigma_z$ |
| $\sigma_x$ | $\sigma_x$ | $\sigma_I$ | $i\sigma_z$ | $-i\sigma_y$ |
| $\sigma_y$ | $\sigma_y$ | $-i\sigma_z$ | $\sigma_I$ | $i\sigma_x$ |
| $\sigma_z$ | $\sigma_z$ | $i\sigma_y$ | $-i\sigma_x$ | $\sigma_I$ |

**Pauli, Wolfgang Ernst (1900 – 1958).** Distinguished Austrian-born physicist who proposed in 1924 a quantum spin number for electrons. Best known for his (Pauli) exclusion principle, proposed in 1925, which states that no two electrons in an atom can have the same four quantum numbers. He predicted mathematically, in 1931, that conservation laws in beta decay required the existence of a new particle, electrically neutral and with very low mass and named it "neutron". In 1933, he published his prediction and he made the claim that the particle had zero mass. The particle, which we now know as the neutron, has a non-zero mass and was discovered by Chadwick in 1932. Pauli's particle was named the "neutrino" by Fermi in 1934, and at that time, he correctly stated that it was not a constituent of the nucleus of an atom. The neutrino was found experimentally in 1956 by C. Cowan and F. Reines.

**Pauli exclusion principle** Two electrons on the same orbital around the nucleus of an atom cannot be in identical states, including the spin; they must have their spins oriented in opposite directions, because they already share three quantum numbers. If, as a result of an experiment, one of the electrons is made to change the orientation of its spin, then a simultaneous measurement of the other finds it in a state with the *opposite* spin.

**Pauli group** The 1-qubit Pauli group $\mathcal{G}_1$ consists of the Pauli matrices, $\sigma_I$, $\sigma_x$, $\sigma_y$, and $\sigma_z$ together with the multiplicative factors $\pm 1$ and $\pm i$:

$$\mathcal{G}_1 \equiv \{\pm\sigma_I, \pm i\sigma_I, \pm\sigma_x, \pm i\sigma_x, \pm\sigma_y, \pm i\sigma_y, \pm\sigma_z, \pm i\sigma_z\}.$$

The cardinality of $\mathcal{G}$ is $\mid \mathcal{G}_1 \mid = 16 = 2^4$. The n-qubit Pauli group $\mathcal{G}_n$ consists of the $4^n$ tensor products of $\sigma_I$, $\sigma_x$, $\sigma_y$, and $\sigma_z$ and an overall phase of $\pm 1$ or $\pm i$. The group has $4^{n+1}$ elements.

**perfect code capable to correct $e$ errors** Linear code where Hamming spheres of radius $e$ about each code word are disjoint and exhaust the entire space of $n$-tuples.

**perfect quantum code** A quantum code achieving the quantum Hamming bound; see the entry of *quantum Hamming bound*.

**photon** From the Greek word "photos" meaning light; a light particle.

**Plotkin bound** If $\mathcal{C}$ is an $[n, k, d]$ linear code, if $2d > n$ and $M = \mid \mathcal{C} \mid$ then the code satisfies the Plotkin bound:

$$M \leq 2\lfloor \frac{d}{2d-n} \rfloor.$$

**polarization filter** A partially transparent material that transmits light of a particular polarization.

**polarization measurements** Polarization is measured by passing a photon through a polarizer. If the polarizer axis is oriented parallel to the polarization of the photon, the photon passes through unimpeded. If it is oriented perpendicular to the polarization of the photon, the photon is absorbed. At an intermediate angle, the photon will have a certain probability of being transmitted.

**polarization of light** As an *electromagnetic radiation*, light consists of an electric and a magnetic field perpendicular to each other and, at the same time, perpendicular to the direction the energy is transported by the electromagnetic (light) wave. The electric field oscillates in a plane perpendicular to the direction of light and the way the electric field vector travels in this plane defines the polarization of the light. When the electric field oscillates along a straight line, we say that the light is *linearly polarized*. When the end of the electric field vector moves along an ellipse, the light is *elliptically polarized*. When the end of the electric field vector moves around a circle, the light is *circularly polarized*. If the light comes toward us and the end of the electric field vector moves around in a counterclockwise direction, we say that the light has *right-hand polarization*; if the end of the electric field vector moves in a clockwise direction, we say that the light has *left-hand polarization*.

**polynomial time algorithm** Term in computational complexity theory to describe algorithms which require a number of steps, $T(n)$, bounded by a polynomial, when $n$ is the size of the input. $\mathcal{P}$ is the class of polynomial time algorithms. $\mathcal{NP}$ is the class of polynomials which require a number of steps not bounded by a polynomial.

**positive operator** A positive operator $\mathbf{A}$ is any Hermitian operator with non-negative eigenvalues; $\sqrt{A}$ denotes the unique positive root of $\mathbf{A}$.

