Time Operator

Time Operator

Innovation and Complexity

I. Antoniou

B. Misra

Z. Suchanecki

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.

New York / Chichester / Weinheim / Brisbane / Singapore / Toronto

Contents

Preface

This is an example preface. This is an example preface. This is an example preface. This is an example preface.

(?)I. PRIGOGINE

Brussels, Belgium

Acknowledgments

To all who helped us we owe a deep sense of gratitude especially now that this project has been completed.

A. M. S.

Introduction

In recent years there appeared a new powerful tool in spectral analysis and prediction of dynamical systems – the time operator method. The power of this method can be compared with Fourier analysis. It the allows spectral analysis of all dynamical systems for which the time evolution can be rigorously formulated in terms of semigroups of operators on some vector space.

Time operator has been introduced by B. Misra and I. Prigogine [Mi,Pr] as a self-adjoint operator T on a Hilbert space $\mathcal{H}$ which is associated with a group $\{V_t\}_{t\in I}$ of bounded operators on $\mathcal{H}$, called the group of evolution, through the commutation relation

$$V_t^{-1} T V_t = T + tI \,. \tag{I.1}$$

One of the reasons why time operators are important is that the knowledge of the spectral resolution of a time operator, for example the knowledge of the complete family of its eigenvectors, allows to decompose each state on its age components. Because of (I.1) the evolution of a state ρ is then nothing but a shift of its age components.

Rescaling the age eigenvalues, i.e. the internal time of a dynamical system, we can get different kinds of evolutions. This idea has been exploited in the Misra-Prigogine-Courbage theory of irreversibility [MPC] that reconciles the invertible unitary evolution of unstable dynamical systems with observed Markovian evolution and the approach to equilibrium.

Time operator has been originally associated with a particular class of dynamical systems called K-systems. The Hilbert space on which it was defined was the space L^2 of square integrable functions defined on the phase space of a dynamical system.

The condition that a dynamical system is a K-system is sufficient for the existence of a time operator but not necessary. It is also possible to define time operators for larger classes of dynamical systems such as exact systems [AStime,ASS].

Although the time operator theory has been developed for the purpose of statistical physics its applications and connections have gone far beyond this field of physics. One of such application is a new approach to the spectral analysis of evolution semigroups of unstable dynamical systems. Associating time operator with a large class of stochastic processes we can see the problems of prediction and filtering of stochastic processes from a new perspective and connect them directly with physical problems. In this article we shall present an interesting connection of time operator with approximation theory.

The first connection, although indirect, of the time operator with the approximation theory has been obtained through wavelets [AnGu,AStime]. An arbitrary wavelet multiresolution analysis can be viewed as a K-system determining a time operator whose age eigenspaces are the wavelet detail subspaces. Conversely, in the case of the time operator for the Renyi map the eigenspaces of the time operator can be expanded from the unit interval to the real line giving the multiresolution analysis corresponding to the Haar wavelet. However, the connections of time operator with wavelets are much deeper than the above mentioned. As we shall see later time operator is in fact a straightforward generalization of multiresolution analysis.

In order to connect time operator with approximation it is necessary to go beyond Hilbert spaces. One of the most important vector spaces from the point of view of application is the Banach space $\mathcal{C}_{[a,b]}$ of continuous functions on an interval $[a,b]$. The space of continuous functions plays also a major role in the study of trajectories of stochastic processes.

Time operator can be, in principle, defined on a Banach space in the same way as on a Hilbert space. However its explicit construction is in general a non-trivial task. Having given a nested family of closed subspaces of a Hilbert space we can always construct a self-adjoint operator with spectral projectors onto those subspaces. This is not true in an arbitrary Banach space. The reason is that it is not always possible to construct an analog of orthogonal projectors on closed subspaces. Even if a self-adjoint operator with a given family of spectral projectors is defined it can appear additional problems associated with convergence of such expansion and with possible rescalings of the time operator.

For some dynamical systems associated with maps the time operator can be extended from the Hilbert space L^2 to the Banach space L^p. This can be achieved by replacing the methods of spectral theory [MPC,GMC], by more efficient martingales methods. For example, for K-flows it is possible to extend the time operator from L^2 to L^1 including to its domain absolutely continuous measures on the the phase space [SuL1,Su]. Martingales methods can not be, however, applied for the space of continuous functions.

In this book we discuss connections of time operator with wavelets, especially those restricted to the interval $[0,1]$, and the corresponding multiresolutions analysis. We establish a link between the Shannon sampling theorem and the eigenprojectors of the time operator associated with the Shannon wavelet. We construct the time operator associated with the Faber-Schauder system on the space $\mathcal{C}_{[0,1]}$ and study its properties. Such time operator corresponds to the interpolation of continuous functions by polygonal lines. We give the explicit form of the eigenprojectors of this time operator and characterize the functions from its domain in terms of their modulus of continuity.

1

Predictability and innovation

Consider a physical system that can be observed through time varying quantities x_t, where t stands for time that can be discrete or continuous. The set $\{x_t\}$ can be a realization of a deterministic system, e.g. a unique solution of a differential equation, or a stochastic process. In the later case each x_t is a random variable. We are interested in the global evolution of the system, not particular realizations x_t, from the point of view of innovation. We call the evolution *innovative* if the dynamics of the system is such that there is a gain of information about the system when time increases. Our purpose is to associate the concept of internal time with such systems. The internal time will reflect about the stage of evolution of the system.

The concept of innovation is relatively easy to explain for stochastic processes. Consider, for example, the problem of prediction of a stochastic process $X = \{X_t\}$. We want to find the best estimation $\hat{X}$ of the value X_t in moment t_0, knowing some of the values X_s for $s < t_0$. If we can always predict the value X_{t_0} exactly, i.e. if $\hat{X} = X_{t_0}$ then we can say that there is no innovation. If, on the other hand $\hat{X} \neq X_{t_0}$ and if, for $s_1 < s_2$, the prediction $\hat{X} = \hat{X}(s_2)$ based on the knowledge of X_t for $t < s_2$ is "better" than the prediction $\hat{X} = \hat{X}(s_1)$ then we shall call such process innovative. As an example consider stochastic process $\{X_n\}$ where X_n is the number of heads after n independent tosses a coin. Suppose we know, say, the values of the first N tosses, i.e. $x_1, \ldots, x_N$ and want to predict the random variable X_{N+M}, $M \geq 1$. Because of the nature of this process ($\{X_n\}$ has independent increments) the best prediction $\hat{X}$ will not be exact, $\hat{X} \neq X_{N+M}$. Moreover, knowing some further values of $\{X_n\}$, say x_{N+1}, we can improve the prediction of X_{N+M}.

Consider now a deterministic dynamical system in which points of the phase space evolve according to a specified transformation. It means that the knowledge of the position x_{t_0} of some point at the time instant t_0 determines its future positions x_t for $t > t_0$. In principle, there is no place for innovation for such deterministic dynamical system. Let us however, consider two specific examples. Consider first the dynamics of pendulum (harmonic oscillator). The knowledge of its initial position x_0 and the direction of the movement at $t = 0$ allows to determine all the future positions x_t. It is obvious that, in this case, the knowledge of x_s, for $0 < s < t$ carries the same information about the future position x_t as x_0. Thus there is indeed no innovation in this dynamical system.

As another example consider billiards where x_t denote the position of some selected ball at the moment t. In principle, the same arguments as for pendulum can be applied. According to laws of classical mechanics the knowledge of the initial position and the direction of the ball at $t = 0$ determines its position at any time instant $t > 0$ (we neglect the friction). This is, however, not true in practice. Contrary to the harmonic oscillator the dynamics in billiards is highly sensitive on initial conditions. Even a very small change of initial conditions may lead very fast to big differences in the position of the ball. Compare, for example, 10 swings of the pendulum and 10 scattering of the billiards balls and suppose that in both cases it requires the same amount of time t_0. In the case of pendulum we can predict with same accuracy the position $x_{t_0+\Delta t}$ if x_{t_0} is known to any given accuracy, while in the case of billiards this is impossible in practice because initial accuracy can get amplified due to sensitivity on initial conditions. It is also obvious that the additional knowledge, say the position of the ball after 5 scattering will improve significantly our prediction.

The above examples show that while there is no innovation in the harmonic oscillator there must be some intrinsic innovation in highly unstable systems like billiards. Innovation is also connected with the observed direction of time. Indeed, suppose that knowing the position x_t of an evolving point at the time instant t we want to recover the position x_s, for some $s < t$. This is possible for harmonic oscillator but for billiards, because of sensitive dependence on initial conditions such time reversal is practically possible only for short time intervals $t - s$.

Our aim is to introduce criteria that will allow to distinguish innovative systems. Then we shall show that systems with innovations have their internal time that can be expressed by the existence of time operator. First, however, let us introduce rigorously some basic concepts and tools.

1.1 DYNAMICAL SYSTEMS

An abstract dynamical system consists of a phase space X of pure states x and a semigroup (or group) $\{S_t\}_{\in I}$ of transformations of $\mathcal{X}$ which describes the dynamics. We assume that $\mathcal{X}$ is equipped with a measure structure, which means that it is given a σ-algebra Σ of subsets of $\mathcal{X}$ and a finite measure μ on $(\mathcal{X}, \Sigma)$. The variable $t \in I$, which signifies time, can be either discrete or continuous. We assume that either $I = \mathbb{Z}$ or $I = \mathbb{N} \cup \{0\}$, for discrete time, and $I = \mathbb{R}$ or $I = [0, \infty)$, for continuous time.

For a given state $x \in \mathcal{X}$ the function

$$I \ni t \longmapsto S_t x \in \mathcal{X}$$

is the evolution of the point x (pure state) in time. We assume that $S_0 x = x$, i.e. S_0 is the identity transformation of $\mathcal{X}$. The semigroup property means that

$$S_{t_1+t_2} = S_{t_1} \circ S_{t_2} \ , \ \text{for all } t_1, t_2 \in I \ ,$$

where $\circ$ denotes the composition of two transformations. The semigroup property is the reflection of the physical property that the lows governing the behavior of the system do not change with time. If $\{S_t\}$ is a group then all the transformations are invertible and we have $S_t^{-1} = S_{-t}$. If this holds then we say that the dynamics on the phase spaces is *reversible* (or that the dynamical system is reversible).

Every transformation S_t is supposed to be measurable. If time is discrete, i.e. I is a subset of integers, then we shall consider a single transformation $S = S_1$ instead of the group $\{S_t\}$ because, according to the semigroup property we have

$$S_n = \underbrace{S_1 \circ \ldots \circ S_1}_{(n-\text{times})} = S^n \, .$$

If time is continuous then we assume additionally that the map

$$\mathcal{X} \times I \ni (x,t) \longmapsto S_t x \in \mathcal{X}$$

is measurable, where the product space is equipped with the product σ-algebra of Σ and the Borel σ-algebra of subsets of I.

Especially important from the point of view of ergodic theory are the dynamical systems with *measure preserving* transformations. This means that for each t

$$\mu(S_t^{-1} A) = \mu(A) \ , \ \text{for every } A \in \Sigma \, .$$

In particular, for reversible system $\{S_t\}$ is the group of measure automorphisms. In the case $\mu(\mathcal{X}) = 1$ the measure μ represents an *equilibrium distribution*.

Throughout this book it will be usually assumed that the measure μ is invariant with respect to the semigroup $\{S_t\}$ although we do not want to make such general assumption. However, we shall always assume that every S_t is *nonsingular*. This means that if $A \in \Sigma$ is such that $\mu(A) = 0$ then also $\mu(S_t^{-1} A) = 0$.

An important class of dynamical system arises from differential equations. For example, suppose it is given a system of equations:

$$\frac{dx_i}{dt} = F_i(x_1, \ldots, x_N) \ , \ i = 1, \ldots, N \tag{1.1}$$

where $x_k = x_k(t), \ k = 1, \ldots, N$ are differentiable real or complex valued function on $[0, \infty)$ and $F_i, \ i = 1, \ldots, N$ real or complex valued functions on $\mathbb{R}^N$. Suppose that for each $(x_1^0, \ldots, x_N^0) \in \mathbb{R}^N$ there exists a unique solution $(x_1(t), \ldots, x_N(t))$ of the system (1.1) that satisfies the initial condition

$$(x_1(0), \ldots, x_N(0)) = (x_1^0, \ldots, x_N^0) \, .$$

Then we can define the semigroup $\{S_t\}$ of transformations of $\mathbb{R}^N$ by putting

$$S_t(x_1^0, \ldots, x_N^0) = (x_1(t), \ldots, x_N(t)) \ , \ \text{for } t \geq 0 \, .$$

To be more specific, let us consider a physical system consisting of N particles contained in a finite volume. The state of the system at time t can be specified by the three coordinates of position and the three coordinates of momentum of each particle, i.e. by a point in $\mathbb{R}^{6N}$. Thus the phase space is a bounded subset of $6N$ dimensional Euclidean space. Simplifying the notation, let the state of the system be described by a pair of vectors (q, p) where $q = (q_1, \ldots, q_N)$, $p = (p_1, \ldots, p_N)$, thus by a point in $\mathbb{R}^{2N}$. Assume further that it is given a Hamiltonian function (shortly a *Hamiltonian*) $H(q, p)$, which does not depend explicitly on time, satisfying the following equations:

$$\frac{\partial q_k}{\partial t} = \frac{\partial H}{\partial p_k} \ , \quad \frac{\partial p_k}{\partial t} = -\frac{\partial H}{\partial q_k} \ , \ k = 1, \ldots, N \tag{1.2}$$

If the initial system state at $t = 0$ is (q, p), then the *Hamilton equations* (1.2) determine the state $S_t(q, p)$ at any time instant t. This is the result of the theorem on the existence and uniqueness of the solutions of first-order ordinary differential equations. In other words, the Hamiltonian equations determine uniquely the evolution in time on the phase space.

It follows from the Hamiltonian equations that $\frac{dH}{dt} = 0$. This implies that the dynamical system is confined to a surface in $\mathbb{R}^{2N}$ that corresponds to some constant energy E. Such surface is usually a compact manifold in $\mathbb{R}^{2N}$. Moreover it follows from the *Liouville theorem* that the Hamiltonian flow $\{S_t\}$ preserves the Lebesgue measure on this surface.

1.2 DYNAMICAL SYSTEMS ASSOCIATED WITH MAPS

There is a large class of dynamical systems associated with maps of intervals. The dynamics of such a system is determined by a function S, mapping an interval $[a, b]$ into itself. The phase space $\mathcal{X}$ is the interval $[a, b]$ and the dynamical semigroup consists of transformations S_n:

$$S_n(x) = \underbrace{S \circ \ldots \circ S}_{\text{(n−times)}}(x) \ , \ n = 1, 2 \ldots \ , \ S_0 = I$$

The time is, of course, discrete $n = 1, 2 \ldots$. If the map S is invertible, then the family $\{S_n\}_{n \in \mathbb{Z}}$ forms a group. The general assumption is that the map S is Borel measurable but in specific examples S turns out to be at least piecewise continuous.

Some dynamical systems associated with maps can be used as simplified models of physical phenomena. However one of the reasons for the study of such dynamical system is their relative simplicity. This allows to obtain analytical solutions of some problems, in particular, to test new tools of the analysis of dynamical systems. One of such tools is the time operator method that will be also tested on dynamical systems associated with maps.

We begin the presentation of dynamical systems arising from maps with, perhaps the simplest one, the Renyi map.

Renyi map

The *Renyi map* S on the interval $[0, 1]$ is the multiplication by 2, modulo 1.

$$Sx = 2x \,(\mathrm{mod}\ 1) \,.$$

A slightly more general is the β*-adic Renyi map*

$$Sx = \beta x \,(\mathrm{mod}\ 1) \,, \tag{1.3}$$

where β is an integer, $\beta \geq 2$.

The measure space corresponding to the dynamical system determined by the Renyi map consists of the interval $\Omega = [0, 1]$, the Borel σ-algebra of subsets of $[0, 1]$ and the Lebesgue measure. It can be easily verified that the Lebesgue measure is invariant with respect to S.

Logistic map

The *Logistic map* is the quadratic map

$$Sx = rx(1 - x) \tag{1.4}$$

on the interval $[0, 1]$, where r is a (control) parameter, $0 < r \leq 4$. Actually (1.4) defines the whole family of maps which behavior depends on the control parameter r. This behavior ranges from the stable contractive, for $r < 1$, to fully chaotic, for $r = 4$. For a thorough study of the logistic map we refer the reader to Schuster (1988). The logistic map has many practical applications. For example, in biology where it describes the growth of population in a bounded neighborhood. In this book we shall only consider the case of fully developed chaos $(r = 4)$. The map

$$Sx = 4x(1 - x) \,, \; x \in [0, 1]$$

is "onto" and admits an invariant measure with the density function

$$f(x) = \frac{1}{\pi\sqrt{x(1 - x)}} \,.$$

Cusp map

The *cusp map* is defined as the map S interval $[-1, 1]$

$$Sx = 1 - 2\sqrt{|x|} \,. \tag{1.5}$$

This map is an approximation of the Poincaré section of the Lorentz attractor. The absolutely continuous invariant measure of the cusp map has the density

$$f(x) = \frac{1 - x}{2}.$$

A characteristic feature of the Renyi, logistic and the cusp map is that all these maps of the interval are noninvertible. Let us present now one of the simplest invertible chaotic map.

Baker transformation

Let the phase space will be the unit square $\mathcal{X} = [0,1] \times [0,1]$. Define

$$S(x,y) = \begin{cases} \left(2x, \dfrac{y}{2}\right) , & \text{for } 0 \le x < \frac{1}{2},\ 0 \le y \le 1 \\ \left(2x-1, \dfrac{y}{2} + \dfrac{1}{2}\right) , & \text{for } \frac{1}{2} \le x \le 1,\ 0 \le y \le 1\,. \end{cases}$$

The action of S can be illustrated as follows. In the first step S compresses the square $\mathcal{X}$ along the y-axis by $\frac{1}{2}$ and stretches $\mathcal{X}$ along the x-axis by 2. Such compressed and stretched rectangle is then vertically divided on two equal parts and the right-hand part is placed on the left hand part. The inverse of the baker transformation

$$S^{-1}(x,y) = \begin{cases} \left(\dfrac{x}{2}, 2y\right) , & \text{for } 0 \le x < 1,\ 0 \le y < \frac{1}{2} \\ \left(\dfrac{x}{2} + \dfrac{1}{2}, 2y-1\right) , & \text{for } 0 < x \le 1,\ \frac{1}{2} < y < 1\,. \end{cases}$$

is defined everywhere on $\mathcal{X}$ except the lines $y = \frac{1}{2}$ and $y = 1$, i.e. except the set of the Lebesgue measure 0. Thus taking as a measure space the unit square $\mathcal{X}$ with the Borel σ-algebra Σ and the planar Lebesgue measure μ we obtain a reversible dynamical system $(\mathcal{X}, \Sigma, \mu; \{S_n\}_{n\in\mathbb{Z}})$. It can be also shown easily (see LM), what is obvious from the above illustration, that the baker transformation is invariant with respect to the Lebesgue measure.

Further examples of dynamical systems will appear successively in this book. Now we shall introduce the basic tools that will allow to study their behavior and introduce the ergodic hierarchy.

1.3 THE ERGODIC HIERARCHY

The time evolution of dynamical systems can be classified according to different ergodic properties that correspond to various degree of irregular behavior. We shall list below the most significant ergodic properties. A more detailed information can be found in textbooks on ergodic theory (Halmos 1956, Arnold and Avez 1968, Parry 1981, Cornfeld et al. 1982).

Let us consider a dynamical system $(\mathcal{X}, \Sigma, \mu, \{S_t\}_{t\in I})$, where the measure μ is finite and S_t - invariant, for every $t \in I$. We shall distinguish the following ergodic properties:

(I) Ergodicity

Ergodicity expresses the existence and uniqueness of an equilibrium measure. This means that for any t there is no nontrivial S_t-invariant subset of $\mathcal{X}$, i.e. if for some $t \in I$ and $A \in \Sigma$ $S_t(A) = A$ μ-a.e., then either $\mu(A) = 0$ or $\mu(A) = 1$. The ergodicity is equivalent to the condition

$$\lim_{\tau \to +\infty} \frac{1}{\tau} \int_0^\tau \mu(S_t^{-1}(A) \cap B) = \mu(A)\mu(B)\,, \tag{1.6}$$

for all $A, B \in \Sigma$. If time is discrete then condition (1.6) has to be replaced by:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(S^{-k}(A) \cap B) = \mu(A)\mu(B)\,. \tag{1.7}$$

(II) Weak mixing

Weak mixing is a stronger ergodic property than ergodicity. The summability of the integral in (1.6) is replaced by absolute summability

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau |\mu(S_t^{-1}(A) \cap B) - \mu(A)\mu(B))|\, dt = 0 \tag{1.8}$$

If time is discrete then condition (1.8) is replaced by

$$\lim_{n \to \infty} \frac{1}{n} \sum_{n=0}^{n-1} |\mu(S^{-k}(A) \cap B) - \mu(A)\mu(B)| = 0\,, \text{ for all } A, B \in \Sigma\,. \tag{1.9}$$

Note that condition (1.7) means the Cesaro convergence of the sequence $\{\mu(S^{-n}(A) \cap B)\}$. Similarly, condition (1.9) means the absolute Cesaro convergence of this sequence.