**POVM** The positive operator-valued measure is a set of generalized measurement positive operators concerned only with the statistics of the measurement; POVM operators are not necessarily orthogonal or commutative and allow the possibility of measurement outcomes associated with non-orthogonal states. The number of POVM operators may differ from the dimension of the Hilbert space, while the number of projective operators is precisely equal to the dimension of the Hilbert space.

**primitive element of a finite field.** Given a finite field $F(q)$ with $q$ elements there is an element $\alpha \in F$ such that the set on nonzero elements of $F$ is

$$F - \{0\} = \{\alpha^1, \alpha^2, \ldots, \alpha^{q-2}\} \quad \text{and} \quad \alpha^{q-1} = 1.$$

**projection operator, projector** The outer product of a unit vector with itself as

$$\mathbf{P}_a = \mid \Psi_a \rangle \langle \Psi_a \mid .$$

It has the defining property: $\mathbf{P}_a^2 = \mathbf{P}_a$. A complete set of orthogonal projectors in $\mathcal{H}_n$ is a set $\{\mathbf{P}_0, \mathbf{P}_1, \ldots, \mathbf{P}_i, \ldots, \mathbf{P}_{m-1}\}$ such that $\sum_{i=0}^{m-1} \mathbf{P}_i = 1$.

**projective measurement** In a von Neumann-type projective measurement the measurement operators are Hermitian and idempotent. The number of such operators is equal to the dimension of the Hilbert space. Since orthogonal measurement operators commute, they correspond to simultaneous observables. A projective measurement is characterized by a set of projectors, $\mathcal{M}_i$, such that $\sum_i \mathcal{M}_i = I$ and $\mathcal{M}_i \mathcal{M}_j = \delta_{ij} \mathcal{M}_i$. The outcome of the measurement is the one with the index $i$ associated with $\mathcal{M}_i$. The probability of outcome $i$ for a system in state $\mid \psi \rangle$ is $p_i = \mid \mathcal{M}_i \mid \psi_i \rangle \mid^2$. Given the outcome $i$, the quantum state "collapses" to the state $M_i \mid \psi_i \rangle / \sqrt{p_i}$.

**pure state of a single qubit** Pure states are characterized by maximum knowledge, they can be expressed as superpositions of basis vectors of an orthonormal basis. Pure states of a single qubit are represented by points on the Bloch sphere. This implies that the trace of the square of their density matrix is one, $\mathrm{tr}(\rho^2) = 1$. See also *mixed state* or *impure states of a single qubit* .

**purification** Purification is the process of finding a reference system $\mathcal{B}$ given a quantum system $\mathcal{A}$ in a mixed state $\rho_{\mathcal{A}}$. If $\mid \psi_{\mathcal{C}} \rangle$ is the pure state of the composite system $\mathcal{C} = \mathcal{AB}$ we wish that $\rho_{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}} (\mid \psi_{\mathcal{C}} \rangle \langle \psi_{\mathcal{C}} \mid)$. Also, a purification is any pure state in the extended Hilbert space $\mathcal{H}_{\mathcal{C}} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ having $\rho_{\mathcal{A}}$ as the reduced state for the subsystem.

**quadratic form** Let $A(a; b)$ be a symmetric bilinear function. Then $A(a; a)$ is called a quadratic form.

**quantitative version of the Church-Turing principle** "Any physical computing device can be simulated by a Turing Machine in a number of steps polynomial in the resources used by the computing device." This thesis allows us to relate the behavior of the abstract model of computation provided by the Turing Machine concept with the physical computing devices used to carry out a computation.

**quantum** Latin word meaning "some defined quantity." In Physics, it is used with the same meaning as *discrete* in mathematics.

**quantum circuit complexity** The smallest number of quantum gates necessary to implement an operation on a fixed number of qubits.

**quantum communication channel** Physical media allowing two parties to exchange quantum information. For example, an optic fiber allowing photons with a certain polarization to circulate from a source to a destination.

**quantum erasure** A phenomenon caused by entanglement; the state of a system is characterized jointly by the density matrix and the information regarding the components of a composite system.

**quantum errors** The most general type of error of a qubit can be described a combination of two basic types of quantum errors: amplitude, or *bit-flip* and phase, or *phase-flip*. Classical errors correspond to quantum bit-flip errors, changes of the form $\mid 0\rangle \leftrightarrow \mid 1\rangle$; the Pauli $\sigma_x$ transformation applied to a qubit leads to a bit-flip. Phase-flips are changes of the form $\mid 0\rangle + \mid 1\rangle \leftrightarrow \mid 0\rangle - \mid 1\rangle$; the Pauli $\sigma_z$ transformation applied to a qubit leads to a phase-flip. There is no correspondence of quantum phase-flip in classical error correction.

**quantum data processing inequality** Consider a quantum system in a state characterized by the density matrix $\rho$ transmitted through a quantum channel Figure 86; the resulting state $\mathcal{E}_1(\rho)$ becomes the input to a second channel and the resulting state is $\mathcal{E}_2 \circ \mathcal{E}_1(\rho)$. The two transformation $\mathcal{E}_1$ and $\mathcal{E}_2$ are trace-preserving. Then the coherent information between the original input and the output of the second channel, $I(\rho, \mathcal{E}_2 \circ \mathcal{E}_1(\rho))$, cannot be larger than the coherent information between the original input and the output of the first channel, $I(\rho, \mathcal{E}_1(\rho))$ and, in turn, $I(\rho, \mathcal{E}_1(\rho))$ cannot be larger than $S(\rho)$, the von Neumann entropy of the source; $S(\rho) \geq I(\rho, \mathcal{E}_1) \geq I\,(\rho, \mathcal{E}_2 \circ \mathcal{E}_1)$.

**quantum Fano inequality** Relates $S_e\,(\rho, \mathcal{E})$, the entropy exchange, and $F_s(\rho, \mathcal{E})$, the entanglement fidelity, $S_e\,(\rho, \mathcal{E}) \leq h(p) F_s(\rho, \mathcal{E}) + [1 - F_s(\rho, \mathcal{E})] \log(d^2 - 1)$. Here $d$ is the dimension of the Hilbert space of the quantum system $Q$ and $h(p)$ is the binary entropy associated to the probability $p$, as $h(p) = -p \log p - (1-p) \log(1-p)$.

**quantum Hamming bound** Gives the smallest number $n$ of physical qubits required to encode a logical qubit $2(3n+1) \leq 2^n$. Thus, $n = 5$ is the smallest number of qubits required to encode the two superposition states $\mid 0_L\rangle$ and $\mid 1_L\rangle$, and then be able to recover them regardless of the qubit in error and the type of error.

**quantum parallelism** Term capturing the fact that a quantum computer can manipulate an exponential set of inputs simultaneously.

**quantum particle** Particle obeying the laws of Quantum Mechanics.

**quantum accuracy threshold theorem** An arbitrarily long computation can be carried out reliably provided that all sources of "noise" affecting a system are weaker than the accuracy threshold; a lower bound on this threshold, $\epsilon > 1.9 \times 10^{-4}$ was derived in [9].

**qubit** Quantum bit; mathematical abstraction for a quantum system capable of storing one bit of information.

**qudit** Multi-level quantum states used for quantum information processing; for example for $d = 3$ we have a quantum system with three basis states.

**Rabi oscillation** The oscillation of the level population of a two-state quantum system in the presence of an oscillatory driving field. The optical Rabi oscillation is the result of a coherent non-linear light-matter interaction; the oscillation results from the quantum interference between the probability amplitudes of two atomic dipole eigenstates involved in the interaction. When an atom or some other two-level system is illuminated by a coherent beam of photons, it will cyclically absorb photons and re-emit them by stimulated emission; such a cycle is a *Rabi cycle* and the inverse of its duration is the *Rabi frequency* of the photon beam.

**Raman transition** The result of inelastic scattering of monochromatic light, usually laser light in the visible, near infrared or near ultraviolet range; in a Raman process, the parent atom or molecule is excited from the ground state to a virtual energy state and relaxes into a vibrational excited state by emitting a photon. The emitted photon is shifted up or

down in energy in relation to the incident one; this shift in energy (Raman shift) is equal to the vibrational level involved in the transition and gives information about the phonon modes in the system. In a stimulated Raman transition a two color (frequency) laser pulse transfers the atom/molecule from ground to a vibration excitation state if the difference in energy matches an allowed Raman transition.

**rank of a linear transformation** The dimension of the range of the transformation. where the range (image) is the set of vectors the transformation maps to.

**rank of a matrix** The maximum number of linearly independent rows or columns of matrix $A$.

**Rationalism, continental rationalism** Theory of knowledge based upon the idea that the criterion of truth is not sensory but intellectual and deductive. The origins of this philosophy whose main credo is self-sufficiency of reason can be traced back to Pythagoras and Plato and was later embraced by Descartes, Spinoza, and Leibnitz who advocated the introduction of mathematical methods in philosophy. Spinoza, and Leibnitz believed that in principle, all knowledge, including scientific knowledge, could be gained through the use of reason alone. Rationalism is often contrasted to Empiricism; see also the entry for *Empiricism*.

**ray in a Hilbert space** A mathematical abstraction that exhibits only direction. It can be represented as a straight line through the origin of the coordinate system.