It is also interesting to note that condition (1.9) can be equivalently expressed as follows. For each $A, B \in \Sigma$

$$\lim_{\substack{n \to \infty \\ n \notin J}} \mu(S^{-n}(A) \cap B) = \mu(A)\mu(B)\,, \tag{1.10}$$

where J is a set of density zero, which may vary for different choices A and B. Recall that a set $J \subset \mathbb{N}$ has density zero if

$$\lim_{n \to \infty} \frac{\mathrm{card}(J \cap \{1, \ldots, n\})}{n} = 0\,.$$

Condition (1.10) has a straightforward generalization that leads to a stronger ergodic property, which is called mixing.

(III) Mixing

Mixing means that

$$\lim_{n \to \infty} \mu(S^{-n}(A) \cap B) = \mu(A)\mu(B)\,, \tag{1.11}$$

for all $A, B \in \Sigma$. For continuous time mixing means that

$$\lim_{t\to\infty} \mu(S_t^{-1}(A) \cap B) = \mu(A)\mu(B)\,.$$

The above three ergodic properties has been formulated regardless the transformations of the phase space are invertible or not. We introduce below two more ergodic properties that are complementary to each other. First property involves only non invertible transformations the second invertible transformations.

(IV) Exactness

A semigroup $\{S_t\}_{t\geq 0}$ of measure preserving transformation is called exact if

$$\bigcap_{t\geq 0} S_t^{-1}(\Sigma) \;\text{ is the trivial } \sigma\text{−field}\,, \tag{1.12}$$

i.e. consists only of the sets of measure 0 or 1. It is obvious that condition (1.12) is the same for both discrete and continuous time. An equivalent condition for exactness is the following. Suppose that the measure μ is normalized, i.e. $\mu(\mathcal{X}) = 1$, and let $\{S_t\}_{t\geq 0}$ be a semigroup of measure preserving transformations such that $S_t(A) \in \Sigma$ for $A \in \Sigma$. Then the dynamical system is exact if and only if

$$\lim_{t\to\infty} \mu(S_t(A)) = 1\;,\; \text{for all } A \in \Sigma,\; \mu(A) > 0\,. \tag{1.13}$$

It can be proved (see Lasota and Mackey 1994) that exactness implies mixing. This property will also follow from the equivalent characterizations of ergodic properties that will be presented below.

(V) Kolmogorov systems

The term *Kolmogorov system* will mean either a K-flow, when time is continuous, or a K-system, when time is discrete. An invertible dynamical system $(\mathcal{X},\Sigma,\mu,\{S_t\}_{t\in\mathbb{R}})$ is called the *K-flow* if there exist a sub-σ-algebra Σ_0 of Σ such that for $\Sigma_t \stackrel{\text{df}}{=} S_t(\Sigma_0)$ we have

(i) $\Sigma_s \subset \Sigma_t$, for all $s < t$, $s, t \in \mathbb{R}$
(ii) $\sigma(\bigcup_t \Sigma_t) = \Sigma$
(iii) $\bigcap_t \Sigma_t$ is the trivial σ-algebra denoted by $\Sigma_{-\infty}$.

A discrete counter part of *K-flow* will be called *K-system* (the terms *K-cascade* is also sometimes used).

Each Kolmogorov system is mixing (this fact will also become clear later on). Thus both exact and Kolmogorov system are the strongest in the ergodic hierarchy. An example of K-system is the baker map. This will be shown in the next section.

For illustration of differences between ergodic properties of dynamical system we refer to Halmos (1956).

1.4 EVOLUTION OPERATORS

The idea of using operator theory for the study of dynamical systems is due to Koopman (1931). He replaced the time evolution $x_0 \mapsto x_t = S_t x_0$ of single points from the phase space $\mathcal{X}$ by the evolution of linear operators $\{V_t\}$ (*Koopman operators*)

$$V_t f(x) \stackrel{\mathrm{df}}{=} f(S_t x)\,, \quad x \in \mathcal{X}\,,$$

acting on square integrable functions f.

Using evolution operators we do not lose any crucial information about the behavior of the considered dynamical systems because the underlying dynamics can be, as we shall see below, recovered from the evolution operators. But in operator approach we gain new methods of analysis of dynamical systems.

Another reason of using operators for the study of dynamical system is that for unstable dynamical systems it is easier to study the evolution of ensembles of points than the evolution of single points. Even for relatively simple dynamical system, such as the system associated with the logistic map, it is practically impossible to trace a single trajectory for a longer time, due to its erratic behavior and very sensitive dependence or initial conditions. Roughly speaking, we consider an initial set of points $\{x_k^0\}_{k=1}^N$, which can be described by a probability density, i.e. by a nonnegative integrable function ρ_0 such that

$$\int_A \rho(x)\,\mu(dx) \cong \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_A(x_k^0)\,,$$

for each $A \in \Sigma$.

Under the action of S_t the set $\{x_k^0\}_{k=1}^N$ is transformed into $\{x_k^t\}_{k=1}^N$ that can be described, in the above sense, by another density ρ_t. The transformation U_t that establishes the correspondence

$$\rho_t \longmapsto U_t \rho_0$$

can be defined as a linear operator (Frobenius-Perron operator) on the space of integrable functions.

Rigorously speaking, in the operator approach the phase space $(\mathcal{X}, \Sigma, \mu; \{S_t\})$ is replaced by the space of p-integrable functions $L^p = L^p(\mathcal{X}, \Sigma, \mu)\,,\ 1 \le p \le \infty$. The choice $p = 2$ is the most common since L^2 is a Hilbert space, where we have at our disposal a whole variety of powerful tools. However, when considering the evolution of probability densities the most natural and unrestrictive is the choice $p = 1$. In the operator approach we consider instead of the transformation S of the phase space $\mathcal{X}$ the evolution of probability densities under the transformation U defined on L^1 as follows. If $f \in L^1$, then Uf denotes such a function from L^1 which satisfies the equality

$$\int_A Uf(x)\mu(dx) = \int_{S^{-1}A} f(x)\mu(dx)\,. \tag{1.14}$$

Since we consider only non singular transformation, the proof of the existence of the transformation U follows easily from the Radon-Nikodym theorem (Lasota and Mackey 1994). For discrete time it is, of course, sufficient to consider a single transformation U. For continuous time we have a family $\{U_t\}_{t\in I}$ of transformation on L^1. The transformation U has the following properties

(1) $U(af + bg) = aUf + bUg$, for all $a, b \in \mathbb{R}$ and $f, g \in L^1$

(2) $f \geq 0 \Rightarrow Uf \geq 0$, for all $f \in L^1$

(3) $\int_{\mathcal{X}} Uf d\mu = \int_{\mathcal{X}} f d\mu$, for all $f \in L^1$.

Thus the transformation U is a bounded and positivity preserving linear operator on L^1. If the measure μ is normalized then U has additionally the property:

(4) $U1 = 1 \Leftrightarrow \mu$ is S-invariant

U is called the *Frobenius-Perron operator*. An arbitrary operator U on L^1 which satisfies (1)-(4) is called the *doubly stochastic operator*. The family $\{U_t\}_{t\in I}$ forms a semigroup or group on L^1 accordingly to the group or semigroup property of the flow $\{S_t\}_{t\in I}$. The operator U_0 is the identity operator.

It follows easily from (1)-(3) that U is a contraction on L^1, i.e.

$$\|Uf\|_{L^1} \leq \|f\|_{L^1} \,.$$

Therefore U is, in particular, a *Markov operator* (see Section 12) and $\{U_t\}_{t\in I}$ is a *Markov semigroup* on each L^1.

By the Riesz convexity theorem (Brown 1976) every doubly stochastic operator U on L^1 also maps L^p into L^p, for each p, $1 \leq p \leq \infty$, with $\|U\|_{L^p} \leq 1$, and with $\|T\|_{L^1} = \|T\|_{L^\infty} = 1$. The operator U defined on L^p, where $1 \leq p < \infty$, has a well-defined adjoint, i.e. a continuous linear operator U^* defined on L^q where $\frac{1}{p} + \frac{1}{q} = 1$ ($q = \infty$ for $p = 1$), such that

$$\langle Uf|g\rangle = \langle f|U^* g\rangle \ , \text{ for all } f \in L^p, \ g \in L^q \,.$$

Consider now the transformation V associated with a non singular transformation S of the phase space $\mathcal{X}$ by the formula

$$Vf(x) = f(Sx) \,,$$

where f is a measurable function on $\mathcal{X}$. It is easy to check that V, when considered as an operator on L^q, $1 < q \leq \infty$, is the adjoint of the Frobenius-Perron operator U defined on L^p, $1 \leq p < \infty$, i.e. $V = U^*$.

The operator V is called the *Koopman* operator associated with S. If the transformation S is measure preserving then the Koopman operator is a positivity preserving contraction on L^q. Moreover, V is, as the adjoint of a doubly stochastic operator, also doubly stochastic (Brown 1976). In the case of flows we shall consider groups or semigroups $\{V_t\}_{t\in I}$ of Koopman operators.

The operators U and V considered on the Hilbert space L^2 are adjoint to each other. If the transformation S is measure preserving then V is a partial isometry. The Koopman and Frobenius-Perron operators carry a similar information about the dynamical system. However, from the physical point of view the semigroup $\{U_t\}$

describes the evolution of states while $\{V_t\}$ describes the evolution of observables. This is an analog of Schrödinger versus Heisenberg picture of evolution in quantum mechanics. The advantage of using the Koopman operator V is that given a transformation S of the phase space $\mathcal{X}$ the explicit form of V is also known. The explicit form of the Frobenius-Perron operator can be derived in special cases. For example, if the phase space is an interval, say $\mathcal{X} = [a, b]$, then formula (1.14) implies that

$$Uf(x) = \frac{d}{dx} \int_{S^{-1}[a,x]} f(y)\, dy\ ,\ \text{for } x \in [a, b]\,. \tag{1.15}$$

In particular, if the transformation S is differentiable and invertible with the continuous derivative $\frac{dS^{-1}}{dx}$ we obtain the following explicit formula for U

$$Uf(x) = f(S^{-1}(x)) \left| \frac{d}{dx} S^{-1}(x) \right| . \tag{1.16}$$

An analog of equation (1.16) can be derived (Lasota and Mackey 1994) for invertible transformation S of $\mathbb{R}^n$, only the last factor of (1.16) has to be replaced by the Jacobian of the inverse S^{-1} of S. Particularly simple is the explicit form of the Frobenius-Perron operator in the case when transformation S is both invertible and μ-invariant. In this case it follows from (1.14) that $Uf(x) = f(S^{-1}x)$.

Applying formula (1.15) we obtain easily the explicit form of the Frobenius-Perron operator U of each noninvertible map from the above examples:

1. *The β-adic Renyi map* (1.3)

$$Uf(x) = \frac{1}{\beta} \sum_{k=0}^{\beta-1} f\left(\frac{x+k}{\beta}\right) \tag{1.17}$$

2. *The logistic map* (1.4) with the control parameter $r = 4$

$$Uf(x) = \frac{1}{4\sqrt{1-x}} \left[f\left(\frac{1-\sqrt{1-x}}{2}\right) + f\left(\frac{1+\sqrt{1-x}}{2}\right) \right] \tag{1.18}$$

3. *The cusp map* (1.5)

$$\begin{aligned} Uf(x) = {} & \frac{1}{2}\left(1 - \left(\frac{1-x}{2}\right)^2\right) f\left(\left(\frac{1-x}{2}\right)^2\right) \\ & + \frac{1}{2}\left(1 + \left(\frac{1-x}{2}\right)^2\right) f\left(-\left(\frac{1-x}{2}\right)^2\right) \end{aligned} \tag{1.19}$$

4. *The baker transformation* S is invertible and preserves the Lebesgue measure. Therefore the Frobenius Perron operator is

$$Uf\,(x, y) = \begin{cases} f\left(\frac{x}{2}, 2y\right)\ , & \text{for } 0 \le x \le 1,\ 0 \le y < \frac{1}{2} \\ f\left(\frac{x}{2} + \frac{1}{2}, 2y - 1\right)\ , & \text{for } 0 < x \le 1,\ \frac{1}{2} \le y \le 1\,. \end{cases} \tag{1.20}$$

1.5 ERGODIC PROPERTIES OF DYNAMICAL SYSTEMS – OPERATOR APPROACH

Now, let us focus our attention on the characterization of ergodic properties of dynamical systems in terms of evolution operators. We begin with the question of existence of invariant measures.

For a given measure space $(\mathcal{X}, \Sigma, \mu)$ and a measurable transformation $S : \mathcal{X} \to \mathcal{X}$ the Frobenius-Perron operator is correctly defined regardless the measure μ is S-invariant or not. It follows immediately from (1.14) that μ is S-invariant if and only if $U1 = 1$. If ν is another measure on $(\mathcal{X}, \Sigma)$ which is absolutely continuous with respect to μ then by the Radon-Nikodym theorem there is an integrable function f_ν (the Radon-Nikodym derivative) such that

$$\nu(A) = \int_A f_\nu(x)\mu(dx)\ , \text{ for each } A \in \Sigma\,. \tag{1.21}$$

It is also an east consequence of (1.14) that measure ν is S invariant if and only if the function f_ν is a fixed point of U – the Frobenius-Perron operator with respect to μ. Indeed, if $Uf_\nu = f_\nu$, then $\int_A Uf_\nu\, d\mu = \int_{S^{-1}A} f_\nu\, d\mu$, and consequently

$$\nu(A) = \int_A f_\nu\, d\mu = \int_{S^{-1}A} f_\nu\, d\mu = \nu(S^{-1}A)\,,$$

for each $A \in \Sigma$. This means that ν is S-invariant. Repeating these arguments in the opposite direction we conclude the proof.

Let us consider now Frobenius-Perron and Koopman operators on the Hilbert space L^2. Note first that if the underlying transformation S is measure preserving then the Koopman operator V is an isometry on L^1:

$$\|Vf\|_{L^1} = \|f\|_{L^1}\,. \tag{1.22}$$

This fact can be proved directly by taking first as f the indicator of a set $A \in \Sigma$. In such case equality (1.22) reduces to

$$\mu(S^{-1}A) = \mu(A)\,.$$

Then the isometry property of a simple function can be derived by the linearity of V, and, for an arbitrary $f \in L^1$ by the standard arguments of the approximation of integrable functions by simple functions. From this we can derive that V is also an isometry on L^2

$$\|Vf\|_{L^2} = \|f\|_{L^2}\,, \quad f \in L^2\,. \tag{1.23}$$

This follows from the fact that L^2 norm of a function f is the square root of the L^1 norm of f^2 and the following obvious property of Koopman operators:

$$|Vf|^2 = |f|^2\,.$$

If the transformation S is measure preserving and invertible then V is an invertible isometry on L^2. Thus V is a unitary operator on L^2. This implies that

$$V = U^{-1}\,.$$

The description of evolution in terms of Koopman or Frobenius-Perron operators on L^p-spaces does not lead to any loss of information about the underlying dynamics. Indeed, according to the Banach-Lamperti theorem (Banach 1932, Lamperti 1958) any isometry on L^p, $1 \leq p < \infty$, $p \neq 2$, can be implemented by a measure preserving transformation. In the case $p = 2$, an isometry V on L^2 is implemented by a measure preserving point transformation if and only if V is positivity preserving (Goodrich et al. 1980).

The ergodic properties of measure preserving transformations can be expressed as properties of the corresponding Koopman operators as follows. Let f, g be two arbitrary functions from L^2 then
(I') $\{S_t\}_{t\geq 0}$ is *ergodic* if and only if $\langle f|V_t g\rangle$ is Cesaro summable:

$$\lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau \langle f|V_t g\rangle\, dt = \langle f|1\rangle\langle 1|g\rangle$$

(II') $\{S_t\}$ is *weak mixing* if $\langle f|V_t g\rangle$ is absolutely Cesaro summable:

$$\lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau |\langle f|V_t g\rangle - \langle f|1\rangle\langle 1|g\rangle|\, dt = 0$$

(III') $\{S_t\}$ is *mixing* if $\langle f|V_t\rangle$ is convergent:

$$\lim_{t\to\infty} \langle f|V_t g\rangle = \langle f|g\rangle\,.$$

(IV') $\{S_t\}$ is *exact* if

$$\bigcap_{t\geq 0} V_t(L^2) = [1]\,.$$

(V') The property of K-flow can be described in terms of $\{V_t\}$ in a similar way as for transformation S_t. Namely, denote by $\mathcal{H}_t$ the orthogonal complement of 1 (constant) in the space $L^2(\mathcal{X}, \Sigma_t, \mu)$ and by $\mathcal{H}$ the orthogonal complement of 1 in $L^2(\mathcal{X}, \Sigma, \mu)$. Then $\mathcal{H}_t = V_{-t}\mathcal{H}_0$ and conditions (i)-(iii) characterizing K-flows take the form
(i) $\mathcal{H}_s \subset \mathcal{H}_t$, for $s < t$
(ii) $\overline{\lim \bigcup_{t\in\mathbb{R}} \mathcal{H}_t} = \mathcal{H}$ (bar denotes the closure)
(iii) $\bigcap_{t\in\mathbb{R}} \mathcal{H}_t = \{0\}$.

If time is discrete the integrals in (I') and (II') have to be replaced by sums. Since the Frobenius-Perron operator is the adjoint of the Koopman operator the ergodic properties (I)-(IV) can be equivalently formulated in terms of the Frobenius-Perron operators $\{U_t\}$. Exactness can be defined as follows
(IV') $\{U_t\}$ is exact if

$$\lim_{t\to\infty} \left\| U_t f - \int f d\mu \right\|_{L^2} = 0\,,\ \text{for each } f \in L^2\,.$$

1.6 INNOVATION AND TIME OPERATOR

We begin with the concept of information and the strictly related concept of entropy as the measure of disorder. Both these concepts have a long history. The concept of entropy was introduced in thermodynamics already in 1854 by Clausius and then Boltzmann found its logarithmic form. The first step towards the definition of information were taken by Hartley (1928) and elaborated by Shannon (1948) and Wiener (1948). Finally Kolmogorov (1958) introduced the concept of entropy to ergodic theory.

There is several axiomatic characterizations of information and entropy. We shall introduce information as a function associated with a finite or countable partition and entropy as its mean value. First, however, we start from some intuitive considerations.

In the dynamical system $(\mathcal{X}, \Sigma, \mu, \{S_t\})$ the σ-algebra Σ of subsets of $\mathcal{X}$ can be interpreted as all available information about the system. Similarly, for a given probability space $(\Omega, \mathcal{F}, P)$, which will be called the *system* for a while, the σ-algebra $\mathcal{F}$ represents all possible events that can occur as the outcomes of an experiment. By distinguishing σ-algebra $\mathcal{F}$ we assume tacitly that no other events are possible. Thus we can say that $\mathcal{F}$ carries the whole information about the system. If we now consider a sub-σ-algebra $\mathcal{F}_0$ of $\mathcal{F}$ then it is natural to expect that $\mathcal{F}_0$ carries some partial information about the system. If $\mathcal{F}_1, \mathcal{F}_2$ are two sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_1 \subset \mathcal{F}_2$ then we expect that $\mathcal{F}_2$ carries more (or at least not less) information then $\mathcal{F}_1$. We may assume that the information carried by $\mathcal{F}_0$ is 0, when $\mathcal{F}_0$ is the trivial σ-algebra $\{\emptyset, \Omega\}$.

In probability theory sub-σ-algebras are often generated by random variables or by families of random variables, e.g. if $X : \Omega \to \mathbb{R}$ is a random variable then $\mathcal{F}_0$ may be defined as the counterimage of the Borel σ-algebra on $\mathbb{R}$. In this case the information carried by $\mathcal{F}_0$ may be identified with the information associated with the observation of the random variable X. Such association of the information with random variables leads to one more natural postulate. Consider two random variables X_1 and X_2. If X_1 and X_2 are independent then it is natural to expect that the information carried by $\mathcal{F}_1 \cup \mathcal{F}_2$ (the smallest σ-algebra containing $\mathcal{F}_1$ and $\mathcal{F}_2$) is the sum of information carried by $\mathcal{F}_1$ and $\mathcal{F}_2$ separately.

In order to ascribe rigorous meaning to this intuitive concept of information let us confine our consideration to σ-algebras generated by finite or countable partitions. Thus let us consider a partition $\pi = \{A_1, A_2, \ldots\}$ of the probability space Ω on pairwise disjoint sets $A_i \in \mathcal{F}$, where $\mu(A_i) > 0$ for each i. Suppose now we want to localize a point $\omega_0 \in \Omega$ but the only available information is that the point is in the set $A_{i_0} \in \pi$.

This allows to locate ω_0 with the accuracy that depends only on the measure of A_{i_0} i.e. on $P(A_{i_0})$. Therefore the available information depends only on $P(A_{i_0})$ and can be defined as some function of $P(A_{i_0})$, say $f(P(A_{i_0}))$.

Using the above postulate about the information carried by two σ-algebras corresponding to independent random variables we can find the explicit form of the function f. Note that the independence expressed in terms of partitions π_1 and π_2 means $P(A \cap B) = P(A)P(B)$ for each $A \in \pi_1$ and $B \in \pi_2$.

The partition π generated by π_1 and π_2 is in fact generated by the intersections $A \cap B$, $A \in \pi_1, B \in \pi_2$ with $P(A \cap B) > 0$. If π_1 and π_2 are independent we should have

$$f(P(A)P(B)) = f(P(A \cap B)) = f(P(A)) + f(P(B))$$

.