**reciprocal polynomial** Given a polynomial of degree $m$, $h(x) = \sum_{k=0}^{m} c_k x^k,\ c_m \neq 0$, its reciprocal is the polynomial

$$\bar{h}(x) = \sum_{k=0}^{m} c_{m-k} x^k.$$

**Reed-Solomon code** A Reed-Solomon code over $GF(q)$ of length $n = 2^k - 1$, dimension $k$, and distance $d$ is a cyclic code with the generator polynomial

$$g(x) = (x - \beta^j)(x - \beta^{j+1}) \ldots (x - \beta^{j+d-2})$$

with $\beta$ be a primitive element of order $n = 2^k - 1$ of the finite field $GF(2^k)$.

**reduced density operator** Let $\mathcal{C} = \mathcal{AB}$ be a composite system described by the density operator $\rho_{\mathcal{C}}$. The reduced density operator of subsystem $\mathcal{A}$ is $\rho^{\mathcal{A}} = \mathrm{tr}_{\mathcal{B}}\,(\rho_{\mathcal{C}})$.

**relative entropy** The relative entropy between two distributions $p(x)$ and $q(x)$ quantifies how close the two distributions are from each other

$$H(p(x) \mid\mid q(x)) = \sum p(x) \log \frac{p(x)}{q(x)}.$$

If the real probability distribution of events is $p(x)$ but we erroneously assume that it is $q(x)$ then the surprise when an event occurs is $\log(1/q(x))$ and the average surprise is $\sum p(x) \log \frac{1}{q(x)}$.

**reversible physical process** A physical process is said to be reversible if it can evolve forwards as well as backwards in time; a reversible system can always be forced back to its original state from a new state reached during an evolutionary process.

**Schmidt decomposition** Let $\mid \psi_{\mathcal{C}}\rangle$ be the state of a composite system $\mathcal{C} = \mathcal{AB}$ with $\mathcal{A} \in \mathcal{H}_A$ and $\mathcal{B} \in \mathcal{H}_B$ There are two orthonormal bases, called Schmidt bases $\{\mid e_j^A\rangle\} \in \mathcal{H}_A$

and $\{| e_j^B\rangle\} \in \mathcal{H}_B$ such that $| \psi_{\mathcal{C}}\rangle = \sum_j \lambda_j | e_j^A\rangle \otimes | e_j^B\rangle$; the Schmidt coefficients $\lambda_j$ are real numbers and $\sum_j \lambda_j^2 = 1$.

**Schmidt number** Consider a bipartite pure state $| \psi_{\mathcal{C}}\rangle$ of a composite system $\mathcal{C} = \mathcal{AB}$. The number of non-zero eigenvalues $\lambda_j$ of the density matrix $\rho_{\mathcal{C}}$, thus the number of terms in the Schmidt decomposition of $| \psi_{\mathcal{C}}\rangle$ is called the Schmidt number of state $| \psi_{\mathcal{C}}\rangle$. The Schmidt number provides a clear distinction between bipartite pure states $| \psi_{\mathcal{AB}}\rangle$ that are separable and those that are entangled: *a bipartite pure state* $| \psi_{\mathcal{AB}}\rangle$ *is entangled if its Schmidt number is greater than one, otherwise it is separable.*

**Schrödinger, Erwin (1887 - 1961).** Eminent Austrian-born physicist. He is one of the founders of Quantum Physics and has made significant contributions to statistical mechanics and the general theory of relativity. Schrödinger transferred the idea of a wave associated to a particle, predicted by de Broglie, to Bohr's atomic model. In 1927, he showed the mathematic equivalence of his equation $\mathbf{H\Psi} = \mathbf{E\Psi}$ and Heisenberg's matrix mechanics. That same year he became Max Planck's successor for the theoretical physics chair at the University of Berlin. He was awarded the Nobel Prize for Physics in 1933 (together with Paul Dirac) for his contributions to the development of quantum mechanics.

**Schwartz inequality** Inequality satisfied by any state vectors $(| \psi_a\rangle, | \psi_b\rangle) \in \mathcal{H}_n$ where

$$\langle\psi_a \mid \psi_a\rangle\langle\psi_b \mid \psi_b\rangle \;\geq\; | \langle\psi_a \mid \psi_b\rangle |^2 \,.$$

**Schumacher compression** We can compress the quantum information produced by the quantum source $\mathcal{A}$ with density matrix $\rho^{\mathcal{A}}$ up to $\log\left[2^{nS(\rho^{\mathcal{A}})+\mathcal{O}(n)}\right] \approx nS(\rho^{\mathcal{A}})$ qubits.

**second law of thermodynamics** The entropy of a system is a non-decreasing function of time.

**self-adjoint operator or matrix** A linear operator or a matrix $A$ with the property that $A = A^\dagger$, where $A^\dagger$ is the dual, or the adjoint of $A$. See also the entry for *adjoint matrix*.

**self-dual code** A linear code $\mathcal{C}$ is called weakly self-dual if $\mathcal{C}^\perp \subset \mathcal{C}$. When $\mathcal{C} = \mathcal{C}^\perp$ the code is called strictly self-dual.

**self-dual basis** See the entry for *trace-orthogonal basis*.

**Shannon, Claude Elwood (1916 - 2001).** Eminent American mathematician and electrical engineer; founder of the modern information theory. In 1948, Shannon published his landmark work *A Mathematical Theory of Communication*, and in 1949, he published the *Communication Theory of Secrecy Systems*, generally credited with transforming cryptography from an art to a science.

**Shannon entropy** Given a random variable $X$, with a probability density function $p_X(x)$, the entropy is a positive real number $H(X) = -\sum_x \; p_X(x) \times \log_2 p_X(x)$.

**Shor, Peter** Distinguished American mathematician, well known for his quantum factoring algorithm and for his pioneering work on quantum error-correcting codes.

**singlet electron state** The anti-symmetric state of a pair of electrons with *anti-parallel* spins $1/\sqrt{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$. The electrons have different spin quantum numbers $+1/2$ and $-1/2$, and the total spin of the state is zero. See also the entry for *triplet electron state*.

**single value decomposition theorem** An $n \times n$ matrix $\Gamma$ over a field $F$ of real or complex numbers, can be decomposed into a product of three $n \times n$ matrices over $F$: $\Gamma = A\Delta B^\dagger$ with: $A = [\alpha_{ij}]$ and $B = [\beta_{ij}]$ unitary matrices, and $\Delta = [\delta_{ij}]$ a diagonal matrix.

**Singleton bound** Every $[n, k, d]$ linear code $\mathcal{C}$ over the field $GF(q)$ satisfies the Singleton bound: $k + d \leq n + 1$.

**Solovay-Kitaev theorem** $\mathcal{G} \subset SU(d)$ is a universal family of gates if (i) $\mathcal{G}$ is closed under the inverse operation, $\forall g \in \mathcal{G} \leftrightarrow g^{-1} \in \mathcal{G}$, and (ii) $\mathcal{G}$ generates a dense subset of $SU(d)$; then

$$\forall U \in SU(d), \epsilon > 0, \exists g_1, g_2, \ldots, g_q \in \mathcal{G} : \| U - U_{g_1, g_2, \ldots g_q} \| \leq \epsilon \quad \text{and} \quad q = \mathcal{O}(\log^2 1/\epsilon)$$

where $\| U \|$ is the norm of the linear operator $U$ and $U_{g_1,g_2,\ldots,g_q}$ is an implementation of $U$ using only gates from the set $\mathcal{G}$. For example, if a set $\mathcal{G}$ of single -qubit quantum gates generates a dense subset of $SU(2)$ than it is possible to construct good approximations of any gate using short sequences of the set $\mathcal{G}$ of one-qubit gates

**special unitary group of degree $n$, $SU(n)$** The multiplicative group of $n \times n$ unitary matrices with determinant equal to 1. $SU(n)$ is a subgroup of the *unitary group* $U(n)$ of all $n \times n$ unitary matrices. In turn, $U(n)$ is a subgroup of the *general linear group* $GL(n, \mathbb{C})$.

**spectral decomposition of a normal operator in a Hilbert space** In $\mathcal{H}_n$, every normal operator $\mathbf{A}$ has $n$ eigenvectors $\mid a_i \rangle$, and, correspondingly, $n$ eigenvalues $\lambda_i, 1 \leq i \leq n$. If $\mathbf{P}_i$ are the projectors corresponding to these eigenvectors, $\mathbf{P}_i = \mid a_i \rangle \langle a_i \mid$, then the operator $\mathbf{A}$ has the spectral decomposition $\mathbf{A} = \sum_{\mathbf{i}} \lambda_{\mathbf{i}} \mathbf{P}_{\mathbf{i}}$.

**spectrum of a vector in an extension finite field** The spectrum $\bar{c}$ of a vector $c = \sum_{i=1}^{n} \alpha_i c_i$ with elements $\alpha_i \in GF(2^k)$ is $\bar{c} = DFT(c)$ with $DFT$ the cyclic Discrete Fourier Transform.

**spin** The observable associated with the intrinsic rotation of a quantum particle is the intrinsic angular momentum, also called the *spin angular momentum*. The "spin" is the quantum number characterizing the intrinsic angular momentum of the quantum particles. There are two types of particles: *bosons*, particles whose spin quantum number can be $s = 1$, $s = -1$, or $s = 0$, e.g., photons and mesons, and *fermions*, particles whose spin quantum number can be $s = 1/2$ or $s = -1/2$ e.g., electrons, protons, and neutrons.

**Stirling approximation** $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$.

**superposition probability rule** In quantum mechanics, if an event may occur in two or more indistinguishable ways, then the probability amplitude of the event is the sum of the probability amplitudes of each case considered separately.