This restricts our choice of f to the logarithmic function (at least in the class of continuous functions). Therefore the *information function* $\mathcal{I}_\pi$ *of the partition* π has the form

$$\mathcal{I}_\pi(x) \stackrel{\text{df}}{=} -\sum_{A \in \pi} \mathbb{1}_A(x) \log P(A)\,.$$

The mean value of the information $\mathcal{I}_\pi$ is called the *entropy of the partition* π (or simply the *entropy of* π) and will be denoted by $H(\pi)$. Thus we have

$$H(\pi) = \sum_{A \in \pi} P(A) \log P(A)\,. \tag{1.24}$$

Formula (1.24) is an analog of Shannon's entropy associated with random variables. Indeed, if X is a random variable assuming values $x_1, \dots, x_n$ with probability $p_1, \dots, p_n$ respectively, then the Shannon's entropy $H(X)$ is

$$H(X) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}\,. \tag{1.25}$$

Using the concept of coding, instead of random variables, Shannon's formula (1.25) can be interpreted as follows. If we have N independent observations of X and code them with $0-1$ sequences then the average amount of zeros and ones for the coding of one information can be expressed by (1.25).

It is therefore the entropy, which is the quantitative measure of information carried by partitions or, generalizing, by σ-algebras. It can be proved that H, as a function of partitions, satisfies all the above postulates on information. The proof requires, however, an elaboration of the concept of information function and entropy. Namely the introduction of the concept of conditional information and entropy. We shall not pursue this subject further. The interested reader can find the proof and many other interesting results in Ref. [MartinEngland].

We have shown above how to associate σ-algebras with the information about a system. We shall show now how to associate information with the flow of time.

Consider first a stochastic process $\{X_t\}$ on a probability space $(\Omega, \mathcal{F}, P)$ and its natural filtration $\{\mathcal{F}_t\}$. Each σ-algebra $\mathcal{F}_t$ of the filtration $\{\mathcal{F}_t\}$ expresses the information about the "past" until the moment t. In particular, $\{\mathcal{F}_t\}$ expresses the gain of information in time about the process $\{\mathcal{X}_t\}$ when time increases.

Actually the concept of filtration refers not only to an increasing family of σ-algebras but can be introduced on an arbitrary measure space. We have just associated filtrations with stochastic processes that is with sequences or flows of random events occurring with time. We have, however, already encountered such increasing families of events in dynamical systems. For example, each K-system possesses a natural filtration. Therefore, in more generals terms, by a filtration on a probability space $(\Omega, \mathcal{F}, P)$ we shall mean a family $\{\mathcal{F}_t\}$ of sub-σ-algebras of $\mathcal{F}$ such that

$$\mathcal{F}_s \subset \mathcal{F}_t \quad \text{for } s < t\,.$$

The σ-algebra $\mathcal{F}_t$ will be interpreted as the set of events observed up to the time instant t, or, as the information about the system at the moment t.

Suppose now that each σ-algebra contained in filtration $\{\mathcal{F}_t\}$ is generated by a countable partition. This allows to define entropy $H(\mathcal{F}_t)$ as the entropy of the generating partition. Since the entropy is a quantitative measure of information carried by a σ-algebra we can say that the filtration is *innovative* if

$$H(\mathcal{F}_s) < H(\mathcal{F}_t)\ , \quad \text{for each } s < t\,.$$

In this way, by introducing entropy, we give a rigorous meaning to innovation in each system having countably generated filtration. However, from the practical point of view entropy, as the criterion of innovations, is not so useful as it seems to be. This concerns especially stochastic processes. Although each stochastic process determines a filtration it does not mean that the σ-algebras are generated by countable partitions. Even if they are the task of finding the partitions and then calculating their entropy is not feasible. Innovative stochastic processes will be described by other means in forthcoming sections. For the rest of this section we shall focus our attention only on dynamical systems where the above introduced ideas can be easier applied.

Let us consider a phase space $(\mathcal{X}, \Sigma, \mu)$ with the normalized measure μ and a measure preserving transformation S on $\mathcal{X}$. In other words we consider the discrete time dynamics described by the semigroup $\{S^n\}_{n\geq 0}$ (or the group $\{S^n\}_{n\in\mathbb{Z}}$ if S is invertible).

Our first task is to distinguish innovative dynamical systems. We know already that K-systems are good candidates for innovative systems, because they have natural filtrations. However, it would be desired to have a more general approach applicable for other dynamical systems. Indeed, it turns out that using the concept of entropy it is possible to verify whether the dynamics, i.e. the transformation S is sufficiently unstable to give rise to innovations.

Consider a finite partition $\pi = \{A_1, \ldots, A_n\}$ of the phase space $\mathcal{X}$. Denote by $S^{-1}\pi$ the partition $\{S^{-1}A_1, \ldots, S^{-1}A_n\}$. Let $\pi \vee S^{-1}\pi \vee \ldots \vee S^{-k+1}$ denotes, for each $k = 1, 2, \ldots$, the refinement of the partitions $\pi,\ S^{-1}\pi, \ldots,\ S^{-k+1}\pi$. The value

$$k(\pi, S) = \lim_{k\to\infty} \frac{1}{k} H(\pi \vee S^{-1}\pi \vee \ldots \vee S^{-k+1}\pi) \tag{1.26}$$

is called the *entropy of the measure preserving transformation* S with respect to the partition π. The proof that the limit of (1.26) exists can be found in many textbooks on ergodic theory (see Arnold 1978 or Cornfeld et al. 1982, for example).

The maximal value $h(S)$ of $h(\pi, S)$ over all finite partitions, i.e.

$$h(S) = \sup_{\pi} h(\pi, S)$$

is called the *entropy of the transformation S* or the *Kolmogorov-Sinai entropy* (shortly *K-S entropy*). The intuitive meaning of the value $h(S)$ is a quantitative measure of disorder resulting from consecutive applications of transformation S. The innovative dynamical systems can be characterized as those which have positive K-S entropy. It can be proved that each K-system has positive K-S entropy. But the class of dynamical system with positive K-S entropy is larger. Exact systems also have positive K-S entropy (Rokhlin 1967).

The notion of K-S entropy is of great significance in ergodic theory mainly due to the fact it is an invariant of isomorphisms. If two dynamical systems are isomorphic then they have the same K-S entropy. One of the crucial facts concerning K-S entropy is the Kolmogorov-Sinai theorem which says that if π is a generating partition for S, then

$$h(S) = h(\pi, S)\,.$$

The term *generating partition for S* means that the σ-algebra generated by π and all the iterations. $S^{-k}\pi$ coincides with Σ. This theorem allows the calculation of K-S entropy in practice.

It follows from the above considerations that K-S entropy can be, in principle, applied for verifying whether a dynamical system is innovative. However, we shall not make a wider use of K-S entropy in this book. The reason is that there is no satisfactory characterization of dynamical systems with positive K-S entropy. We shall thus consider separately two particular cases: K-systems and exact systems, associating with them time operators. Nevertheless, we hope that by relating entropy with innovations, the later concept has acquired clear intuitive meaning.

Summarizing, recall that in an innovative system we distinguish a filtration, i.e. an increasing family of σ-algebras $\{\Sigma_t\}$. Since the variable t signifies time, called the *external time*, each σ-algebra Σ_t represents the information about the system that is available at the time instant t. Another possible interpretation is that each Σ_t represents the stage of development or age. This is because of a constant development of the system as time t increases. Each σ-algebra Σ_t represents the *internal age* at the time instant t and the family $\{\Sigma_t\}$ represents the *internal time* of the system.

It has been argued in Section 2.4 that a convenient approach to the study of dynamical systems is to replace the phase space by the space of integrable functions and to consider time evolution of Koopman or Frobenius-Perron operators instead of evolution of single points. We shall show below that the concept of internal time can also be conveniently expressed in terms of operators.

Note first that since σ-algebra Σ_t represents the (internal) age we can say accordingly that each function f, which is Σ_t measurable, has the age at most Σ_t, or, identifying Σ_t with its label t, we can say that the age of f, or the stage of development of f is at most t.

In order to describe the internal time on the space $L^p_{\mathcal{X}}$, $p \geq 1$, let us consider instead of $\{\Sigma_t\}$ the family of conditional expectations $\{E_t\}$, where

$$E_t f = \; E(f\,|\Sigma_t), \text{ for } \; f \in L^p_{\mathcal{X}}\,.$$

Confining our consideration to the Hilbert space $\mathcal{H} = L^2_{\mathcal{X}} \ominus [1]$ we can say that the flow of internal time is associated with the family $\{E_t\}$ of orthogonal projectors on $\mathcal{H}$. Now the condition that the age of f is at most t amounts to checking that $E_t f = f$ a.e.

Recall that time ranges either between 0 and ∞ or between $-\infty$ and $+\infty$. It follows from the assumption that $\{\Sigma_t\}$ is a filtration that

$$E_s E_t = E_{s \wedge t}\ , \text{ where } s \wedge t = \min\{s, t\}\,. \tag{1.27}$$

Without a loss of generality we may also assume that the projectors E_t satisfy additionally:

$$E_0 = 0 \;\; (\text{or } E_{-\infty} = 0) \;\; \text{and} \;\; E_\infty = I \tag{1.28}$$

If time is continuous we shall also assume the right continuity of the family $\{E_t\}$ in the sense of strong convergence, i.e.

$$E_{t+0} = E_t\ , \text{ for each } t \tag{1.29}$$

If time is discrete we just extend E_t on $\mathbb{R}$ in such way that (1.29) is satisfied.

Condition (1.28 amounts to assuming that: The "first" σ-algebra in the filtration is trivial, and the "last" coincides with the whole σ-algebra Σ. The condition (1.29) is equivalent to

$$\bigcap_{t'>t} \Sigma_{t'} = \Sigma_t\,.$$

If the condition (1.27)-(1.29) are satisfied then $\{E_t\}$ is a resolution of identity in $\mathcal{H}$. Thus the family $\{E_t\}$ determines the operator

$$T = \int_{-\infty}^{\infty} t dE_t \tag{1.30}$$

called the *time operator associated with the filtration* $\{\Sigma_t\}$ (if t ranges from 0 to ∞ then we put $E_t = 0$ for $t \leq 0$).

In order to clarify the concept of time operator let us assume for a while that time is discrete, say $t \in \mathbb{Z}$. Denote by P_n the orthogonal projection on the space

$$\mathcal{H}_n = L^2_{\mathcal{X}}(\Sigma_n) \ominus L^2_{\mathcal{X}}(\Sigma_{n-1})\,,$$

i.e.

$$P_n = E_n - E_{n-1} \quad , \quad n \in \mathbb{Z}\,.$$

In this case T has the form

$$T = \sum_{n=-\infty}^{\infty} n P_n\,.$$

Consider now a function $\rho \in \mathcal{H}$ such that $\|\rho\| = 1$. If $P_{n_0}\,\rho = \rho$, for some n_0, then according to the above identification of Σ_{n_0} with n_0 we may say that the age of ρ is n_0 and we also have $n_0 \langle P_{n_0}\rho | \rho\rangle \; = n_0 \|\rho\|^2 = n_0$.

If $P_{n_0}\rho \neq \rho$, then ρ does not have well-defined age since there is also a contribution from another age components $P_n\rho,\ n \neq n_0$. The value of this contribution is $n\langle P_n\rho|\rho\rangle$. The numbers $\langle P_n\rho|\rho\rangle$ are non-negative and $\Sigma_n\langle P_n\rho|\rho\rangle = 1$. Thus the meaning of the expression:

$$\langle T\rho|\rho\rangle = \sum_{n=-\infty}^{\infty} n\,\langle P_n\rho|\rho\rangle$$

is the *average age* of ρ.

The value $n \in \mathbb{Z}$ are thus *eigenvalues* of T and $\{P_n\}$ the eigenprojectors. In general, since $\{E_t\}$ is a spectral family, the operator T (1.30) is self-adjoint. This fact follows from the general properties of spectral families (see Appendix 1). It will be useful however to show explicitly a dense subset of the domain $D(T)$ of T. If time ranges from 0 to ∞ then each simple function $f \in \mathcal{H}$, which is $\bigcup_n \Sigma_n$ measurable, belongs to $D(T)$. Indeed, if $f = \sum_{i=1}^{k} a_i \mathbb{1}_{A_i}$, $A_i \in \bigcup_t \Sigma_t$, then there is t_0 such that $A_i \in \Sigma_{t_0}$, for $i = 1, \dots,\ k$. Thus Tf is well defined and finite

$$Tf = \int_0^{t_0} t\, dE_t f\,.$$

If time ranges from $-\infty$ to $+\infty$ then taking f as above and putting

$$g_s = f - E_s f\,,$$

where $s < t_0$, we check easily that

$$Tg_s = \int_s^{t_0} t\, dE_t g_s$$

The density of the family $\{g_s\}$ in $\mathcal{H}$ follows from the property: $E_s f \to 0$ as $s \to -\infty$.

In this way we have shown that an innovative system endowed with a filtration determines a self-adjoint operator on the Hilbert space of square integrable functions. Such operator will be called the internal time operator. In the next sections we shall elaborate on this concept.

2

Time operator of Kolmogorov systems

Let us recall that dynamical system is a*K-flow* if there exists a sub-σ-algebra Σ_0 of Σ such that for $\Sigma_t = S_t(\Sigma_0)$ we have

(i) $\Sigma_s \subset \Sigma_t$, for $s < t$

(ii) $\sigma(\bigcup_{t\in\mathbb{R}}\Sigma_t) = \Sigma$

(iii) $\bigcap_{t\in\mathbb{R}}\Sigma_t = \Sigma_{-\infty}$ – the trivial σ-algebra, i.e. the algebra of sets of measure 0 or 1

where $\sigma(\bigcup_{t\in\mathbb{R}} \Sigma_t)$ stands for σ-algebra generated by all $\Sigma_t,\ t \in \mathbb{R}$.

If time is discrete then we shall use the term *K-system* instead. Recall that both K-flows and K-systems are called *Kolmogorov systems*. Kolmogorov systems are time reversible dynamical systems which exhibit the strongest form of instability among irreversible systems, according to the ergodic hierarchy introduced in the previous section. In the previous section we also presented a simple example of a K-system that results from the baker transformation. Let us, however, add that a great variety of systems of physical interest, including infinite ideal gas, hard rods system, a hard sphere gas and geodesic flow on a compact, connected smooth Riemanian manifold with negative curvature are known to be K-flows.

Geodesic flows have been studied since the beginning of the 20th century, starting with Hadamard. In the case the manifold $\mathcal{X}$ has $\dim \mathcal{X} = 2$, the geodesic flow has an especially simple interpretation. It describes the notion of a point that moves on the surface $\mathcal{X}$ in the absence of external forces and without friction. The behavior of a geodesic flow depends on the properties of the manifold $\mathcal{X}$ and most of all on its curvature. When the curvature of $\mathcal{X}$ is negative, this behavior is highly chaotic.

Kolmogorov systems are natural examples of innovative dynamical system, which we defined as the systems with positive K-S entropy. The proof of the well known fact that Kolmogorov systems have positive K-S entropy can be found, for example, in Cornfeld et al. 1982.

The time evolution in a K-flow can be described by the unitary group $\{U_t\}$ on the Hilbert space $L^2_{\mathcal{X}}$

$$U_t\rho(x) = \rho(S_{-t}x) \tag{2.1}$$

The generator L of this group is called the *Liouvillian* operator

$$U_t = e^{-itL}\,.$$

For Hamiltonian systems generator L is given by the Poisson brackets

$$L\rho = i\{H, \rho\}\,.$$

With any K-flow we associate the family $\{E_t\}$ of conditional expectations regarded as orthogonal projectors on the Hilbert space $\mathcal{H} = L^2_{\mathcal{X}} \ominus [1]$. Because of K-flow's properties $\{E_t\}$ is a spectral family (see Appendix 2) and the operator

$$T = \int_{-\infty}^{\infty} t\, dE_t\,, \tag{2.2}$$

which is the time operator of the K-flow. In the case of a K-system, i.e. discrete time, the time operator assumes the form

$$T = \sum_{n=-\infty}^{\infty} n(E_n - E_{n-1})\,. \tag{2.3}$$

We shall now relate the time operator T of a K-flow with the dynamics described by the group $\{U_t\}$. We begin with a simple but important imprimitivity condition.

Proposition 2.1 *For each $s, t \in \mathbb{R}$ we have the following relation on $L^2_{\mathcal{X}}$.*

$$E_{s+t}\, U_t = U_t E \tag{2.4}$$

Remark. In the proof of the above proposition and in the proofs of other properties of time operators on $L^2_{\mathcal{X}}$ we shall use rather probabilistic techniques and reasoning instead of the spectral properties of operators on Hilbert spaces. For example, properties of conditional expectations or martingale techniques. This will allow to simplify proofs in some cases, but a more important reason is that this technique will also allow to extend, without major problems, the results concerning time operator on Banach spaces $L^p,\ 1 \le p < \infty$.

Proof of Proposition 2.1 Recall that E_t is the conditional expectation with respect to Σ_t, $E_t\rho = E(\rho|\Sigma_t)$. It follows from (2.1) that U_t transforms Σ_s measurable functions into Σ_{s+t} measurable. If $\rho \in L^2_{\mathcal{X}}$, then both $E_{s+t}U_t\rho$ and $U_tE_s\rho$ are Σ_{s+t} measurable. Thus in order to prove (2.4) it suffices to show that for each $A \in \Sigma_{s+t}$

$$\int_A E_{s+t}U_t\rho\, d\mu = \int_A U_tE_s\rho\, d\mu$$

Moreover, since the $sigma$-algebra Σ_{s+t} is the same as $S_t\Sigma_s$, we can assume that $A = S_t(B)$, where $B \in \Sigma_s$. Therefore, using the fact that S_t is a measure preserving transformation and $E_t\rho$ is a regular martingale we have

$$\begin{aligned}
\int_A U_t E_s \rho(\omega)\,\mu(d\omega) &= \int_A E_s\rho(S_{-t}\,\omega)\,\mu(d\omega) \\
&= \int_B E_s\rho(\omega)\,\mu(d\omega) \\
&= \int_B E_\infty\rho(\omega)\,\mu(d\omega) \\
&= \int_A U_t E_\infty\rho(\omega)\,\mu(d\omega) \\
&= \int_A E_\infty U_t\rho(\omega)\,\mu(d\omega) \\
&= \int_A E_{s+t} U_t\rho(\omega)\,\mu(d\omega)\,,
\end{aligned}$$

which proves (2.4). The next proposition follows directly from the properties of spectral families presented in Section 2.

Proposition 2.2 *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a Borel measurable function and let*

$$D(f(T)) = \{\rho \in L^2_{\mathcal{X}} : \int_{-\infty}^{\infty} f^2(t)\, d\langle E_t\rho|\rho\rangle\}\,.$$

Then $D(f(T))$ *is a linear dense subspace of* $L^2_{\mathcal{X}}$ *and the operator* $f(T)$*:*

$$f(T) = \int_{-\infty}^{\infty} f(t)\, dE_t$$

is well defined on $D(f(T))$*.*

Corollary 2.1 *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a bounded Borel measurable function. Then* $f(T)$ *is a bounded operator on* $L^2_{\mathcal{X}}$

Proposition 2.3 *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a bounded and Borel measurable function then we have*

$$U_t\left(\int_{-\infty}^{\infty} f(s)\, dE_s\right) = \left(\int_{-\infty}^{\infty} f(s-t)\, dE_s\right) U_t \tag{2.5}$$

on $\mathcal{H}$*.*

Proof. Since f is bounded, $D(f(T)) = \mathcal{H}$ and also $D(f(sI - T)) = \mathcal{H}$, for each $s \in \mathbb{R}$ (I stands for the identity operator). Taking as f a simple function

$$f = \sum_{i=1}^{k} a_i \mathbb{1}_{[s_i, s_{i+1})}$$

we have

$$\begin{aligned} U_t\left(\int_{-\infty}^{\infty} f(s)dE_s\right) &= \sum_{i=1}^{k} a_i\left(U_tE_{s_{i+1}} - U_tE_{s_i}\right) \\ &= \sum_{i=1}^{k} a_i(E_{s_{i+1}+t} - E_{s_i+t})U_t \\ &= \left(\int_{-\infty}^{\infty} f(s-t)dE_s\right)U_t \end{aligned}$$

An arbitrary bounded and measurable function f is the limit in any L^2-norm of function

$$f_M(t) = f(t)\,\mathbb{1}_{[-M,M]}(t), \quad \text{as M} \to \infty\,.$$

In particular, for each $\rho \in \mathcal{H}$, $\int_{-\infty}^{\infty} |f_M - f|^2\, d\langle E_t\rho|\rho\rangle \to 0$, as $M \to \infty$. On the other hand the function f_M can be approximated in L^2-norm by simple functions. Combining these two facts we see that equality (2.5) is true for any bounded and Borel measurable function.

The following theorem relates time operator of a K-flow with its unitary dynamics.

Theorem 2.1 *Let $\{U_t\}$ be the unitary group of a K-flow and T the corresponding time operator. Then we have*

$$U_t(D(T)) \subset D(T) \tag{2.6}$$

$$TU_t = U_tT + tU_t \tag{2.7}$$

for each $t \in \mathbb{R}$.