**superposition state of a quantum system** If the states $\mid 0\rangle, \mid 1\rangle, \ldots, \mid n-1\rangle$ of a quantum system are distinguishable and if the complex numbers $\alpha_i$ satisfy the condition $\sum_i \mid \alpha_i \mid^2 = 1$, then the state $\sum_i \alpha_i \mid i\rangle$ is a valid quantum state called a superposition state. See also *basis states of a quantum system.*

**stabilizer** A complete description of a particular class of quantum error correcting codes by the set of tensor products of Pauli matrices that fix every state in the coding space.

**stabilizer code** A class of quantum error correcting codes described by an Abelian subgroup of the group $\mathcal{G}$ of Pauli matrices.

**Stark effect** Splitting and shifting of an atomic spectral line into several components in the presence of an electric field, discovered by Johannes Stark in 1913; analogous to the Zeeman effect, the splitting of an atomic spectral line into several components in the presence

of a magnetic field. The effect is due to the interaction of the electric dipole moment with the external electric field. In a semiconductor heterostructure where a small bandgap material is sandwiched between layers of a material with larger bandgap, like that used to manufacture the quantum dots, the effect can be significantly enhanced by bound excitons. The electron and the hole forming the excitons are pushed in opposite directions by the electric field, but are confined in the smaller bandgap material and can not be pulled apart.

**stationary, ergodic source of classical information** Source that emits symbols with a probability that does not change over time; an *ergodic* source emits information symbols with a probability equal to the frequency of their occurrence in a long sequence. Stationary, ergodic sources have a finite but arbitrary and potentially long correlation time.

**Stern-Gerlach experiment** Experiment revealing the spin of quantum systems.

**Stinespring dilation theorem** If $A$ is a unital $C^*$ algebra and $B(\mathcal{H})$ is a bounded operator on the Hilbert space $\mathcal{H}$, then for every completely positive map $T : A \mapsto B(\mathcal{H})$ there exists another Hilbert space $\mathcal{K}$ and a unital *-homomorphism $U : A \mapsto B(\mathcal{K})$ such that $T(a) = VU(a)V^*, \quad \forall a \in A$ if $V : \mathcal{K} \mapsto \mathcal{H}$ is a bounded operator. Then $\| T(1) \| = \| V \|^2$ [412]. The interpretation of Stinespring's Dilation Theorem is that any completely positive and trace-preserving map in a Hilbert space can be constructed from three operations: (i) tensoring with a second Hilbert system in a specified state; (ii) a unitary transformation on the larger space obtained as a result of step (i), and (iii) reduction to a subsystem.

**strong superposition principle** Any complete set of mutually orthogonal vectors has a physical realization as a maximal quantum test [325].

**super selection rule (SSR)** An SSR is a restriction on the allowed local operations on a system, not on its allowed states, and it is associated with a group of physical transformations [24]. Such restrictions could be imposed by the properties of the underlying theory, or arise due to physical restrictions. The operations it applies to include unitary transformations, $\mathbf{O}\rho = U\rho U^\dagger$, and measurements, $\mathbf{O}_r\rho = \mathcal{M}_r\rho\mathcal{M}_r^\dagger$ with $\sum \mathcal{M}_r\mathcal{M}_r^\dagger = 1$. SSRs prohibit entangled states involving different particle numbers, thus a two-qubit system consisting of two quantum dots and an electron in a superposition state as being on one or the other quantum dot would not satisfy the first requirement [131].

**support of a normal operator** $\mathrm{supp}\,[\rho]$, the *support of a normal operator* $\rho \in \mathcal{L}(\mathcal{H})$ is defined as the subspace of $\mathcal{H}$ spanned by the eigenvectors of $\rho$ having nonzero eigenvalues $\mathrm{supp}\,[\rho] = \mathrm{spann}\{| \xi_i\rangle : 1 \leq i \leq n, \lambda_i \neq 0\}$ where $\rho = \sum_{i=1}^n \lambda_i \mid \xi_i\rangle\langle\xi_i \mid$ is a spectral decomposition of $\rho$.

**standard Wiener process** A standard Wiener process (often called Brownian motion) on the interval $[0, T]$ is a continuous random variable $W(t)$ on $t \in [0, T]$ such that $W(0) = 0, \quad W(t) - W(s) \sim \sqrt{t-s}N(0,1), \quad 0 \leq s < t \leq T$ and $W(t) - W(s)$ and $W(v) - W(u)$ are independent for $0 \leq s < t < u < v \leq T$. $N(0,1)$ is a normal distribution with zero mean and unit variance. The process is also referred to as Gaussian.

**symmetric bilinear function (form)** Let $A(a; b)$ be a bilinear function. We say that $A(a; b)$ is a symmetric bilinear function if $A(a; b) = A(b; a)$.

**symmetry** A symmetry is a transformation of the state of a quantum system that leaves all the observables of the system unchanged. To be a symmetry, the mapping should preserve the absolute value of inner products:
$| \varphi^{(i)}\rangle \mapsto | \psi^{(i)}\rangle$ and $| \varphi^{(j)}\rangle \mapsto | \psi^{(j)}\rangle \implies | \langle\varphi^{(i)} \mid \varphi^{(j)}\rangle | = | \langle\psi^{(i)} \mid \psi^{(j)}\rangle | \quad \forall(\varphi^{(i)}, \psi^{(i)}) \in \mathcal{H}_n$.

**Szilárd, Leó (1898 – 1964).** Distinguished Hungarian-born physicist who proposed the idea of nuclear chain reactions before nuclear fission was discovered. In 1939 Szilárd and Fermi conducted a simple experiment at Columbia University and discovered significant neutron multiplication in uranium, proving that the chain reaction was possible and opening the way to nuclear weapons. Szilárd was directly responsible for the creation of the Manhattan Project. In 1929, Leó Szilárd stipulated that information is physical while trying to explain Maxwell's demon paradox [415].

**Szilárd's engine** Imaginary engine powered by Maxwell's demon.

**teleportation** In a science fiction context, making an object or person disintegrate in one place and have it reembodied as the same object or person somewhere else. In the context of quantum information theory: "a way to scan out part of the information from an object A, which one wishes to teleport, while causing the remaining, unscanned, part of the information to pass, via the Einstein-Podolsky-Rosen effect, into another object C, which has never been in contact with A. Later, by applying to C a treatment depending on the scanned-out information, it is possible to maneuver C into exactly the same state as A was in before it was scanned" (see `http://www.research.ibm.com/quantuminfo/teleportation)`. In this process, the original state is destroyed.

**thermodynamic entropy** The thermodynamic entropy of a gas, $S$, quantifies the notion that a gas is a statistical ensemble and it measures the randomness (or the degree of disorder) of the ensemble. The entropy is larger when the vectors describing the individual movements of the molecules of gas are in a higher state of disorder than when all of them are well organized and moving in the same direction with the same speed. Ludwig Boltzmann postulated that $S = k_B \ln \Omega$, with $k_B$ the Boltzmann's constant, $k_B = 1.3807 \times 10^{-23}$ Joules per degree Kelvin, and $\Omega$ the number of microstates.

**thermodynamics** Thermodynamics is the study of energy, its ability to carry out work, and the conversion between various forms of energy, such as the internal energies of a system, heat, and work. Thermodynamic laws are derived from statistical mechanics. C. P. Snow, the English physicist and novelist, summarized the three laws of thermodynamics as follows:

I. You cannot win; matter and energy are conserved, thus you cannot get something for nothing.

II. You cannot break even; you cannot return to the same energy state because there is always an increase in disorder. The entropy always increases.

III. You cannot get out of the game; absolute zero is unattainable.

**Toffoli gate** A three-qubit gate with two control inputs, $a$ and $b$, and one target input, $c$. The outputs are: $a' = a$, $b' = b$ and $c'$. When $c = 1$, then $c' = 1 \oplus (a \texttt{ AND } b) = \texttt{NOT}(a \texttt{ AND } b)$, otherwise $c' = c$. The `Toffoli` gate is a universal gate and it is reversible. There are both classical and quantum versions of the `Toffoli` gate.

**trace distance between two operators** The trace distance between **A** and **B** is

$$D(A, B) = \frac{1}{2}\text{tr} \mid A - B \mid = \frac{1}{2}\text{tr}\sqrt{(A-B)^\dagger(A-B)}$$

with $A$ and $B$ be the matrices corresponding to operators **A** and **B**, respectively.

**trace distance between two probability density functions** The trace distance between $p_X(x)$ and $p_Y(x)$ is denoted by $D(p_X(x), p_X(x))$; also called Kolmogorov, or L1

distance and is defined as $D(p_X(x), p_Y(x)) = \frac{1}{2}\sum_x \mid p_X(x) - p_Y(x) \mid$.

**trace of a linear operator** The trace of an operator is the trace of the matrix associated with the operator.

**trace of a matrix** The sum of the diagonal elements of matrix $A$, $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$; $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

**trace of an extension Galois field $GF(p^k)$ over $GF(p)$** A linear mapping from $GF(p^k)$ to $GF(p)$, $\mathrm{tr} : GF(p^k) \mapsto GF(p)$ defined as $\mathrm{tr}(\alpha) = \alpha + \alpha^k + \ldots + \alpha^{p^{k-1}}$.

**trace-orthogonal basis** The elements $\alpha_1, \alpha_2, \ldots, \alpha_k$ form a basis of $GF(p^k)$ over $GF(p)$ if and only if

$$\det \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \\ \alpha_1^p & \alpha_2^p & \cdots & \alpha_k^p \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{p^{k-1}} & \alpha_2^{p^{k-1}} & \cdots & \alpha_k^{p^{k-1}} \end{pmatrix} \neq 0.$$

The set $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a *trace orthogonal basis* if $\mathrm{tr}(\alpha_i \alpha_j) = 0, \quad 1 \leq i, j \leq k, i \neq j$. If in addition $\mathrm{tr}(\alpha_i^2) = 1$ then $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a *self-dual basis*.