Proof. Using unitarity of U_t and Proposition 2.1 we have

$$\begin{aligned} \langle E_sU_t\rho|U_t\rho\rangle &= \langle E_sU_t\rho|E_sU_t\rho\rangle \\ &= \langle U_tE_{s-t}\rho|U_tE_{s-t}\rho\rangle \\ &= \langle E_{s-t}\rho|E_{s-t}\rho\rangle \\ &= \langle E_{s-t}\rho|\rho\rangle\,, \end{aligned}$$

for each $\rho \in D(T)$. Then using the above equality and the elementary inequality: $(s+t)^2 \le 2s^2 + 2t^2$ we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} s^2\, d\langle E_sU_t\rho|U_t\rho\rangle &= \int_{-\infty}^{\infty} (s+t)^2\, d\langle E_s\rho|\rho\rangle \\ &\le 2\int_{-\infty}^{\infty} s^2\, d\langle E_s\rho|\rho\rangle + 2t^2\int_{-\infty}^{\infty} d\langle E_s\rho|\rho\rangle \\ &< \infty\,. \end{aligned}$$

This proves that $U_t\rho \in D(T)$ for each t. Note that we have also shown that for each t

$$D(T+tI) = D(T)\,.$$

Now, put $g(t) = t$ and $g_M(t) = t\,\mathbb{1}_{[-M,M]}(t)$. Applying Proposition 2.3 we have

$$U_t \left(\int_{-\infty}^{\infty} g_M(s)\, dE_s \right) = \left(\int_{-\infty}^{\infty} g_M(s-t)\, dE_s \right) U_t\,.$$

We also have

$$\lim_{M\to\infty} \left\| \int_{-\infty}^{\infty} g_M(s)\, dE_s\rho - \int_{-\infty}^{\infty} s\, dE_s\rho \right\|_{L^2_{\mathcal{X}}} = 0$$

and

$$\lim_{M\to\infty} \left\| \int_{-\infty}^{\infty} g_M(s-t)\, dE_s\rho - \int_{-\infty}^{\infty} (s-t)\, dE_s\rho \right\|_{L^2_{\mathcal{X}}} = 0\,,$$

for each $\rho \in D(T)$ (see Appendix 3 ???, about the strong convergence of truncated operators). Since U_t preserves the domain of $D(T)$ we have for each $\rho \in D(T)$

$$U_t \left(\int_{-\infty}^{\infty} s dE_s \rho \right) = \int_{-\infty}^{\infty} (s-t) dE_s U_t \rho$$

or, equivalently,

$$U_t T = T U_t - t U_t\,,$$

which gives (2.7).

In order to give the physical meaning to (2.7) let us consider a state $\rho \in \mathcal{H}$, i.e. $||\rho|| = 1$ and its time evolution $\rho_t = U_t\rho$ by (2.7) we have

$$\langle T\rho_t|\rho_t\rangle = \langle T\rho|\rho\rangle + t\,.$$

The above equality means that the average age of the evolved state keeps step with the external clock time t.

The above theorem suggests the following generalization of time operator. Let $\mathcal{H}$ be a (separable) Hilbert space and $\{U_t\}$ a unitary group of evolution. A self-adjoint operator T acting on $\mathcal{H}$ is called the time operator with respect to $\{U_t\}$ if it satisfies the conditions of Theorem 2.1, i.e.: $U_t D(T) \subset D(T)$ and $TU_t = U_t T + tU_t$, for each t. This generalization of the notion of time operator can go further. The group $\{U_t\}$ can be replaced by a semigroup, and/or the Hilbert space can be replaced by a Banach space. We shall elaborate on this in next sections. Until the end of this section we shall still consider K-flows and discuss a possibility of extensions of time operators on L^p spaces, $p \geq 1$.

A search for L^p extension of time operators is physically motivated by the fact that operators U_t describe, in particular, the evolution of probability densities, which are integrable functions on the phase space. There is no reason, except technical, to restrict our considerations to square integrable densities. The evolution operators can be, in fact, defined as operators on any L^p-space. The canonical commutation relation (2.7), which actually defines the internal time in dynamical system does not depend on the nature of the space either.

In order to define time operators on other L^p-spaces than L^2 we have to extend the notion of the integral:

$$I_\rho(f) = \int_{-\infty}^{\infty} f(t)\, dE_t\rho$$

which has been defined for square integrable functions ρ. The convenient spectral theory technique, which we have exploited so far, is of little use beyond the Hilbert space L^2. But we can treat $I_\rho(f)$ as an integral with respect to a vector valued measure or as a stochastic integral with respect to the martingale $\{E_t\rho\}$. The extension of the integral presented below is based on both vector measures and stochastic integrals arguments. Using the theory of integral with respect to vector measures we can relatively easy obtain a straightforward generalization. But the stochastic integrals technique is more powerful and allows, in particular, to derive limit theorems needed for applications.

If $\rho \in L^p$ is fixed then the correspondence

$$\mathbf{M} : [s.t) \longmapsto E_s\rho - E_t\rho, \text{ where } \mathrm{s} < \mathrm{t},$$

defines a vector measure on the algebra $\mathcal{B}_0$ generated by intervals $[s.t)$ with values in the Banach space $L^p_{\mathcal{X}}$, $1 \leq p \leq \infty$. This means that $\mathbf{M}$ is an L^p-valued set function on $\mathcal{B}_0$ such that

$$\mathbf{M}(\emptyset) = 0$$

and

$$\mathbf{M}(A \cup B) = \mathbf{M}(A) + \mathbf{M}(B)$$

for each $A, B \in \mathcal{B}_0$, $A \cap B = \emptyset$.

The integral I_ρ can be viewed as an integral of real-Borel measurable functions with respect to the vector measure. However, in order to construct such integral it is first necessary to extend $\mathbf{M}$ to a σ-additive measure on the σ-algebra generated by $\mathcal{B}_0$, i.e. on the σ-algebra $\mathcal{B}_{\mathbb{R}}$ of all Borel subsets of $\mathbb{R}$. The reader is already aware then such extension is possible in the case $p = 2$, because then E_t are spectral projectors. In general, the problem of extension can be formulated in terms of stochastic measures and corresponding stochastic integrals.

Observe that for a fixed $\rho \in L^p_{\mathcal{X}}$, $1 \leq p < \infty$, the family $\{E_t\rho\}$ is a martingale with respect to the family of σ-algebras $\{\Sigma_t\}$ (we assume that the measure μ is normalized). Thus the problem can be formulated as follows: Suppose that $\{m(t)\}$ is a martingale with respect to the filtration $\{\Sigma_t\}$ such that $m(t) \in L^p_{\mathcal{X}},\ p \geq 1$, for each t. Define the stochastic integral

$$\int_{\mathcal{X}} f(t)\, dm(t) \tag{2.8}$$

on real valued functions, such that each bounded Borel measurable function f is integrable and the following Lebesgue dominated convergence theorem holds:

> *If a sequence $\{f_n\}$ of uniformly bounded functions is convergent pointwise to f then the integrals $\int f_n(t)dm(t)$ converge to $\int f(t)\, dm(t)$ in $L^p_{\mathcal{X}}$.*

Indeed, we have

Proposition 2.4 *Let $\{m(t)\}$ be a right continuous martingale in $L^1_{\mathcal{X}}$, which satisfies the property*

$$\int_{\mathcal{X}} \sup_t |m(t)| d\mu < \infty \,.$$

Then the integral in (2.8) *is correctly defined for each bounded function f, and the Lebesgue convergence theorem in $L^1_{\mathcal{X}}$ is satisfied.*

The above proposition is a particular case of a more general theory of stochastic integration with respect to semimartinglales developed by Bichteler (1981). Since the proof requires the technique that is beyond the scope of this book we refer the reader to the above quoted article of Bichteler or to the monograph Kwapień and Woyczyński (1992).

The integrability of $\sup_t |m(t)|$, which is the basic assumption of Proposition 2.4 is satisfied if, for example, $\{m(t)\}$ is L^p bounded martingale, for some $p > 1$. In particular the case of the martingale $\{E_t \rho\}$, the integrability of supremum is guaranteed if

$$\int_{\mathcal{X}} |\rho| \ln |\rho| \, d\mu < \infty \,. \tag{2.9}$$

Recall here that the expression

$$\Omega(\rho) \stackrel{\text{df}}{=} - \int_{\mathcal{X}} \rho(x) \ln \rho(x) \mu(dx) \,,$$

defined for probability densities $\rho \in L^1_{\mathcal{X}}$ is called the *entropy functional*. It is obvious that condition (2.9) is satisfied for ρ bounded. Thus stochastic integral $\int f(t) dE_t \rho$ is correctly defined for any ρ from a dense subset of $L^1_{\mathcal{X}}$. Summarizing the above considerations, we see that:

Corollary 2.2 *Formula* (2.2) *defines time operator T as an operator acting on L^1. The domain of this operator is dense in L^1 since it was dense in L^2 and because the convergence in L^2 implies the convergence in L^1. For an arbitrary bounded Borel measurable function f the operator function $f(T)$ is correctly defined on the set of densities ρ with finite entropy $\Omega(\rho)$.*

3

Time operator for the baker map

In the previous section we have constructed time operators for K-systems. Here we shall concentrate on the simplest of K-systems – the baker transformation. Of course all the general results from the previous section can applied here so the baker map can serve as an explicit illustration. However, thanks to the simple form of this transformation we can refine significantly the results of the previous section. One of such the most important refinements is the explicit form the eigenvectors of the time operator.

Let the phase space $\mathcal{X}$ be the unit square $[0,1] \times [0,1]$. The σ-algebra Σ is the Borel σ-algebra of subsets of $\mathcal{X}$ generated by all possible rectangles of the form $[a,b] \times [c,d]$. The measure μ is the Borel measure defined on the rectangles as the area:

$$\mu([a,b] \times [c,d]) = (b-a)(d-c)\,.$$

we define the transformation S by

$$S(x,y) = \begin{cases} \left(2x\,, \dfrac{y}{2}\right)\,, & 0 \le x < \frac{1}{2}\,,\ 0 \le y \le 1 \\ \left(2x-1\,, \dfrac{y+1}{2}\right)\,, & \frac{1}{2} \le x \le 1\,,\ 0 \le y \le 1\,. \end{cases}$$

The transformation S is compressing in the y direction by $\frac{1}{2}$ and stretching in the x direction by 2. Due to this property S is measure preserving. Moreover it is also invertible and its inverse is given by

$$S^{-1}(x,y) = \begin{cases} \left(\dfrac{x}{2}, 2\right), & 0 \le x \le 1\,,\ 0 \le y < \frac{1}{2} \\ \left(\dfrac{x+1}{2}, 2y-1\right), & 0 \le x \le 1\,,\ \frac{1}{2} \le, y \le 1\,. \end{cases}$$

An convenient way to show that the dynamical system determined by the baker transformation is a K-system is to present as a bilateral shift.

Each point (x, y) of the unit square can be represented as follows

$$x = \sum_{k=0}^{\infty} \frac{x_{-k}}{2^{k+1}}\,,\quad y = \sum_{k=1}^{\infty} \frac{x_k}{2^k}$$

where x_k is either 0 or 1, $k = 0, \pm 1, \pm 2, \ldots$. Therefore the pair (x, y) can be represented by the sequence

$$(x, y) = (\underbrace{\ldots, x_{-2}, x_{-1}, x_0}_{x}, \underbrace{x_1, x_2, \ldots}_{y}) \tag{3.1}$$

of zeros and ones. It is elementary to check that the action of S on the point (x, y) represented as a doubly infinite sequence is the shift to the right

$$\begin{aligned} (x,y) &= (\ldots\ x_{-1}, \quad x_0, \quad\ x_1 \ldots) \\ S(x,y) &= (\ldots\ x_{-2}, \quad x_{-1}, \quad x_0 \ldots) \end{aligned}$$

Consider now the infinite Cartesian product $\mathcal{Y} = \{0,1\}^{\mathbb{Z}}$ that consists of all sequences $\{x_n\}_{n\in\mathbb{Z}}$, where x_k is equal to 0 or 1. Let $\mathcal{A}$ be the σ-algebra of subsets of $\mathcal{Y}$ generated by the cylindrical sets $C^{i_1,\ldots,i_k}_{n_1,\ldots,n_k}$ that consist of all the sequences $\{x_n\}_{n\in\mathbb{Z}}$ such that $x_{n_1} = i_1, \ldots x_{n_k} = i_k$, where $i_1, \ldots, i_k$ is a given finite sequence of zeros and ones.

On the σ-algebra $\mathcal{A}$ we define the product measure ν

$$\nu = \bigotimes_{n=-\infty}^{\infty} \nu_k\,,$$

where ν_k are measures on $\{0,1\}$, such that $\nu_k(\{0\}) = \nu_k(\{1\}) = \frac{1}{2}$. In particular, we have

$$\nu(C^{i_1,\ldots,i_k}_{n_1,\ldots,n_k}) = \frac{1}{2^k}\,.$$

In this way we obtain the measure space $(\mathcal{Y}, \mathcal{A}, \nu)$.

The transformation S on $\mathcal{Y}$ that shifts each sequence one step to the right is measurable and preserves the measure ν. To see this it is enough to note that $S^m(C^{i_1,\ldots,i_k}_{n_1,\ldots,n_k}) = C^{i_1,\ldots,i_k}_{n_1+m,\ldots,n_k+m}$, for each $m \in \mathbb{Z}$. Moreover, S is also a K-automorphism of $(\mathcal{Y}, \mathcal{A})$. Indeed, let us choose as a distinguished σ-algebra $\mathcal{A}_0$ the σ-algebra generated by those cylinders $C^{i_1,\ldots,i_k}_{n_1,\ldots,n_k}$ for which $n_1 < \ldots < n_k = 0$, i.e. the number i_k is located on the 0th coordinate. Then the σ-algebra $S^n\mathcal{A}_0$ is the σ-algebra generated by $C^{i_1,\ldots,i_k}_{n_1+n,\ldots,n_k+n}$ with i_k placed on the nth coordinate. It is now straightforward to check that all three conditions (i)–(iii) of K-system are satisfied, i.e. the dynamical system $(\mathcal{Y},\mathcal{A},\nu;S)$ is a K-system.

The above results can be expressed in terms of the baker transformation, i.e. S as a transformation of the unit square. Note that although the representation (3.1) is not unique, the nonuniqueness concerns only the points of the dyadic division, which is the set of the Lebesgue measure 0. The map that corresponds to each sequence $\{x_n\}\in\mathcal{Y}$ the pair (x,y), as determined by (3.1), is continuous and onto $[0,1]\times[0,1]$ and therefore also measurable. The image through this map of each simple cylinder C^i_n is the set of parallel equally distant strips (let us call them black) that are, for $n\geq 1$, parallel to x-axis and of the width $\frac{1}{2^n}$ and, for $n\leq 0$ parallel to y-axis and of the width $\frac{1}{2^{n+1}}$. If $i=0$ then the first black strip is that adjacent to the x-axis, if $n\geq 1$, or adjacent to the y-axis, if $n\leq 0$. The image of a cylinder $C^{i_1,\ldots,i_k}_{n_1,\ldots,n_k}$ is the intersection of the above described strips. Therefore the family of cylinders corresponds the finite unions of rectangles in $[0,1]\times[0,1]$. The latter family generates the σ-algebra Borel subsets of the unit square. The same correspondence is between the Lebesgue measure on $[0,1]\times[0,1]$ and the measure ν. Finally let us observe that $\pi=\{\Delta_0,\Delta_1\}$, where $\Delta_0=\{(x,y)\in\mathcal{X}\,:\,x<\frac{1}{2}\}$ and $\Delta_1=\{(x,y)\in\mathcal{X}\,:\,x\geq\frac{1}{2}\}$ is a generating partition in the dynamical system $(\mathcal{X},\Sigma,\mu;S)$ (see Section 2.2).

It follows from the previous section that the K-system which arises from the baker transformation admits time operator T of the form

$$T=\sum_{n=-\infty}^{\infty} nP_n$$

where P_n, $n\in\mathbb{Z}$, is the eigenprojection of T corresponding to the eigenvalue n. Time operator T is correctly defined on $\mathcal{H}=L^2_{\mathcal{X}}\ominus[1]$ and its eigenprojections satisfy

$$P_nP_m=\delta_{nm}P_n$$

and

$$\sum_{n=-\infty}^{\infty} P_n=I\,.$$

Our task is to find a complete set of eigenvectors of T.

Consider the above defined partition $\pi=\{\Delta_0,\Delta_1\}$ of the unit square into the left and right half. Define

$$\chi_0=1-2\mathbb{1}_{\Delta_0}$$

and

$$\chi_n=U^n\chi_0\,,\;\; n=\pm1,\pm2,\ldots$$

For any finite ordered set $\mathbf{n}$ of integers

$$\mathbf{n}=\{n_1,\ldots,n_k\}\,,\;\; n_1<\ldots<n_k\,,$$

define the function

$$\chi_{\mathbf{n}} = \chi_{n_1}\chi_{n_2}\cdots\chi_{n_k}\,. \tag{3.2}$$

A complete set of eigenfunctions of T is obtained by taking all possible finite products of χ_n. A product (3.2) is an eigenvector corresponding to the eigenvalue n_k, which is the maximal index among $n_1,\ldots,n_k$. For example $\chi_{-3}\chi_1\chi_0\chi_2$, $\chi_{-1}\chi_1\chi_2$, and χ_2, etc. are all eigenvectors of T corresponding to eigenvalue 2.

In order to prove that $\chi_{\mathbf{n}}$, where $\mathbf{n}$ runs over all ordered sets of integers, together with the constant function 1 form an orthonormal basis in $L^2_{[0,1]\times[0,1]}$ it is sufficient to show that the restrictions of $\chi_{\mathbf{n}}$ to x-axis and y-axis form an orthonormal basis in $L^2_{[0,1]}$. A detailed proof of the later fact will be presented in Section 7 (Lemma 4).

The above constructed eigenvectors of the time operator of the baker map can be equivalently expressed in terms of Walsh function. Namely

The eigenstates of the Baker map are:

$$\{\widetilde{w}_{n_1,\ldots,n_k}\ :\ n_1<\ldots<n_k,\ n_j\in\mathbb{Z}\},$$

where

$$\widetilde{w}_{n_1,\ldots,n_k}(x,y)=\widetilde{r}_{n_1}(x,y)\cdot\ldots\cdot\widetilde{r}_{n_k}(x,y),$$

$$\widetilde{r}_n(x,y)=\begin{cases} r_n(x), \text{ if } \ n=1,2,\ldots,\\ r_{1-n}(y), \text{ if } \ n=0,-1,-2,\ldots.\end{cases}$$

The functions $r_n\in L^2_{[0,1]}$ are the Rademacher functions:

$$r_n(x)=\begin{cases} 1, \text{ if } x_n=0,\\ -1 \text{ if } x_n=1,\end{cases}$$

where x_n is the n-th sign in the binary representation of the number x. The natural projection P is the conditional expectation with respect to the partition of the square into vertical segments, being the k-partition of the Baker transformation:

$$Pf(x,y)=\int_0^1 f(x,t)\,dt.$$

One can easily verify that

$$P\widetilde{w}_{n_1,\ldots,n_k}(x,y)=\begin{cases} w_{n_1,\ldots,n_k}(x),\ \ if\ \ n_1\geq 1,\\ 0,\ \ if\ \ n_1\leq 0\,.\end{cases}$$

4

Time operator of relativistic systems

5

Time operator of exact systems

5.1 EXACT SYSTEMS

In Section 2 we have classified abstract dynamical systems according to different ergodic properties. We have distinguished the following ergodic properties: ergodicity, weak mixing, mixing, exactness and Kolmogorov systems. This classification reflects the degree of irregular behavior, and the property of being Kolmogorov systems or exact system are the strongest in this hierarchy.

Recall that a semigroup $\{S_t\}_{t\geq 0}$ (or the dynamical system) is exact if

$$\lim_{t\to\infty} \mu(S_t(A)) = \mu(\mathcal{X}) \text{ for each } A \in \Sigma\,,\ \mu(A) > 0\,,\ \text{such that } S_t(A) \in \Sigma\,.$$

As it was pointed out in Section 2 exactness implies mixing, so both Kolmogorov systems and exact dynamical systems are subclasses of mixing systems. However these subclasses are disjoint. For a Kolmogorov systems, the set $\{S_t\}$ form a group, while a necessary condition of exactness of a semigroup $\{S_t\}$ is that none of S_t can be invertible. Indeed, if S_{t_0} had the inverse $S_{t_0}^{-1}$, then using the assumption that S_t are measure preserving, we would have

$$\mu(S_{t_0}(A)) = \mu(S_{t_0}^{-1}(S_{t_0}(A))) = \mu(A)\,,$$

for each $A \in \Sigma$, and by induction

$$\mu(S_{nt_0}(A)) = \mu(A)\,, \text{ for } n = 1, 2, \ldots\,.$$

But the last equality contradicts exactness.

Exact transformations often appear when study the behavior of solutions of some differential equations. For example, consider the equation

$$\frac{\partial u}{\partial t} + s\frac{\partial u}{\partial s} = \alpha u\,, \text{ where } (s,t) \in [0,\infty) \times [0,1]\,,$$

with the initial condition

$$u(0,s) = x(s)\,, \quad \text{for } s \in [0,1]\,.$$

Assume that for some r $(r = 0, 1, \ldots)$ the function $x(\cdot)$ is continuous function with all its derivatives up to the order r on the interval $[0,1]$. Then this equation generates the flow $\{S_t\}_{t \geq 0}$ that is determined by its solution $u(s,t)$:

$$S_t x(s) \stackrel{\text{df}}{=} u(t,s) = e^{\alpha t} x(se^{-t})\,.$$

Thus the phase space $\mathcal{X}$ of this flow is

$$\mathcal{X} = \{x \in \mathcal{C}^r[0,1] \,:\, x(0) = x'(0) = \ldots = x^{(r)}(0) = 0\}$$

equipped with the topology of uniform convergence with all derivatives of order $\leq r$. The σ algebra Σ is generated by all cylindrical subsets of $\mathcal{X}$, i.e. the sets of the form

$$\{x \in \mathcal{X} : (x(s_1), \ldots, x(s_n)) \in B\},\ n = 1, 2, \ldots,\ s_1, \ldots, s_n \in [0,1],\ B \in \mathcal{B}_{\mathbb{R}^n}\,.$$

In particular, if $r = 0$, then $\mathcal{X}$ is the space of all continuous function on the interval $[0,1]$ which vanish at the origin. It can be shown (see [LM]) that, , for $\alpha > 0$, the Wiener measure μ_W on $(\mathcal{X}, \Sigma)$, i.e. the measure determined by the Brownian motion, is invariant with respect to $\{S_t\}$. Moreover, the dynamical system $(\mathcal{X}, \Sigma, \mu_W; \{S_t\})$ is exact. In general, if $r > 0$ then it can be shown (see [MyjRud]) that for $\alpha > r$ there exists a Gaussian $\{S_t\}$-invariant measure on $(\mathcal{X}, \Sigma)$ and that the dynamical system is also exact. It is worth to notice that chaotic behavior of the considered flow is only exhibited when $\alpha > r$. For $\alpha \leq r$ all the trajectories $S_t x$, $t \geq 0$, are convergent to 0 in $\mathcal{X}$, as $t \to \infty$. The above type of equations appears in biology, where it describes chaotic behavior in cell populations [Bru]. Incidentally, the family of transformations $\{S_t\}$ corresponding the parameter $\alpha = 1/2$ coincides with the spectral deformation group used in the Aguilar-Balslev-Combes definition of resonance [AgCo,BaCo].

There is a wide class exact dynamical systems associated with maps of intervals. In particular exact are Renyi, the logistic and the cusp map introduced in Section 2.

We have already mentioned that Kolmogorov systems and exact systems are two disjoint subclasses of mixing systems. Kolmogorov systems are reversible exact systems are irreversible. On the other hand there are, however, strict ties between these two systems. We shall show in Section 7 that projecting dynamical system $(\mathcal{X}, \Sigma, \mu; S)$ corresponding to the baker transformation S on the x-coordinate we obtain the exact system corresponding to the 2-adic Renyi map. Conversely it was shown by Rokhlin [Roch] that exact systems can arise as natural projections of Kolmogorov systems and have the same entropy.

In other words, exact dynamical systems can be extended (dilated) to Kolmogorov systems. Let demonstrate briefly this extension assuming for simplicity that time is discrete. Thus instead of a group $\{S_t\}_{t\geq 0}$ let us consider a single measure preserving transformation on the space $(\mathcal{X}, \Sigma, \mu)$ that satisfies the property $\bigcap_{n=1}^{\infty} S^{-n}(\Sigma) =$ the trivial σ-algebra.

Note first that since S is measure preserving it must be "onto" (μ - a.e.). Hence for a given x_0 from $\mathcal{X}$ there is x_{-1} such that $x_0 = Sx_{-1}$. Then there is an x_{-2} such that $x_{-1} = Sx_{-2}$, and so on. Moreover, put $x_n = S^n x_0$, $n = 1, 2, \ldots$, and consider the sequences

$$(\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)\,, \quad \text{for } x_0 \in \mathcal{X}\,. \tag{5.1}$$

If S was invertible, each such sequence could be written simply as

$$(\ldots, S^{-2}x_0, S^{-1}x_0, x_0, Sx_0, S^2x_0, \ldots)$$

and identified with the point x_0. Therefore, in this case the action of S on the point x_0 would be expressed as the left shift in the above sequence. Since S is not invertible, to a given x_0 corresponds a tree of points on the left side of x_0 in the sequence (5.1).

The collection of all such branches, i.e. the set of all sequences (5.1), where $x_0 \in \mathcal{X}$, is treated as a new phase space $\tilde{\mathcal{X}}$ and the transformation $\tilde{S}$ on a sequence $\tilde{x} \in \tilde{\mathcal{X}}$ is defined as the left shift. The σ-algebra $\tilde{\Sigma}$ and measure $\tilde{\mu}$ are then defined in a similar manner as in product spaces (see also Section 12 for a similar construction) and it turns out that $\tilde{S}$ is one-to-one $\tilde{\mu}$ preserving transformation.

Because of the just described relations between exact and Kolmogorov systems it is natural to expect that the time operator associated with a Kolmogorov system should also have a counterpart for exact systems. We shall show that this is indeed the case. Let us assume that $(\Omega, \mathcal{A}\,\mu)$ is a finite measure space and S a single measure preserving transformation of the phase space Ω. As we know there is several equivalent definition of exactness. The most useful for us is the definition of exactness expressed in terms of the Koopman operator V, $Vf(\omega) \equiv f(S\omega)$ and the Frobenius-Perron operator $U \equiv V^\dagger$ regarded as linear operators on $L^2 \equiv L^2(\Omega, \mathcal{A}\,\mu)$. Namely the dynamical system is exact if and only if

$$\lim_{n\to\infty} \|U^n\rho - 1\|_{L^2} = 0\,, \ \ \text{for each density } \rho \in L^2\,, \tag{5.2}$$

or, in terms of the Koopman operator:

$$\bigcap_{n=1}^{\infty} V^n(L^2) = \mathrm{lin}\{1\} \ - \ \text{the linear space spanned by constants}\,. \tag{5.3}$$

For the proof of equivalence of various definition of exactness we refer the reader to the monograph [LM] and Ref. [Lin].In the sequel we assume condition (5.3) as the definition of exactness.

Consider now the Koopman operator V, which is determined by an arbitrary measure preserving transformation S. Observe that V is an isometry on L^2 and consider the intersection

$$\mathcal{H}_\infty = \bigcap_{n=0}^{\infty} V^n(L^2)\,. \tag{5.4}$$

Since $L^2 \supset V(L^2) \supset V^2(L^2) \supset \ldots$, the intersection $\mathcal{H}_\infty$ is V-invariant. The operator V is unitary on the orthogonal complement $\mathcal{H}_0$ of $\mathcal{H}_\infty$,

$$\mathcal{H}_0 \equiv L^2 \ominus \mathcal{H}_\infty\,,$$

and we have

$$\bigcap_{n=0}^{\infty} V^n(\mathcal{H}_0) = \{0\}\,. \tag{5.5}$$

Therefore the space L^2 can be decomposed on a direct sum

$$L^2 = \mathcal{H}_0 \oplus \mathcal{H}_\infty \tag{5.6}$$

in such a way that V is unitary on the second space – $\mathcal{H}_\infty$. The definition (5.3) of exactness means that the operator V is exact if and only if the space $\mathcal{H}_\infty$ consist only of constants. The decomposition (5.6) corresponds to the Wald decomposition from the theory of stationary processes. In the Wald decomposition the analog of the spaces $\mathcal{H}_0$ and $\mathcal{H}_\infty$ are called *purely nondeterministic* and *deterministic* part respectively. Thus exact systems correspond to purely nondeterministic stationary processes.

There is another interpretation of the decomposition (5.6) on the ground of operator theory. Namely, note first that the Koopman operator V has the uniform spectrum on the space $\mathcal{H}_0$. Indeed let $\{g_\alpha\}$ be a orthonormal basis in the space $\mathcal{H}_0 \ominus V(\mathcal{H}_0)$ and denote, for each α, by $\mathcal{H}_\alpha$ the Hilbert space spanned by $\{g_\alpha, Vg_\alpha, V^2g_\alpha, \ldots\}$. Then $(V^kg_\alpha, V^lg_\alpha) = 0$, for each $k, l = 0, 1, 2, \ldots$, $k \neq l$ and, similarly, the spaces $\mathcal{H}_\alpha$ are mutually orthogonal. By (5.5) the system $\{V^ng_\alpha\}$ is complete in $\mathcal{H}_0$, thus we have the decomposition of $\mathcal{H}_0$ on a direct sum of V-invariant subspaces: $\mathcal{H}_0 = \bigoplus_\alpha \mathcal{H}_\alpha$. This shows that V has uniform spectrum on $\mathcal{H}_0$. In particular, if the Koopman operator is exact then it has a uniform spectrum on the orthogonal complement of constants: $L^2 \ominus \{1\}$. Secondly, the definition (5.3) of exactness means in terms of operator theory that V is a unilateral shift on $L^2 \ominus \{1\}$ with the generating space $\mathcal{N}_0 \equiv \mathcal{H}_0 \ominus V(\mathcal{H}_0)$ (see the following section).

The above observations concerning exact systems indicates that the concept of time operator can be generalized for stationary stochastic processes and for shifts. Below we will study the time operator associated with unilateral shifts.

5.2 TIME OPERATOR OF UNILATERAL SHIFT

It was shown in the previous section that the Koopman operator of an exact system is in the case of discrete time a unilateral shift on the space of square integrable functions. The problem of construction of time operator of exact systems can thus be considered in a more general framework. Namely, to construct a time operator associated with a shift operator acting on an arbitrary Hilbert space $\mathcal{H}$. This means (see Section 2) that given a shift operator V on $\mathcal{H}$ we want to construct an operator T that satisfies the following conditions:

$$V(D(T)) \subset D(T) \tag{5.7}$$

and

$$TV_t = V_tT + tV_t\,, \tag{5.8}$$

where $D(T)$ is the domain of T.

We shall show in this section that time operators can be associated with all unilateral shifts. Time operators associated with bilateral shifts on Hilbert spaces will be considered separately. However, before generalizing the construction of time operator we give the notation and some basic facts on shifts.

Definition A linear continuous operator V on a Hilbert space $\mathcal{H}$ is called a *shift* iff there exists a sequence $\{\mathcal{N}_n \,|\, n = 0, 1, 2, \ldots\}$, enumerated by the set of all integers or by the set of all positive integers, of closed linear subspaces of $\mathcal{H}$ such that

(i) $\mathcal{N}_n$ is orthogonal to $\mathcal{N}_m$ if $m \neq n$

(ii) $\mathcal{H} = \oplus_n \mathcal{N}_n$

(iii) V is an isometry from $\mathcal{N}_n$ onto $\mathcal{N}_{n+1}$, for each n, .

The operator V is called *unilateral shift* if $n = 0, 1, 2, \ldots$ and *bilateral shift* if $n \in \mathbb{Z}$. We shall call $\mathcal{N} \equiv \mathcal{N}_0$ the *generating space* of the shift V.

The adjoint $V^\dagger$ of the unilateral shift V vanishes on $\mathcal{N} \equiv \mathcal{N}_0$, moreover $\mathcal{N}_0 = \mathrm{Null} V^\dagger$. The operator $V^\dagger$ maps isometrically $\mathcal{N}_n$ onto $\mathcal{N}_{n-1}$, $n = 1, 2, \ldots$. Therefore the spaces $\mathcal{N}_n$, $n = 1, 2, \ldots$ and the multiplicity m are unique for a given unilateral shift V.

Any unilateral shift V is an isometry, the adjoint shift $V^\dagger$ is a partial isometry and $V^\dagger V = 1$. Any bilateral shift is a unitary operator. The adjoint to a bilateral shift is again a bilateral shift of the same multiplicity.

The spaces $\mathcal{N}_n$ for a bilateral shift V are non unique. Indeed, observe that for any unitary operator B commuting with V the spaces $\mathcal{N}'_n = B(\mathcal{N}_n)$ also satisfy all the conditions of the definition of bilateral shift. The converse is also true: if V is a bilateral shift, $\mathcal{N}_n$ and $\mathcal{N}'_n$ are spaces that satisfy all conditions of the definition of bilateral shift for V then there exists a unitary operator B, commuting with V such that $V(\mathcal{N}_n) = \mathcal{N}'_n$.

Finally, let us introduce the concept of generating basis.

Definition. Let V be a unilateral shift of multiplicity $m \in \mathbb{N} \cup \{\infty\}$ on a Hilbert space $\mathcal{H}$. An orthonormal basis $\{g_a^n \,|\, n = 0, 1, \ldots, 1 \le a < m+1\}$ is called a *generating basis* for V iff $Vg_a^n = g_a^{n+1}$ for all n, a (or equivalently, $V^\dagger g_a^n = g_a^{n-1}$ if $n \geq 1$ and $V^\dagger g_a^0 = 0$).

The generating basis for a bilateral shift can defined in a similar way

Definition. Let V be a bilateral shift of multiplicity $m \in \mathbb{N} \cup \{\infty\}$ on a Hilbert space $\mathcal{H}$. An orthonormal basis $\{g_a^n \,|\, n \in \mathbb{Z},\, 1 \le a < m+1\}$ is called a *generating basis* for V iff $Vg_a^n = g_a^{n+1}$ for all n, a (or equivalently, $V^\dagger g_a^n = g_a^{n-1}$ for all n, a).

The following proposition gives a clear procedure for constructing a generating basis for a given unilateral shift.

Proposition 5.1 *Let V be a unilateral shift on a Hilbert space $\mathcal{H}$ of multiplicity m and $\{g_a \,|\, 1 \leq a < m+1\}$ be an orthonormal basis in $\mathcal{N}_0 = \ker V^\dagger$. Let also $g_a^n = V^n g_a$ for any $k = 0, 1, 2, \ldots$ and any a. Then $\{g_a^n \,|\, n = 0, 1, 2, \ldots, 1 \leq a < m+1\}$ is a generating basis for V.*

Proof. The condition $V g_a^n = g_a^{n+1}$ is obvious. It remains to prove that $\{g_a^n \,|\, n = 0, 1, 2, \ldots, 1 \leq a < m+1\}$ is an orthonormal basis in $\mathcal{H}$. First, from the definition of the shift, the operators V^n are isometries from $\mathcal{N}_0$ onto $\mathcal{N}_n$. Therefore, the set $\{g_a^n \,|\, 1 \leq a < m+1\}$ is an orthonormal basis in $\mathcal{N}_n$. By orthogonality of different $\mathcal{N}_n$ we get orthonormality of the set $\{g_a^n \,|\, n = 0, 1, 2, \ldots, 1 \leq a < m+1\}$. The completeness of this orthonormal system follows from the density of the sum of $\mathcal{N}_n$.

The following theorem gives the general form of time operators for unilateral shifts.

Theorem 5.1 *Let V be a unilateral shift on a Hilbert space $\mathcal{H}$ Then for any self-adjoint time operator T for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$ there exists a unique self-adjoint operator A on $\mathcal{N}_0$ such that the operator T has the form:*

$$T = \sum_{n=0}^{\infty} V^n (A + n)(V^\dagger)^n K_n, \tag{5.9}$$

where $K_n = V^{n+1}(V^\dagger)^{n+1} - V^n (V^\dagger)^n$ are the orthoprojections onto the spaces $\mathcal{N}_n$ from the definition of unilateral shift.

Conversely any operator of the form (5.9), *where A is a self-adjoint operator on $\mathcal{N}_0$, is a time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$.*

Proof. Let T be a self-adjoint time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$.

First, we shall prove that T has the form (5.9) when the spectrum $\sigma(T)$ is a subset of the set $\mathbb{Z}/l = \{m/l \,|\, m \in \mathbb{Z}\}$ for some positive integer l. In this case by the spectral theorem $\mathcal{H}$ is the orthogonal direct sum of the spaces $\mathcal{H}_\tau$ ($\tau \in \mathbb{Z}/l$), where $\mathcal{H}_\tau = \mathrm{Null}(T - \tau)$. Because V preserves the domain of T we have $V(\mathcal{H}_\tau) \subseteq \mathcal{H}_{\tau+1}$ for all $\tau \in \mathbb{Z}/l$. Therefore for any $\tau \in [0, 1) \cap \mathbb{Z}/l$, $V^n(\mathcal{H}_{\tau-n})$ is the decreasing sequence of closed linear subspaces of $\mathcal{H}_\tau$ (here $n = 0, 1, 2, \ldots$) with zero intersection. Choosing orthonormal bases in each space $V^n(\mathcal{H}_{\tau-n}) \ominus V^{n+1}(\mathcal{H}_{\tau-n-1})$ and taking the union of these bases we obtain an orthonormal basis $\mathcal{B}_\tau$ in $\mathcal{H}_\tau$ such that for any $n = 0, 1, 2, \ldots$, $\mathcal{B}_\tau \cap V^n(\mathcal{H}_{\tau-n})$ is an orthonormal basis in $V^n(\mathcal{H}_{\tau-n})$. Then for any $\tau \in \mathbb{Z}/l$, $\tau < 0$

$$\mathcal{B}_\tau = (V^\dagger)^{[\tau]}(\mathcal{B}_{\{\tau\}} \cap V^{[\tau]}(\mathcal{H}_\tau))$$

is an orthonormal basis in $\mathcal{H}_\tau$ (here $[\tau]$ and $\{\tau\}$ are the integer and fractional parts of the number τ respectively). For $\tau \in \mathbb{Z}/l$, with $\tau > 1$ we shall construct orthonormal basis $\mathcal{B}_\tau$ in $\mathcal{H}_\tau$ by induction. The first step of the induction ($\mathcal{B}_\tau$ for $\tau < 1$) is already established. Assume that $j \in \mathbb{Z}$, $j \geq 2$ and the bases $\mathcal{B}_\tau$ for $\tau \in \mathbb{Z}/l$,

$\tau < j-1$ are already constructed. We shall construct the basis B_τ for $\tau \in \mathbb{Z}/l$, $j-1 \le \tau < j$. Choose an arbitrary orthonormal basis $\mathcal{B}'_\tau$ in $\mathcal{H}_\tau \ominus V(\mathcal{H}_{\tau-1})$. The set $\mathcal{B}_\tau = \mathcal{B}'_\tau \cap V(\mathcal{B}_{\tau-1})$ is obviously an orthonormal basis in $\mathcal{H}_\tau$. Let

$$\mathcal{B} = \bigcup_{\tau \in \mathbb{Z}/l} \mathcal{B}_\tau .$$

Then $\mathcal{B}$ is an orthonormal basis in $\mathcal{H}$ and $Vf \in \mathcal{B}$ for any $f \in \mathcal{B}$ (this follows from the construction of $\mathcal{B}$). This means that $\mathcal{B}$ is a generating basis for V. Therefore, $\mathcal{A}_0 = \{f \in \mathcal{B} \,|\, V^\dagger f = 0\}$ is an orthonormal basis in $\mathcal{N}_0$. Consequently $\mathcal{A}_n = V^n(\mathcal{A}_0) \subset \mathcal{B}$ is an orthonormal basis in $\mathcal{N}_n$ for any $n = 0, 1, 2, \ldots$. From the construction of $\mathcal{B}_\tau$ for any $f \in \mathcal{B}$ there exists the unique $\tau(f) \in \mathbb{Z}/l$ that $f \in \mathcal{B}_{\tau(f)}$. Let us define an operator A_n on $\mathcal{N}_n$ by the formula

$$A_n \left(\sum_{f \in \mathcal{A}_n} a_f f \right) = \sum_{f \in \mathcal{A}_n} \tau(f) a_f f .$$

This operator is self-adjoint (because the vectors of the orthonormal basis $\mathcal{A}_n$ are eigenvectors of A_n with real eigenvalues). Since for any $f \in \mathcal{B}$, $f \in \mathcal{B}_{\tau(f)} = \mathrm{Null}(T - \tau(f))$ we have $Tf = \tau(f)f$. Therefore A_n coincides with the restriction of T to $\mathcal{N}_n$. Thus T commutes with projections K_n and

$$T = \sum_{n=0}^{\infty} A_n K_n .$$

To prove (5.9) for $A = A_0$ it remains to verify the equality $A_n = V^n(A_0 + n)(V^*)^n$. From the construction of bases $\mathcal{A}_n$ we have that $\mathcal{A}_n = V^n(\mathcal{A}_0)$. Let $\mathcal{A}_0 = \{g_k \,|\, 1 \le k < m+1\}$ (where m is multiplicity of V). Then $\mathcal{A}_n = \{V^n g_k \,|\, 1 \le k < m+1\}$. So

$$\begin{aligned} A_n \left(\sum_{1 \le k < m+1} c_k V^n g_k \right) &= \sum_{1 \le k < m+1} c_k \tau(V^n g_k) V^n g_k \\ &= \sum_{1 \le k < m+1} c_k (\tau(g_k) + n) V^n g_k = \\ &= V^n \sum_{1 \le k < m+1} c_k (\tau(g_k) + n) g_k = \\ &= V^n (A + n)(V^\dagger)^n V^n \sum_{1 \le k < m+1} c_k g_k = \\ &= V^n (A + n)(V^\dagger)^n \left(\sum_{1 \le k < m+1} c_k V^n g_k \right) . \end{aligned}$$

So formula (5.9) for T (with $A = A_0$) is proved.