**trace-preserving operation** Informally, a trace-preserving quantum information takes place when a quantum system interacts with the environment and when no measurement is performed either on the quantum system, or the environment. When classical information about the quantum system is made available through a measurement the operation is non trace-preserving. If the operation is denoted by $\mathcal{E}$ and the quantum system is in a state characterized by the density matrix $\rho$, then a trace-preserving operations requires that $\mathrm{tr}\,(\mathcal{E}(\rho)) = 1$. A non-trace preserving quantum operation is characterized by the inequality $\sum_k A_k^\dagger A_k \leq I$.

**triplet electron state** The state of a pair of electrons with *parallel* spins, $\mid \uparrow\uparrow\rangle$ or $\mid \downarrow\downarrow\rangle$, or in a symmetric superposition of anti-parallel spins, $(1/\sqrt{2})(\mid \uparrow\downarrow\rangle + \mid \downarrow\uparrow\rangle$. The total spin of a triplet state is $+1$.

**turbo codes** Turbo codes are a refinement of concatenated codes; they use the encoding structure of concatenated codes and an iterative algorithm for decoding.

**Turing, Alan Mathison (1912 - 1954).** Eminent British mathematician regarded as the founder of the modern Computer Science. In 1936, he published the paper "On Computable Numbers, with an Application to the Entscheidungsproblem." Turing defined a computable number as a real number whose decimal expansion could be produced by a Turing machine starting with a blank tape. In March 1946, he submitted a report proposing the Automatic Computing Engine (ACE), an original design for a modern computer. In 1948, he moved to Manchester, and in 1950, he published a paper, "Computing Machinery and Intelligence in Mind" where he proposed the Turing Test; the test used today to answer the question of whether a computer can be intelligent.

**uniform family of (quantum) circuits** A set of circuits, $\{C_n\}$, with one circuit for each input of length $n$. The *uniformity condition* requires that the description of a circuit $C_n$ should be computed in time *poly(n)* on some Turing machine.

**unitary matrix** A matrix $A = [a_{ij}]$ with complex elements; $a_{ij} \in \mathbb{C}$ is said to be unitary if $A^\dagger A = I$. Here $A^\dagger$ is the *adjoint* of $A$, a matrix obtained from $A$ by first constructing $A^T$, the *transpose* of $A$, and then taking the complex conjugate of each element (or by first taking

the complex conjugate of each element and then transposing the matrix). The determinant of a unitary matrix is 1.

**unitary operator** A linear operator $\mathbf{A}$ on a Hilbert space that preserves the inner product, thus the distance. See also *unitary matrix*.

**universal set of quantum gates** A set of gates with the property that there exists a network of them capable to implement every single unitary operation.

**von Neumann, John (1903 – 1957).** Eminent Hungarian-born mathematician who has made contributions to Quantum Physics, functional analysis, set theory, topology, economics, computer science, numerical analysis, hydrodynamics (of explosions), statistics and many other mathematical fields. von Neumann was a pioneer of the application of operator theory to quantum mechanics [442] and the co-creator of game theory and the concepts of cellular automata and the universal constructor. John von Neumann was instrumental in the development of the theory of thermonuclear reactions and the development of the hydrogen bomb, along Edward Teller and Stanislaw Ulam. John von Neumann is one of the pioneers of computing; the "von Neumann architecture" implements a universal Turing machine, and the common "referential model" of specifying sequential architectures. The model uses a processing unit and a single separate storage structure to hold both instructions and data, thus the term "stored-program computer."

**von Neumann entropy** The *von Neumann entropy* $S$ of a quantum system $\mathcal{A}$ is a function of the density matrix: $S(\rho^{\mathcal{A}}) = -\mathrm{tr}\left(\rho^{\mathcal{A}} \log \rho^{\mathcal{A}}\right)$.

**wave function** Function describing the state of a stationary system and the evolution in time of a non-stationary system. See also the entry for *Schrödinger equation*.

**weakly self-dual code** See the entry for *self-dual code*.

**Wigner, Eugene (1902 –1995).** Distinguished Hungarian-born physicist and mathematician. Wigner laid the foundation for the theory of symmetries in quantum mechanics and in 1927 introduced what is now known as the Wigner D-matrix. He and Hermann Weyl introduced group theory into quantum mechanics. He received the Nobel Prize in Physics in 1963. In 1960, Wigner discussed the power of mathematics in his best-known essay outside Physics, the "Unreasonable Effectiveness of Mathematics in the Natural Sciences."