In the general case let $P(\tau)$ be the spectral resolution of T:

$$T = \int_{-\infty}^{\infty} \tau \, dP(\tau)$$

It is straightforward to check that the conditions (5.7) and (5.8) for T (with respect to the semigroup V^n) are equivalent to the fact that for any $\tau_1 < \tau_2$, $V(P_{\tau_2} - P_{\tau_1})(\mathcal{H}) \subseteq (P_{\tau_2+1} - P_{\tau_1+1})(\mathcal{H})$. From this reformulation of the definition of time operator (in our particular case) it immediately follows that for any Borel function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x+1) = f(x) + 1$ for all $x \in \mathbb{R}$, the operator $f(T)$ is also a self-adjoint time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$. The function $f_l(x) = [lx]/l$ ($l \in \mathbb{N}$) satisfies the condition $f_l(x+1) = f_l(x) + 1$ and takes values in $\mathbb{Z}/l$. Hence, $f_l(T)$ is a self-adjoint time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$, with spectrum included in $\mathbb{Z}/l$. From the previous result there exist self-adjoint operators A_l on $\mathcal{N}_0$ such that

$$f_l(T) = \sum_{n=0}^{\infty} V^n (A_l + n)(V^\dagger)^n K_n \, . \tag{5.10}$$

From the definition of f_l it follows that the operators $T - f_l(T)$ are bounded and $\|T - f_l(T)\| \le 1/l$. So, the sequence $f_l(T)$ converges to T with respect to the operator norm. From the other side $A_l = f_l(T)\big|_{\mathcal{N}_0}$. Therefore the operators $A_i - A_j$ are bounded and $\|A_i - A_j\| \le \frac{1}{\min(i,j)}$. So A_i is the Cauchy sequence with respect to the operator norm, and therefore A_l converges to some self-adjoint operator A on $\mathcal{N}_0$. Passing to the limit in the formula (5.10) we obtain the desired formula (5.9).

Formula (5.9) implies that $A = T\big|_{\mathcal{N}_0}$. Therefore A is unique for given T.

Conversely, let T be an operator given by the formula (5.9). Since A is a self-adjoint operator on $\mathcal{N}_0$, each operator $A_n = V^n(A+n)(V^\dagger)^n K_n\big|_{\mathcal{N}_n}$, $n = 0, 1, 2, \ldots$ is self-adjoint on $\mathcal{N}_n$, as it is unitary equivalent to the self-adjoint operator $A + n$. Therefore T is self-adjoint as a direct sum of self-adjoint operators. It remains to verify conditions (5.7) and (5.8) with respect to the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$. Let $f \in D_T \subset \mathcal{H}$. Then by (5.9)

$$TVf = \sum_{n=0}^{\infty} A_n K_n V f = \sum_{n=0}^{\infty} V^n (A+n)(V^\dagger)^n K_n V f \tag{5.11}$$

From the obvious relations $K_0 V = 0$ and $K_n V = V K_{n-1}$ for $n \ge 1$ we obtain

$$\sum_{n=0}^{\infty} V^n (A+n)(V^\dagger)^n K_n V f = \sum_{n=1}^{\infty} V^n (A+n)(V^\dagger)^n V K_{n-1} f =$$

$$= \sum_{n=1}^{\infty} V V^{n-1} (A+n-1)(V^\dagger)^{n-1} K_{n-1} f + \sum_{n=1}^{\infty} V V^{n-1} (V^\dagger)^{n-1} K_{n-1} f =$$

$$= \sum_{n=0}^{\infty} VA_nK_nf + \sum_{n=1}^{\infty} VK_nf\,. \tag{5.12}$$

Since $f \in D_T$, the series $\sum_{n=0}^{\infty} A_nK_nf$ converges. Since V is continuous and the sequence K_n of orthoprojections is a decomposition of the identity both series $\sum_{n=0}^{\infty} VA_nK_nf$ and $\sum_{n=1}^{\infty} VK_nf$ converge in $\mathcal{H}$. Thus, by (5.11) and (5.12) $Vf \in D_T$ and

$$TVf = \sum_{n=0}^{\infty} VA_nK_nf + \sum_{n=1}^{\infty} VK_nf = VTf + Vf.$$

Corollary 5.1 *For a unilateral shift V on a Hilbert space $\mathcal{H}$, there exists a unique self-adjoint time operator T for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$, satisfying the additional condition that the null space of the adjoint shift is included into the null space of T:*

$$V^{\dagger}f = 0 \text{ implies that } Tf = 0\,. \tag{5.13}$$

In this case the null space $\mathcal{N}$ of the adjoint shift $V^{\dagger}$ is identical with the null space of T.

Proof. By Theorem 5.1, formula (5.9) with $A = 0$ defines a time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$. It is straightforward to verify that this time operator satisfies condition (5.13). Now suppose that T is time operator for the semigroup $\{V^n \,|\, n = 0, 1, 2, \ldots\}$, satisfying the condition (5.13). From theorem 5.1, T can be represented in the form (5.9) with some self-adjoint operator A on $\mathcal{N}_0$. Since the restriction of T to $\mathrm{Null}V^*$ coincides with A, condition (5.13) implies that $A = 0$. Therefore T is unique. The inclusion $\mathrm{Null}V^{\dagger} \subseteq \mathrm{Null}T$ is one of our assumptions. The opposite inclusion follows from (5.9) with $A = 0$, since formula (5.9) in the case $A = 0$ provides spectral resolution of T.

The above theorem allows, in addition, to determine the diagonal form of the time operator T associated with unilateral shift. It turns out that T has the diagonal representation in terms of the generating bases for the shift V.

Corollary 5.2 *The time operator T of Corollary 5.1 has diagonal representation for any generating basis g_a^n:*

$$T = \sum_{n=0}^{\infty} n \sum_{1 \le a < m+1} |g_a^n\rangle\langle g_a^n| \tag{5.14}$$

$$K_n = \sum_{1 \le a < m+1} |g_a^n\rangle\langle g_a^n|\,.$$

Therefore the generating basis is a basis of eigenvectors of T: $Tg_a^n = ng_a^n$.

Proof of the formula (5.14) follows from (5.13) and (5.9) immediately.

5.3 TIME OPERATOR OF BILATERAL SHIFT

We shall prove now an analog of Theorem 5.1 for bilateral shifts. This theorem generalizes the construction of time operator from Section 2 because Koopman and Frobenius-Perron operators of invertible K-systems are bilateral shift operators on the Hilbert space of square integrable functions.

The generating basis of multiplicity M for a bilateral shift V is defined in the same way. Only difference is that $n \in \mathbb{Z}$ and the condition $V^\dagger g_a^0 = 0$, if $n \geq 1$, is dropped.

Theorem 5.2 *Let V be a bilateral shift on a Hilbert space $\mathcal{H}$ and T be a self-adjoint time operator for the group $\{V^n \,|\, n \in \mathbb{Z}\}$. Then the time operator T is given by*

$$T = \sum_{n=-\infty}^{\infty} V^n (A + nI)(V^\dagger)^n K_n \,, \tag{5.15}$$

where

1) *K_n are the orthoprojections onto the subspaces $\mathcal{N}_n$, $n \in \mathbb{Z}$;*
2) *the subspaces $\mathcal{N}_n$ satisfy the conditions (1–3) of the definition of shifts;*
3) *A is a self-adjoint operator on $\mathcal{N}_0$ such that the spectrum $\sigma(A)$ is contained in $[0,1]$ and 1 is not an eigenvalue of A.*

The subspaces $\mathcal{H}_n$ and the operator A are uniquely determined by T.

Conversely, if the subspaces $\mathcal{N}_n$, $n \in \mathbb{Z}$ satisfy all conditions from the definition of bilateral shift, K_n are the orthoprojections onto $\mathcal{N}_n$, and A is an arbitrary self-adjoint operator on $\mathcal{N}_0$ then any operator of the form (5.15) *is a time operator for the group $\{V^n : n \in \mathbb{Z}\}$.*

Proof. Let T be a self-adjoint time operator for the group $\{V^n : n \in \mathbb{Z}\}$, and $\{E_\lambda\}$ be the spectral resolution of T:

$$T = \int_{-\infty}^{\infty} \lambda \, dE_\lambda \,. \tag{5.16}$$

Define the function $\varphi : \mathbb{R} \to \mathbb{R}$ by the formula

$$\varphi(t) = \begin{cases} 1, & \text{if } t \in [0,1), \\ 0, & \text{if } t \in \mathbb{R} \setminus [0,1). \end{cases}$$

Let $K_n = \varphi_n(T)$, where $\varphi_n(t) = \varphi(t-n)$ or equivalently

$$K_n = \lim_{\substack{\lambda \to n \\ \lambda < n}} E_{\lambda+1} - E_\lambda .$$

It is clear that K_n are orthoprojections and that the spaces $\mathcal{N}_n = K_n(\mathcal{H})$ satisfy all the properties from the definition of bilateral shift for V and $T(\mathcal{N}_n) \subseteq \mathcal{N}_n$.

Let $A_n : \mathcal{N}_n \to \mathcal{N}_n$ be the restriction of the operator T to the space $\mathcal{N}_n$. The operators A_n are self-adjoint as restrictions of the self-adjoint operator T. Evidently,

$$T = \sum_{n=-\infty}^{\infty} A_n K_n.$$

To prove (5.15) for $A = A_0$ it remains to verify the equality $A_n = V^n(A_0+n)(V^*)^n$ for any $n \in \mathbb{Z}$. This can be done by the same argument as the corresponding statement in the proof of the theorem 5.1. Since $A = TK_0\big|_{\mathcal{N}_0}$ and $K_0 = \varphi(T)$ we have that $\sigma(A)$ is contained in the closure of the set $\varphi(\mathbb{R})$ and the point spectrum of A is contained in the set $\varphi(\mathbb{R})$. Therefore $\sigma(A) \subset [0,1]$ and 1 does not belong to the point spectrum.

Let us verify now the uniqueness of $\mathcal{N}_n$ and A. Let T be the operator defined by the formula (5.15). Since $\sigma(A+n) \subseteq [n, n+1]$ and $n+1$ is not an eigenvalue of $A+n$, we have that $\varphi_n(T) = K_n$ for all $n \in \mathbb{Z}$. So, K_n and therefore $\mathcal{N}_n$ are unique. By (5.15) A is the restriction of T to $\mathcal{N}_0$. Therefore, A is also unique.

The proof of the converse part of Theorem 5.2 is the same as for Theorem 5.1.

Remark. For a given self-adjoint Time Operator T of a bilateral shift V there exist representations of T in the form (5.15) with different K_n. The condition that makes K_n unique is the restriction that the spectrum of A is contained in [0,1] and 1 is not an eigenvalue of A. This condition is actually equivalent to the identification of K_n of formula (5.15) with the eigenprojections $P_{n+1} \ominus P_n$ obtained from the spectral resolution (5.16).

NOTES

1. The null space $\mathcal{N} = \mathcal{N}_0 = \ker V^\dagger$ of the adjoint unilateral shift is the generator of innovations $\mathcal{N}_n = V^n(\mathcal{N})$ which coincide with the eigenspaces of the time operator. The space $\mathcal{N}$ contains all initial information and defines the origin of time.

2. As the bilateral shift is the unitary dilation [SzNFo] of the unilateral shift, $\mathcal{N}$ is still a generator of innovations. However as the null space of the adjoint bilateral shift is the zero subspace, the generating subspace $\mathcal{N}$ does not contain the initial $\lambda = 0$ information. The initial information is included in the space

$$\mathcal{H}^{\text{in}} = E_0(\mathcal{H}) = \bigoplus_{n=-\infty}^{0} V^n(\mathcal{N}).$$

This space is called incoming subspace by Lax and Phillips [LaxPh] as it contains the initial information which will manifest in future. In the case of Kolmogorov systems [CFS] $P(0)$ is the conditional expectation projection onto the K-partition [Mi,Pr].

3. If V is a unilateral shift, then the adjoint semigroup $\{(V^*)^n\}$ does not have any time operator. Therefore the Time Operator does not exist for the evolution of probability densities of exact endomorphisms, where V is the Koopman operator and $U = V^\dagger$ is the Perron–Frobenius operator.

This non-equivalence between states and observables does not appear in automorphisms, like Kolmogorov systems, where the Koopman and Perron–Frobenius operators are both unitary. In the original definition of time operators [Mi,Pr] both pictures in terms of states (Schrödinger) and observables (Heisenberg) were used without discrimination. However the generalization of time operators beyond unitary evolutions is only possible in observables (Heisenberg) evolution picture.

4. The definition of time operator goes beyond the shift evolution. In fact time operators can be defined for more general contracting semigroups [AShdiff].

6

Time operator of the Renyi map and the Haar wavelets

A characteristic feature of the time operator of exact systems constructed in the previous section is that it has uniform infinite multiplicity. This means that the eigenspaces of time operator have the same infinite dimension. We shall show in this section that it is also possible a natural construction of a time operator for an exact systems based on its generating partition. If this partition is finite then the eigenspaces of the time operator are also finite. Time operators of the first kind are called the *time operators with uniform multiplicity,* the second kind the *time operators with non-uniform multiplicity*.

In this section we shall construct a time operator with non-uniform multiplicity for the exact system determined by the Renyi map. One of the reasons for choosing this particular dynamical system is that this will allow to demonstrate in a transparent way the ideas, the details of the construction as well as to compare uniform and non-uniform time operator. This will also clarify the details of the construction of time operator of the baker map. Another reason is that there are close connections between the non-uniform time operator of the Renyi map and the multiresolution analysis of $L^2_{\mathbb{R}}$ based on Haar wavelets. We shall show later (Section 13) that this connection is not incidental. The spectral decomposition of a time operator turns out to be a generalization of multiresolution analysis.

We begin with the construction of the complete family of eigenfunctions of the time operator. The construction is based on a simple but important remark that the Koopman operator acts as a shift on the Haar functions. Then we shall discuss the domain of the time operator and show the connections between eigenfunctions of the time operator and the Haar wavelets on the interval $[0, 1]$. We shall also show relations of the time operator of the Renyi map with the time operator of the baker's transformation constructed previously by Misra, Prigogine and Courbage. Finally we shall construct, for a comparison, a uniform time operator of the Renyi map applying the results of Section 6.

6.1 NON-UNIFORM TIME OPERATOR OF THE RENYI MAP

The 2-adic Renyi map is defined on the unit interval $[0, 1)$ by the formula

$$Sx = 2x \ (\mathrm{mod}\, 1) \, .$$

The Lebesgue measure is invariant with respect to S. The Renyi map is the simplest chaotic system and the prototype of exact endomorphisms presented in the previous section.

The Koopman operator of the Renyi map is the unilateral shift defined by

$$Vf(x) = f(Sx) = \begin{cases} f(2x) \, , & \text{for } x \in [0, \frac{1}{2}) \\ f(2x-1) \, , & \text{for } x \in [\frac{1}{2}, 1) \, , \end{cases} \tag{6.1}$$

The construction of the time operator of the Renyi map is based on the following observation:

Lemma 6.1 *The Koopman operator* (6.1) *acts as a shift on the Haar functions in* $L^2_{[0,1]}$

$$h_{n,k}(x) \stackrel{\mathrm{df}}{=} \mathbb{1}_{[0,1)}(2^{n+1}x - 2k) - \mathbb{1}_{[0,1)}(2^{n+1}x - 2k - 1) \, , \tag{6.2}$$

where $n = 0, 1, 2, ..., \, k = 0, 1, ..., 2^n - 1$*,*

$$V \, h_{n,k} = h_{n+1,k} + h_{n+1,k+2^n} \, ,$$

for $n = 0, 1, 2, ...$ *and* $k = 0, 1, ..., 2^n - 1$.

Proof.

$$Vh_{n,k}(x) = \begin{cases} h_{n,k}(2x), & \text{for } x \in [0, \frac{1}{2}) \\ h_{n,k}(2x-1), & \text{for } x \in [\frac{1}{2}, 1) \end{cases}$$

$$= \begin{cases} \mathbb{1}_{[0,1)}(2^{n+2}x - 2k) - \mathbb{1}_{[0,1)}(2^{n+2}x - 2k - 1), & x \in [0, \frac{1}{2}) \\ \mathbb{1}_{[0,1)}(2^{n+2}x - 2(k + 2^n)) - \mathbb{1}_{[0,1)}(2^{n+2}x - 2(k+2^n) - 1), & x \in [\frac{1}{2}, 1) \end{cases}$$

$$= \begin{cases} h_{n+1,k}(x), & \text{for } x \in [0, \frac{1}{2}), \ k = 0, 1, ..., 2^n - 1 \\ h_{n+1,k}(x), & \text{for } x \in [\frac{1}{2}, 1), \ k = 2^n, ..., 2^{n+1} - 1 \end{cases}$$

$$= h_{n+1,k}(x) + h_{n+1,k+2^n}(x)$$

The last equality follows from the fact that function $h_{n+1,k}$ and $h_{n+1,k+2^n}$ have disjoint supports $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$ respectively.

In order to obtain an orthonormal basis we define the functions

$$\begin{aligned} \varphi_1(x) &= h_{0,0}(x) = \mathbb{1}_{[0,1)}(2x) - \mathbb{1}_{[0,1)}(2x-1) \\ \varphi_{n+1}(x) &= V^n \varphi_1(x)\,, \quad \text{for } n = 1, 2, ... \end{aligned}$$

We observe that

$$\varphi_2(x) = V h_{0,0}(x) = h_{1,0}(x) + h_{1,1}(x)\,.$$

In general, it follows from Lemma 6.1 that

$$\varphi_{n+1} = h_{n,0} + ... + h_{n,2^n-1}\,, \tag{6.3}$$

for $n = 1, 2, \ldots$.

We see therefore that every function $\varphi_{n+1}(x)$ can be expressed in terms of the Haar functions as the sum of all k-components of $h_{n,k},\ k = 0, 1, ..., 2^n - 1$

Now, for a given set of integers $\mathbf{n}$

$$\mathbf{n} = \{n_1, ..., n_m\}\,, \quad n_1 < ... < n_m$$

define the function

$$\varphi_{\mathbf{n}}(x) \stackrel{\text{df}}{=} \varphi_{n_1}(x) ... \varphi_{n_m}(x)\,. \tag{6.4}$$

Lemma 6.2 *Any function $\varphi_{\mathbf{n}}$, where $\mathbf{n} = \{n_1, ..., n_m\}\,,\ \ n_1 < ... < n_m$, can be expressed as*

$$\varepsilon_0 h_{n_m,0} + ... + \varepsilon_{2^{n_m}-1} h_{n_m, 2^{n_m}-1}, \tag{6.5}$$

for some choice of $\varepsilon_i = -1$ or 1, $i = 0, \ldots, 2^{n_m} - 1$.

Proof. Indeed, since n_m corresponds to the finest division of the interval and the functions φ_i can only assume values +1 or -1, the multiplication of φ_{n_m} by φ_{n_i}, for $n_i < n_m$, can only change signs of some $h_{n_m,j}$. This ends the proof.

Lemma 6.3 *The Koopman operator act as a shift on $\varphi_{\mathbf{n}}$*

$$V \varphi_{\mathbf{n}} = \varphi_{\mathbf{n}+1}\,,$$

where $\mathbf{n} + 1 \stackrel{\text{df}}{=} \{n_1 + 1, ..., \ n_m + 1\}$.

Proof. Since the Koopman operator is multiplicative

$$(Vfg)(x) = f(Sx)g(Sx) = (Vf)(x)(Vg)(x)\,,$$

we have

$$V\varphi_{\mathbf{n}} = V(\varphi_{n_1} ... \varphi_{n_m}) = V\varphi_{n_1} ... V\varphi_{n_m} = \varphi_{n_1+1} ... \varphi_{n_m+1} = \varphi_{\mathbf{n}+1}\,.$$

Lemma 6.4 *The functions $\varphi_{\mathbf{n}}$, where $\mathbf{n}$ runs over all ordered subsets of $\mathbb{N}$, together with the constant $\equiv 1$, form an orthonormal basis in $L^2_{[0,1]}$.*

Proof. Note first that $\varphi_n^2 \equiv 1$, for each n. Therefore $\varphi_{\mathbf{n}}^2 \equiv 1$ and

$$\varphi_{\mathbf{n}'}\varphi_{\mathbf{n}''} = \varphi_{n'_1}...\varphi_{n'_{m_1}}\varphi_{n''_1}...\varphi_{n''_{m_2}} = \varphi_{n_1}...\varphi_{n_m} ,$$

where $n_1, ...n_m$ is the rearrangement of the numbers $n', ..., n'_{m_1}, n''_1, ..., n''_{m_2}$ in the increasing order (if $n'_i = n''_j$ then $\varphi_{n'_i}\varphi_{n''_j} \equiv 1$, thus we can eliminate these numbers). This implies that $\varphi_{\mathbf{n}'}\varphi_{\mathbf{n}''}$ is again of the form (6.5) and consequently $\int_0^1 \varphi_{\mathbf{n}'}(x)\varphi_{\mathbf{n}"}(x)dx = 0$.

We shall show next that all indicators of sets $[\frac{k}{2^n}\,,\,\frac{k+1}{2^n})$ can be obtained as finite linear combinations of products $\varphi_{n_1}...\varphi_{n_m}$, with $n_m \leq n$ and the constant function 1. This will imply, in particular, that any Haar function is such a linear combination and that the completeness in $L^2_{[0,1]}$ of the family $\{\varphi_{\mathbf{n}}\}$ will be a consequence of the completeness of the Haar basis. Indeed applying the mathematical induction we see that for $n = 1$

$$\mathbb{1}_{[0,\frac{1}{2})} = \frac{1}{2}\cdot 1 + \frac{1}{2}\varphi_1 \quad \text{and} \quad \mathbb{1}_{[\frac{1}{2},1)} = \frac{1}{2}\cdot 1 - \frac{1}{2}\varphi_1.$$

Let us assume now that the statement is true for n and consider an interval $\Delta^{(n+1)} = [\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}})$. It is the left or the right part of some interval $\Delta^{(n)} = [\frac{\beta}{2^n}, \frac{\beta+1}{2^n})$ indicator of which is by the assumption a linear combination of products $\varphi_{n_1}...\varphi_{n_m},\ n_m \leq n$. Therefore the indicator

$$\mathbb{1}_{\Delta^{(n+1)}} = \frac{1}{2}\mathbb{1}_{\Delta^{(n)}} \pm \frac{1}{2}\mathbb{1}_{\Delta^{(n)}}\varphi_{n+1} ,$$

where the choice of + or $-$ depends whether $\Delta^{(n+1)}$ is the left or the right part of $\Delta^{(n)}$, is again a linear combination of products $\varphi_{n_1}...\varphi_{n_m}$, with $n_m \leq n+1$.

From Lemma 6.3 and 6.4 we can construct the innovation spaces $\mathcal{W}_n, n = 1, 2, \ldots$ for the Koopman operator V.

Let us define

$$\mathcal{W}_0 = \mathrm{span}\{1\} \text{ and } \mathcal{W}_n = \mathrm{span}\{\varphi_{n_1}...\varphi_{n_m}\},$$

for all choices of nonnegative integers $n_1 < ... < n_m = n,\ n = 1, 2, \ldots$. Let us also denote by $\mathcal{B}_n$ the σ-algebra which corresponds to the n-th dyadic division and by $L^2(\mathcal{B}_n)$ the subspace of $L^2_{[0,1]}$ consisting of all $\mathcal{B}_n$-measurable functions. By construction we have the standard properties of innovation spaces for shifts, namely:

1) the spaces $\mathcal{W}_n$ are mutually orthogonal and

$$\mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \ldots \oplus \mathcal{W}_n = L^2(\mathcal{B}_n)\,.$$

2) the space $L^2_{[0,1]}$ has the following direct sum decomposition

$$L^2_{[0,1]} = \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \ldots\,.$$

3) the Hilbert space $L^2_{[0,1]} \ominus \{1\} \equiv \mathcal{H}$ of non-equilibrium deviation is:

$$\mathcal{H} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \dots$$

4) the Koopman operator shifts the innovations

$$V\mathcal{W}_n \subset \mathcal{W}_{n+1}\,, \quad n = 1, 2, \dots . \tag{6.6}$$

We should remark here that in the unitary case we have

$$V\mathcal{W}_n = \mathcal{W}_{n+1}\,.$$

Putting P_n as the projection on $\mathcal{W}_n$, $n = 1, 2, \dots$, we obtain the time operator

$$T = \sum_{n=1}^{\infty} nP_n \tag{6.7}$$

The union $\bigcup_{n=1}^{\infty} \mathcal{W}_n$, which is dense in $\mathcal{H}$, is in the domain of T and that for $\mathbf{n} = \{n_1, \dots, n_m\}$

$$T\varphi_{\mathbf{n}} = n_m \varphi_{\mathbf{n}}\,.$$

Therefore each $\varphi_{\mathbf{n}}$ is an eigenvector of T corresponding to the eigenvalue $n = \max\{m : m \in \mathbf{n}\}$ which is called the *age* of $\varphi_{\mathbf{n}}$. The age n corresponds to the partition of [0,1) on 2^n intervals and can be identified with the information carried by this uniform partition. The space $\mathcal{W}_0 \oplus \mathcal{W}_1 \oplus ... \oplus \mathcal{W}_n$ consists of all elements having the age $\leq n$ or, using the computational language, all functions (with the mean value 0) which can be computed with the accuracy $\frac{1}{2^n}$. Then $\mathcal{W}_{n+1}$ represents the necessary contribution which is needed if one wants to describe all functions which can be computed with the accuracy $\frac{1}{2^{n+1}}$. We shall show below that T is indeed the Koopman operator with respect to V.

Theorem 6.1 *Each vector $f \in \mathcal{H}$ has the following expansion in the basis $\{\varphi_{\mathbf{n}}\}$*

$$f = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a_{n,k} \varphi_{\mathbf{n}_n^k}\,, \tag{6.8}$$

where $\mathbf{n}_n^k$ denotes the set $\{n_1, ..., n_i\}$ with fixed $n_i = n$ while k runs through all 2^{n-1} possible choices of integers $n_1 < ... < n_{i-1} < n$. The operator T acts on f having finite expansion (6.8) *as follows*

$$Tf = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} n a_{n,k} \varphi_{\mathbf{n}_n^k}\,. \tag{6.9}$$

Moreover T satisfies

$$TV^m = V^m T + mV^m\,, \quad m = 1, 2, \dots \tag{6.10}$$

Proof. The expansion (6.8) is a direct consequence of Lemma 6.4 and equality (6.9) from the definition of P_n and T.
The intertwining relation (6.10) follows follows from the equality

$$P_{n+m}V^m = V^m P_n, \tag{6.11}$$

where we put $P_n = 0$ for $n \leq 0$. Indeed, for any function f of the form (6.8)

$$V^m f = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a_{n,k} \varphi_{\mathbf{n}_n^k + m} \,.$$

Therefore $V^m f \in \mathcal{W}_{m+1} \oplus \mathcal{W}_{m+2} \oplus \ldots$ and consequently

$$P_n V^m f = 0 = V^m P_{n-m} f \,,$$

for $n \leq m$. If $n > m$ then

$$P_n V^m f = \sum_{k=1}^{2^{n-m-1}} a_{n-m,k} \varphi_{\mathbf{n}_n^k} \,.$$

On the other hand

$$V^m P_{n-m} f = V^m \left(\sum_{k=1}^{2^{n-m-1}} a_{n-m,k} \varphi_{\mathbf{n}_n^k - m} \right) = \sum_{k=1}^{2^{n-m-1}} a_{n-m,k} \varphi_{\mathbf{n}_n^k}$$

which proves (6.11), for all $n, m \in \mathbb{N}$.

6.2 THE DOMAIN OF THE TIME OPERATOR

What functions belong to the domain of our time operator? An immediate consequence of the definition of T is that it is densely defined on L^2 because all finite linear combinations of Haar functions are in the domain. On the other hand it is easy to find a function $f \in L^2$ of the form (6.8) for which

$$\|f\|_{L^2} < \infty \,,$$

while

$$\|Tf\|_{L^2} = \infty \,.$$

We shall show below that the domain of the time operator contains also "smooth" functions. For this purpose we shall consider expansions of functions in the Haar basis instead of of the expansion in the basis $\{\varphi_{\mathbf{n}_m^k}\}$ which is in fact the Walsh basis.

Although the Walsh basis allows to describe in a simple way the dynamics on L^2, i.e. the action of the Koopman operator, it not so convenient when dealing with smooth functions. It is known, for example, that the Haar expansion of a continuous function on the interval $[0,1]$ converges uniformly, while its Walsh series may be even pointwise divergent.

From now on till the end of this section we shall only consider those functions from L^2 which have continuous extensions on $[0,1]$. The following theorem gives a necessary condition on a function f to belong to the domain of the time operator.

Theorem 6.2 *Any function $f \in \mathcal{C}_{[0,1]}$ such that its modulus of continuity satisfies the property*

$$\int_0^1 \omega_f(x) \frac{\log x}{x} dx > -\infty$$

belongs to the domain of T.

The modulus of continuity ω_f [Zy] is defined by:

$$\omega_f(\delta) = \sup_{\substack{x,y\in[0,1] \\ |x-y|\le\delta}} |f(x) - f(y)| \,, \quad 0 \le \delta \le 1 \,,$$

for any $f \in \mathcal{C}_{[0,1]}$.

In order to prove this theorem we need the following lemma

Lemma 6.5 *The Haar function $h_{n,k}$* (6.2) *is an eigenfunction of T corresponding to the same eigenvalue n.*

Proof. Consider for a fixed m the functions $\varphi_{\mathbf{n}_m^k}$, $k = 0, 1, \ldots, 2^m - 1$. By Lemma 6.2

$$\varphi_{\mathbf{n}_m^k} = \varepsilon_0 h_{m,0} + \ldots + \varepsilon_{2^m-1} h_{m,2^m-1} \,, \tag{6.12}$$

where the choice of ε_j depends on k. A simple geometric argument shows that each $h_{m,j}$ can be represented as a linear combination of $\varphi_{\mathbf{n}_m^k}$, $k = 0, 1, \ldots, 2^m - 1$. Indeed, we can identify $h_{m,j}$ with the unit vector of the 2^m dimensional Euclidean space $(0, \ldots, 0, 1, 0, \ldots, 0)$, where 1 is on the $j+1$-coordinate. Since $(\varphi_{\mathbf{n}_m^k} | \varphi_{\mathbf{n}_m^{k'}})_{L^2} = \sum_j \varepsilon_j \varepsilon_{j'}$, the L^2-scalar product can also be identified with the scalar product in the Euclidean space. From (6.12) follows that there are 2^m mutually orthogonal, thus linearly independent vectors $\varphi_{\mathbf{n}_m^k}$ in $\mathbb{R}^{2^m}$. Therefore each $h_{m,j}$ is a linear combination of $\varphi_{\mathbf{n}_m^k}$. Since $\varphi_{\mathbf{n}_m^k}$ is an eigenvector of T corresponding to the eigenvalue m, also $\varphi_{\mathbf{n}_m^k}$ is en eigenvector of T with the same eigenvalue.

Proof of Theorem 6.2. Without a loss of generality we may assume that $\int_0^1 f(x)\,dx = 0$. Then f has the expansion in the Haar basis

$$f = \lim_{N\to\infty} f_N = \lim_{N\to\infty} \sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} c_{n,j} \psi_{n,j} \,,$$

with $\psi_{n,j} = 2^{n/2} h_{n,j}$ and $c_{n,j} = \int_0^1 f(x) \psi_{n,j}(x) dx$. Since f is continuous the above series is not only L^2-convergent but also uniformly on $[0,1]$. We shall show that also Tf_N converges uniformly on $[0,1]$, as $N \to \infty$. Thus we can define f as the limit $\lim_{n\to\infty} Tf_N$. By Lemma 6.5

$$Tf_N = \sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} nc_{n,j}\psi_{n,j} \,.$$

Therefore applying the inequality [Ci]

$$|c_{n,j}| \leq \frac{1}{22^{n/2}} \omega_f \left(\frac{1}{2^{n+1}} \right) ,$$

valid for all $k = 0, 1, \ldots, 2^{n-1}$, we have

$$\begin{aligned} \left| \sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} nc_{n,j}\psi_{n,j}(x) \right| &\leq \sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} |nc_{n,j}\psi_{n,j}(x)| \\ &\leq \sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} 22^{n/2} \sup_{j=0,\ldots,2^{n-1}} |c_{n,j}| \\ &\leq \sum_{n=1}^{N} n\omega \left(\frac{1}{2^{n+1}} \right) \,. \end{aligned}$$

However $\sum_{n=1}^{N} n\omega\left(\frac{1}{2^{n+1}}\right)$ is nothing but the Riemann sum of the integral $-\int_0^1 \omega_f(x) \frac{\log x}{x} dx$ which is finite by the assumption. therefore there is a constant K such that

$$\sum_{n=1}^{N} \sum_{j=0}^{2^{n-1}} |nc_{n,j}\psi_{n,j}(x)| \leq K < \infty \,,$$

for each $N = 1, 2, \ldots$ which shows the uniform convergence of the functional series $\sum_{n=1}^{\infty} \sum_{j=0}^{2^{n-1}} nc_{n,j}\psi_{n,j}(x)$.

6.3 THE HAAR WAVELETS ON THE INTERVAL

We show that the above constructed spectral decomposition of the time operator coincides with the multiresolution analysis of $L^2_{[0,1]}$ [Ch,LMR].

A *multiresolution analysis* (MRA) of $L^2_{\mathbb{R}}$ is a sequence $\{\mathcal{H}_n\}_{n\in\mathbb{Z}}$ of closed subspaces of $L^2_{\mathbb{R}}$ such that

$$\{0\} \subset \ldots \subset \mathcal{H}_{-1} \subset \mathcal{H}_0 \subset \mathcal{H}_1 \subset \ldots \subset L^2_{\mathbb{R}} \,,$$

$$\bigcap_{n\in\mathbb{Z}} \mathcal{H}_n = \{0\} \,,$$

$$\overline{\bigcup_{n\in\mathbb{Z}} \mathcal{H}_n} = L^2_{(\mathbb{R})}\,,$$

$$f(\cdot) \in \mathcal{H}_n \Longleftrightarrow f(2^n\,\cdot) \in \mathcal{H}_0\,.$$

Moreover it is assumed that there is a function $\phi \in L^2_{(\mathbb{R})}$ (the *scaling function*) whose integer translates form a Riesz basis of $\mathcal{H}_0$, i.e. the set $\{\phi(\,\cdot - k)\ : k \in \mathbb{Z}\}$ is dense in $L^2_{(\mathbb{R})}$ and there exist positive constants A and B such that

$$A\sum_{k\in\mathbb{Z}} c_k^2 \le \|\sum_{k\in\mathbb{Z}} c_k\phi(\,\cdot - k)\|^2_{L^2_{(\mathbb{R})}} \le B\sum_{k\in\mathbb{Z}} c_k^2\,,$$

for each $\{c_k\}_{k\in\mathbb{Z}} \subset \mathbb{R}$ such that $\sum_{k\in\mathbb{Z}} c_k^2 < \infty$.

It follows from the definition of MRA that the space $\mathcal{H}_n$ is spanned by the functions

$$\phi_{n,\alpha}(x) \stackrel{\text{df}}{=} 2^{n/2}\phi(2^n x - \alpha)\,,\ \ \alpha \in \mathbb{Z}\,.$$

Let us denote by $\mathcal{W}_n$ the orthogonal complement of $\mathcal{H}_n$ in $\mathcal{H}_{n+1}$, i.e.

$$\mathcal{H}_{n+1} = \mathcal{W}_n \oplus \mathcal{H}_n\,,\ \ n \in \mathbb{Z}\,.$$

It can be shown [Ch,LMR,Da] that there exists a function h such the translated versions of

$$h_{n,\alpha}(x) \stackrel{\text{df}}{=} 2^{n/2}h(2^n x - \alpha)\,,\ \ \alpha \in \mathbb{Z}\,, \tag{6.13}$$

generate an orthonormal basis of $\mathcal{W}_n$. The function h is called the *wavelet*.

As we have seen above there is a strict analogy between wavelets and shift operators. This analogy leads to a definition of a time operator for wavelets [AnGu]. Here we shall compare the eigenfunctions of the time operator (6.12) of the Renyi map with the wavelets on the interval $[0,1)$.

The Haar wavelet

$$\psi(x) = \begin{cases} 1\,, & \text{for } 0 \le x < 1/2 \\ -1\,, & \text{for } 1/2 \le x < 1 \\ 0\,, & \text{otherwise} \end{cases}$$

generates the orthonormal basis $\psi_{n,\alpha}$ of $L^2_{(\mathbb{R})}$ with formula (6.13). The restriction of the Haar wavelet to the unit interval $[0,1)$ can be achieved [LMR,Da] by *periodizing* the underlying wavelet. Namely, for a given wavelet h we define

$$b\psi^{\text{per}}_{n,\alpha}(x) = \sum_{\beta\in\mathbb{Z}} h_{n,\alpha}(x+\beta)\,.$$

We shall denote by $\mathcal{H}_n^{\text{per}}$ and $\mathcal{W}_n^{\text{per}}$ the periodic counterparts of the spaces $\mathcal{H}_n$ and $\mathcal{W}_n$ defined above. In this way we can show the the analogy mentioned previously is an exact fact namely

Theorem 6.3 *The wavelet spaces $\mathcal{W}_n^{\text{per}}$ of the Haar wavelet on the interval $[0,1)$ coincide with the eigenspaces $\mathcal{W}_n$ of the time operator T of the Renyi map.*

Proof. Since the functions $\psi^{\rm per}_{n,\alpha}$ are periodic with the period 1, they can be regarded as functions on the interval $[0,1)$. It is enough to focus our attention on the wavelet ψ which is, as it was assumed above, the Haar wavelet. In such case the corresponding functions $h_{n,\alpha}$ have, for fixed n and different α, different supports. Moreover, if $n<0$ then the functions $h_{n,\alpha}$ are constant on the interval $[0,1)$. This implies that also the periodic functions $\psi^{\rm per}_{n,\alpha}$ are constant on $[0,1)$, for $n\geq 0$ and that only $\psi^{\rm per}_{n,\alpha}$ with $0\leq\alpha\leq 2^n-1$ may be different. The latter fact can be easily checked directly. Indeed, each integer α can be written as $\alpha=l2^n+\alpha'$, where $l\in\mathbb{Z}$ and $0\leq\alpha'<2^n$. Then

$$\begin{aligned}\psi^{\rm per}_{n,\alpha}(x) &= \textstyle\sum_{i\in\mathbb{Z}} h_{n,\alpha}(x+i)\\ &= \textstyle\sum_{\beta\in\mathbb{Z}} 2^{n/2}h(2^n(x+\beta)-\alpha)\\ &= \textstyle\sum_{\beta\in\mathbb{Z}} 2^{n/2}h(2^n(x+\beta-l)-\alpha')\\ &= \textstyle\sum_{\beta\in\mathbb{Z}} h_{n,\alpha'}(x+\beta-l)\\ &= \psi^{\rm per}_{n,\alpha'}(x)\,.\end{aligned}$$

Finally observe that

$$\psi^{\rm per}_{n,\alpha}|_{[0,1)} = h_{n,\alpha}\,.$$

Indeed, recall that

$$\psi^{\rm per}_{n,\alpha}(x) = \sum_{i\in\mathbb{Z}} h_{n,\alpha}(x+i) = \sum_{i\in\mathbb{Z}} 2^{n/2}h(2^n(x+i)-\alpha)\,.$$

However, h, being the Haar function, is zero outside $[0,1)$. Therefore, if $n\geq 0$ and $0\leq\alpha\leq 2^{n-1}$, the function $h(2^{n/2}(x+i)+\alpha)$ may assume non-zero values only if $i=0$. This implies that above sum is reduced to one component $h_{n,\alpha}(x)$. Consequently the space $\mathcal{W}^{\rm per}_n = \overline{{\rm span}\{\psi^{\rm per}_{n,\alpha} : 0\leq\alpha\leq 2^n-1\}}$ may be identified with the space $\mathcal{W}_n = \overline{{\rm span}\{h_{n,\alpha} : 0\leq\alpha\leq 2^n-1\}}$, $n\geq 0$. By Lemma 6.2 the latter space coincides with $\mathcal{W}_n$.

6.4 RELATIONS BETWEEN THE TIME OPERATORS OF THE RENYI AND BAKER MAPS

The Renyi map results as a canonical projection of the baker's transformation on the unit square $[0,1)\times[0,1)$

$$B(x,y) \overset{\rm df}{=} \begin{cases} \left(2x, \dfrac{y}{2}\right)\,, & \text{for } 0\leq x\leq\frac{1}{2} \\ \left(2x-1, \dfrac{y}{2}+\dfrac{1}{2}\right)\,, & \text{for } \frac{1}{2}\leq x\leq 1\,. \end{cases}$$

Therefore the Koopman operator V of the Renyi map is the projection of the Koopman operator V_B of the baker's transformation. Equivalently V_B is a positive dilation of V. We ask therefore whether the time operator T of the Renyi map is also a projection of the time operator of the baker's transformation. This is indeed the case.

The Koopman operator of the baker transformation is

$$V_B f(x,y) = f(B(x,y))\,.$$

V_B denotes the unitary group of $\{V_B^n\}_{n=0,\pm1,\pm2,\ldots}$. From the function

$$h(x,y) \stackrel{\rm df}{=} \left[\mathbb{1}_{[0,\frac{1}{2})}(x) - \mathbb{1}_{[\frac{1}{2},1)}(x)\right]\mathbb{1}_{[0,1)}(y) \tag{6.14}$$

we construct the family of functions $\{h_{\mathbf{n}}\}$

$$\{h_{\mathbf{n}}\} \stackrel{\rm df}{=} V_B{}^{n_1} h V_B{}^{n_2} h \ldots V_B{}^{n_m} h\,, \tag{6.15}$$

where $\mathbf{n} = \{n_1, n_2, \ldots n_m\}$ ($n_1 < n_2 < \ldots < n_m$) runs through all finite subsets of all integers. Observe that for $n \geq 0$, $V_B{}^n h$ is of the form $h_1(x)\mathbb{1}_{[0,1)}(y)$ while for $n < 0$ it is of the form $\mathbb{1}_{[0,1)}(x)h_2(y)$. Therefore each function $h_{\mathbf{n}}$ can be represented as a product of two functions depending on x and y respectively. Moreover the family $\{h_{\mathbf{n}}\}$ is an orthonormal basis in $L^2_{[0,1]\times[0,1]}$. Indeed, the orthonormality can be proved precisely in the same way as in Lemma 6.4, while completeness follows from the fact that both 'x and y parts' are complete in L^2 (Lemma 6.4). Define the spaces $\mathcal{W}_n^B$ in a similar way as $\mathcal{W}_n$, i.e.

$$\mathcal{W}_n^B = \mathrm{lin}\{V_B{}^{n_1} h V_B{}^{n_2} h \ldots V_B{}^{n_m} h\}\,,$$

for all choices of integers $n_1 < \ldots < n_m = n$, $n = 0, \pm1, \pm2, \ldots$. The space $\mathcal{H}^B \stackrel{\rm df}{=} L^2_{[0,1]\times[0,1]} \ominus \{1\}$

$$\mathcal{H}^B = \bigoplus_{n=-\infty}^{\infty} \mathcal{W}_n^B\,.$$

The time operator T_B associated with the group $\{V_B{}^n\}$ is defined as

$$T_B = \sum_{n=-\infty}^{\infty} n P_n\,,$$

where P_n is the orthogonal projection from $\mathcal{H}^B$ onto $\mathcal{W}_n^B$. It is straightforward to check that T_B is a time operator with the group $\{V_B^n\}$, similarly as we checked (6.10) for T and V.

Let us now define the projection P from $L^2_{[0,1]\times[0,1]}$ onto the x-coordinate – $L^2_{[0,1]}$ by

$$Pf(x)g(y) \stackrel{\rm df}{=} f(x)\int_0^1 g(y)dy$$

and extending by linearity on all functions of the form

$$\sum_{i=1}^{N} a_i f_i(x) g_i(y)\,. \tag{6.16}$$

Since the operator T is bounded and functions of the form (6.16) are dense in $L^2_{[0,1]\times[0,1]}$ P extends on the whole space $L^2_{[0,1]\times[0,1]}$. Now note that $Ph_{\mathbf{n}} = 0$, for all $\mathbf{n}$ which contain a number $n < 0$. This implies that

$$P(\mathcal{W}^B) = P\left(\bigoplus_{n=-\infty}^{\infty} \mathcal{H}_n^B\right) = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \ldots \mathcal{W}$$

and that

$$PT^B = TP\,.$$

6.5 THE UNIFORM TIME OPERATOR FOR THE RENYI MAP

The Koopman operator $V : L^2_{[0,1]} \to L^2_{[0,1]}$ of an exact endomorphism restricted to the orthocomplement to constants

$$\mathcal{H} = L^2_{[0,1]} \ominus \mathbf{1}$$

is a unilateral shift. Therefore we can also construct the time operator T of $V\big|_{\mathcal{H}}$ applying the general construction of the time operator of the exact endomorphism.

Note first, that the adjoint V^* of V is the Frobenius-Perron operator:

$$V^*f(x) = Uf(x) = \frac{1}{2}\left[f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)\right]. \tag{6.17}$$

It is straightforward to verify that $\operatorname{Null} V^*$ consists of all functions f of the form

$$f(x) = \begin{cases} \phi(x), \text{if } x \in (0,1/2) \\ -\phi(x-1/2), \text{if } x \in (1/2,1) \end{cases}$$

where $\phi \in L^2_{[0,1/2]}$. This implies that for a given set of integers $\mathbf{n} = \{n_1, ..., n_m\}$, $n_1 < ... < n_m$, the function $\varphi_{\mathbf{n}}(x) = \varphi_{n_1}(x)...\varphi_{n_m}(x)$ is in $\operatorname{Null} V^*$ if $n_1 = 1$. Since each function $\varphi_{\mathbf{n}}$ with $n_1 > 1$ is orthogonal to $\operatorname{Null} V^*$, the set

$$\{\varphi_{\mathbf{n}} = \varphi_1\varphi_{n_1}...\varphi_{n_m} \;:\; 1 < n_1 < n_2 < \ldots < n_m\}$$

is an orthonormal basis in $\mathcal{N}_0 = \operatorname{Null} V^*$. By Proposition 6.1 the set of functions

$$\{\varphi_{\mathbf{n}} = \varphi_n\varphi_{n_1}...\varphi_{n_m} = V^{n-1}\varphi_1\varphi_{n_1-n+1}...\varphi_{n_m-n+1} : 1 < n_1 < n_2 < \ldots < n_m\}$$

is a generating basis of V. This basis coincides with the well known Walsh basis.

From Theorem 6.1 we have that the operator T with the following diagonal representation in the basis $\{\varphi_{\mathbf{n}}\}$:

$$T\varphi_{n_1}...\varphi_{n_m} = n_1\varphi_{n_1}...\varphi_{n_m}$$

is a (uniform) time operator for the semigroup $\{V^n : n = 0, 1, 2, \ldots\}$. Indeed, let $\mathcal{H}_n$, $n = 0, 1, \ldots$ be the subspace of $L^2_{[0,1]}$ consisting of the functions constant on the intervals $[k2^{-n}, (k+1)2^{-n})$, $k = 0, 1, \ldots 2^n - 1$ and P_n be orthoprojections on $\mathcal{H}_n \ominus \mathcal{H}_{n-1} \subset \mathcal{H}$, $n = 1, 2, \ldots$. It was shown that the operator

$$T = \sum_{k=0}^{\infty} nP_n \tag{6.18}$$

is a self-adjoint time operator for the semigroup $\{V^n : n = 0, 1, 2, \ldots\}$. By Theorem (6.1) it can be represented by the formula (6.9) through a self-adjoint operator A on the space $\mathcal{N}_0 = \ker V^\dagger$, where A is simply the restriction of T to $\mathcal{N}_0$. It is straightforward to verify that this restriction A acts on Walsh functions from $\mathcal{N}_0$ as follows

$$A\varphi_{1,n_1,\ldots,n_k} = n_k \varphi_{1,n_1,\ldots,n_k}.$$

Note that the dimensions of the age eigensubspaces $\ker(T - nI)$ of the time operator (6.18) are finite and increase with n.

NOTES

1) The age eigenfunctions $\varphi_{\mathbf{n}}$ allow for probabilistic prediction because any square integrable function f can be expanded as

$$f = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a_{n,k} \varphi_{\mathbf{n}_n^k}.$$

The evolution of f can be easily computed as a shift in the coordinates $f_{\tau,\alpha(\tau)}$

$$\begin{aligned} V^m f &= \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a_{n,k} V^m \varphi_{\mathbf{n}_n^k} \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} a_{n,k} \varphi_{\mathbf{n}_n^k + m}. \end{aligned}$$

2) The representation of evolution in term of the age eigenstates is the generalized Fourier transform of the spectral representation which has also been constructed in the context of irreversibility [AMSS].

3) The time operator (6.9) is the analogue of the canonical time operator of Misra, Prigogine and Courbage (see Section 15). However there is significant difference in the multiplicity of the age eigenspaces. The eigenvalue n corresponds to the eigenvectors $\varphi_{\{n_1,\ldots,n_m\}}$ with $n_\alpha = n$. The dimension of the space $\mathcal{W}_n$ is therefore equal to the number of different sets which can be composed using numbers $1, \ldots, n-1$, which is 2^{n-1}. Therefore the time operator T has a nonuniform multiplicity in contradistinction with the canonical time operator of Misra, Prigogine and Courbage which has uniform multiplicity equal to the multiplicity of the Koopman operator which is countably infinite. We shall show elsewhere that it is possible to define time operators of exact systems with uniform multiplicity.

4) Since the multiplicity of the age eigenspaces increases we have (6.6)

$$V\mathcal{W}_n \subset \mathcal{W}_{n+1}$$

in contradistinction to the uniform case where

$$V\mathcal{W}_n = \mathcal{W}_{n+1}$$

We distinguish as a result two kinds of functions in each innovation space $\mathcal{W}_n$, namely: the elements $\varphi_{\mathbf{n}} \in \mathcal{W}_n$ for which there is an "ancestor" $\varphi_{\mathbf{n}'} \in \mathcal{W}_{n-1}$ such that $\varphi_{\mathbf{n}} = V\varphi_{\mathbf{n}'}$, and those which do not arise from $\mathcal{W}_{n-1}$. These function can be characterized with respect to their symmetry with respect to $\frac{1}{2}$. Namely the functions with "ancestors" in $\mathcal{W}_n$ are symmetric

$$\varphi_{\mathbf{n}}(x+\frac{1}{2}) = \varphi_{\mathbf{n}}(x) \pmod 1 \quad \Leftrightarrow \quad 1 \notin \mathbf{n} \tag{6.19}$$

while the functions without "ancestors" in $\mathcal{W}_n$ are antisymmetric

$$\varphi_{\mathbf{n}}(x+\frac{1}{2}) = -\varphi_{\mathbf{n}}(x) \pmod 1 \quad \Leftrightarrow \quad 1 \in \mathbf{n}\,. \tag{6.20}$$

Indeed the function $\varphi_1(x)$, which is by the definition equal to 1, for $0 < x < \frac{1}{2}$, and -1, for $\frac{1}{2} < x < 1$, satisfies (6.19). On the other hand for any function h we have

$$(Vh)(x+\frac{1}{2}) = h(2x+1) \underset{(\mathrm{mod}\ 1)}{=} h(2x) = (Vh)(x)\,.$$

This implies (6.19), since $\varphi_{\mathbf{n}} = V\varphi_{\mathbf{n-1}}$ provided $1 \notin \mathbf{n}$. Condition (6.20) follows from the fact any function $\varphi_{\mathbf{n}}$ with $1 \in \mathbf{n}$ can be written in the form

$$\varphi_{\mathbf{n}}(x) = \varphi_1(x)(V\varphi_{\mathbf{n}'})(x)\,,$$

for some $\mathbf{n}'$.

Each space $\mathcal{W}_n$, and consequently the space $\mathcal{H} = \bigoplus_{n=1}^{\infty} \mathcal{W}_n$, can be decomposed on orthogonal classes of functions – the symmetric and antisymmetric functions with respect to $\frac{1}{2}$. The symmetric functions are simply in the space $V\mathcal{H}$ while the antisymmetric functions are in the space $\mathcal{H} \ominus V\mathcal{H}$.

5) The eigenfunction (6.12) of the time operator are everywhere identical with Walsh functions except at the jump points. Recall that the Walsh functions are defined as all possible products of Rademacher functions

$$r_n(x) \overset{\mathrm{df}}{=} \mathrm{sign}(\sin 2^n \pi x)\,.$$

6) From the previous remark we see that the Koopman operator of the Renyi map generates the Rademacher functions from the first Haar function

$$h(x) = \mathbb{1}_{[0,1)}(2x) - \mathbb{1}_{[0,1)}(2x-1)$$

like the "creation" operator in the case of the Harmonic oscillator in quantum mechanics

7

Time operator of the cusp map

Time operator of the cusp map

The purpose of this section is to construct a Time Operator for the so-called cusp map

$$S : [-1, 1] \to [-1, 1], \quad \text{where} \quad S(x) = 1 - 2\sqrt{|x|}\,, \tag{7.1}$$

which is an approximation of the Poincaré section of the Lorenz attractor [Ott]. The absolutely continuous invariant measure of the cusp map has the density:

$$\rho(x) = \frac{1-x}{2}\,. \tag{7.2}$$

Therefore the measure space $(\mathcal{X}, \Sigma, \mu)$ of the dynamical system determined by the cusp map consists of the interval $[-1, 1]$, the Borel σ-algebra and the measure μ with the density function (7.2).

The Koopman operator of the cusp map is the linear operator V acting on $L^2 = L^2(\mathcal{X}, \Sigma, \mu)$ of the form:

$$Vf(x) = f(1 - 2\sqrt{|x|})\,. \tag{7.3}$$

The dynamical system determined by the cusp map is an exact system. However the proof of exactness is based on other technique than used so far and will not be demonstrated here. We refer the interest reader to [AnShYa].

We shall construct a time operator of the cusp map constructing a generating basis of V and then using the general results of Section 6. For this we have to find an orthonormal basis of the space $\mathcal{N} = \ker V^*$ generating the innovations. We start with the following convenient characterization of the space $\mathcal{N}$:

Lemma 7.1 *The space $\mathcal{N}$ coincides with the following space of functions:*

$$\{f \in L^2 : f(x) = (1+x)g(x), \text{ where } g \text{ is odd}\}\,. \tag{7.4}$$

Proof. One can easily verify that the Frobenius–Perron operator $U = V^*$ is

$$\begin{aligned} Uf(x) = \;& \left(\frac{1}{2} - \frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(\left(\frac{1-x}{2}\right)^2\right) \\ &+ \left(\frac{1}{2} + \frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(-\left(\frac{1-x}{2}\right)^2\right). \end{aligned}$$

Let $f \in \mathcal{N}$ and assume, changing f if necessary on a set of zero Lebesgue measure, that $f(1) = f(-1) = 0$ and. Then

$$\left(\frac{1}{2}-\frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(\left(\frac{1-x}{2}\right)^2\right) + \left(\frac{1}{2}+\frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(-\left(\frac{1-x}{2}\right)^2\right) = 0\,,$$

for all $x \in [-1, 1]$.

Denoting $y = (\frac{1-x}{2})^2$ we obtain the equation:

$$(1-y)f(y) + (1+y)f(-y) = 0\,, \text{ for all } y \in [0, 1]\,. \tag{7.5}$$

Let $g(x) = \frac{f(x)}{1+x}$ for $x \in (-1, 1]$ and $g(-1) = 0$. Then $f(x) = g(x)(1 + x)$ and equation (7.5) can be rewritten as

$$(1-y^2)(g(y) + g(-y)) = 0 \text{ for all } y \in [0, 1]\,. \tag{7.6}$$

Equation (7.6) together with the equalities $g(1) = g(-1) = 0$, imply that g is an odd function. Hence, f is an element of space (7.4).

Conversely if f has the form $f(x) = (1 + x)g(x)$, where g is odd, then we have

$$\begin{aligned} Uf(x) &= \left(\frac{1}{2} - \frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(\left(\frac{1-x}{2}\right)^2\right) \\ &\quad + \left(\frac{1}{2} + \frac{1}{2}\left(\frac{1-x}{2}\right)^2\right) f\left(-\left(\frac{1-x}{2}\right)^2\right) \\ &= \frac{1}{2}\left(1 - \left(\frac{1-x}{2}\right)^4\right)\left(g\left(\left(\frac{1-x}{2}\right)^2\right) + g\left(-\left(\frac{1-x}{2}\right)^2\right)\right) \\ &= 0 \end{aligned}$$

since g is odd. Hence $f \in \mathcal{N}$.

Lemma 7.2 *Let $g_n : [-1, 1] \to \mathbb{C}$, $n = 1, 2, \ldots$, be a sequence of odd measurable functions. Then the sequence $f_n(x) = (1+x)g_n(x)$ is an orthonormal basis in $\mathcal{N}$, if and only if the sequence $e_n(x) = \sqrt{1-x^2}g_n(x)$, $x \in [0, 1]$ is an orthonormal basis in $L^2_{[0,1]}$.*

Proof. Taking into account Lemma 7.1 we obtain

$$
\begin{aligned}
\langle f_n | f_m \rangle_{L^2([-1,1],\mu)} &= \int_{-1}^{1} \frac{1-x}{2}(1+x)^2 g_n(x) g_m(x)\, dx \\
&= \frac{1}{2}\int_{-1}^{1} (1-x^2) g_n(x) g_m(x)\, dx \\
&\quad + \frac{1}{2}\int_{-1}^{1} (1-x^2) x g_n(x) g_m(x)\, dx \\
&= \int_{0}^{1} (1-x^2) g_n(x) g_m(x)\, dx \\
&= \int_{0}^{1} e_n(x) e_m(x)\, dx \\
&= (e_n | e_m)_{L^2_{[0,1]}} ,
\end{aligned}
$$

where we have used the fact that function $(1-x^2)g_n(x)g_m(x)$ is even and function $(1-x^2)xg_n(x)g_m(x)$ is odd. This proves that orthonormality and completeness of $\{f_n\}$ in $\mathcal{N}$ is equivalent to orthonormality and completeness of $\{e_n\}$ in $L^2_{[0,1]}$.

Corollary 7.1 *The set of functions*

$$
\chi^k(x) = \sqrt{\frac{2(1+x)}{1-x}} \sin \pi k x, \quad k = 1, 2, \dots , \tag{7.7}
$$

is an orthonormal basis in the space $\mathcal{N} = \ker V^*$.

Proof. It is well known that the sequence $e_k(x) = \sqrt{2} \sin \pi k x$, $k = 1, 2, \dots$ is an orthonormal basis in $L^2_{[0,1]}$. It remains to apply Lemma 7.2.

The proof of the following theorem follows immediately from Corollaries 6.1, 6.2 and 7.1.

Theorem 7.1 *The set* $\chi^k_n = V^n \chi^k = \chi^k(S^n(x))$, $k = 1, 2, \dots, n = 0, 1, \dots$ *is a generating basis for the Koopman operator* V *of the cusp map acting on* $L_2 \ominus \mathbf{1}$. *The operator with eigenvectors* χ^k_n *and eigenvalues* n *is the time operator of the shift* V*:*

$$
T\chi^k_n = n\chi^k_n;
$$

$$
T = \sum_{n=0}^{\infty} n \sum_{k=1}^{\infty} |\chi^k_n\rangle\langle\chi^k_n| \,. \tag{7.8}
$$

The form of the age eigenfunctions (7.7) of the cusp map is shown in Fig. 1.

Remark:

In the construction of the generating basis (7.7) for V, we used the basis $e_n(x) = \sqrt{2}\sin \pi n x$ in $L^2_{[0,1]}$. Different orthonormal bases in $L^2_{[0,1]}$ lead to different generating bases for V (but the operator T defined by formula (7.8) is of course the same).

8

Time operator of stationary stochastic processes

We begin this section with a brief reminder of basic concepts and notions from the theory of stochastic processes. More details the reader will find in Appendix 1.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(\mathcal{X}, \Sigma)$ a measurable space and I an index set. An $\mathcal{X}$-valued stochastic process on $(\Omega, \mathcal{F}, P)$ is a family $\{X_t\}_{t\in I}$ of $(\mathcal{F}, \Sigma)$ measurable functions

$$X_t : \Omega \to \mathcal{X}, \;\; t \in I \, .$$

The index set I is assumed to be a totally ordered subset of integers or real numbers. In this section we shall only consider real or complex valued stochastic processes. Intuitively speaking a stochastic process is a family of random variables $\{X_t\}$ defined on a common probability space, where each random variable $X_t = X_t(\cdot), \; t \in I$, represents the outcome of an experiment performed at the time instant t. For each $\omega \in \Omega$ the correspondence

$$t \longmapsto X_t(\omega)$$

is a function defined on I called a *realization* of the process $\{X_t\}$.

A *finite dimensional distribution* of the $\mathcal{X}$-valued stochastic process $\{X_t\}_{t\in I}$ is defined for a given set $\{t_1, \ldots, t_n\} \in I, \; t_1 < \ldots < t_n$, as the probability measure $\mu_{t_1,\ldots,t_n}$, on the product space $(\mathcal{X}^n, \Sigma^n)$

$$\mu_{t_1,\ldots,t_n}(B) = P\{\omega : (X_{t_1}(\omega), \ldots, X_{t_n}, (\omega)) \in B\} \tag{8.1}$$

for all $B \in \Sigma^n$.

Recall that a family of sub-σ-algebras of $\mathcal{F}$ is *adapted* if every random variable $X_t, \; t \in I$, is $\mathcal{F}_t$ measurable. A family $\{\mathcal{F}_t\}_{t\in I}$ called a *filtration* of the process $\{X_t\}_{t\in I}$ if:

(1) $\mathcal{F}_s \subset \mathcal{F}_t$, for each $s < t$

(2) $\{X_t\}$ is $\{\mathcal{F}_t\}$ adapted.

In particular, $\{\mathcal{F}_t\}_{t\in I}$ is called the *natural filtration* of the process $\{X_t\}_{t\in I}$ if each $\mathcal{F}_t$ is defined as the smallest σ-algebra generated by all random variables X_s, for $s \leq t$.

In this book we shall be mostly concerned with stochastic processes on which some additional conditions are imposed. These additional conditions may concern random variables $X_t,\ t \in I$, the finite dimensional distributions, or the realizations of of $\{X_t\}_{t\in I}$.

A condition that concerns individual random variables X_t is their integrability. We distinguish the class of L^p-processes defined as follows. An L^p stochastic process, $p > 0$, is defined as a family $\{X_t\}_{t\in I}$ of real or complex random variables on Ω such that

$$\|X_t\|_{L^p} = (E|X_t|^p)^{1\wedge\frac{1}{p}} < \infty, \text{ for all } t \in I\,.$$

Thus an L^p-process $\{X_t\}_{t\in I}$ can be interpreted as a trajectory

$$t \longmapsto X_t$$

in the Fréchet space L^p (in the Banach space if $p \geq 1$, or in the Hilbert space if $p = 2$). Particularly important is the class of L^p processes with $p \geq 1$, which will be also called *integrable processes*.

Integrable processes appear in a natural way in connection with filtrations. Indeed, if $\{\mathcal{F}_t\}$ is a filtration of some stochastic process, then it can be associated with $\{\mathcal{F}_t\}$ the family of conditional expectation $\{E_t\}$

$$E_t \stackrel{\text{df}}{=} E(\,\cdot\,|\mathcal{F}_t)\,,\ \ t \in I\,.$$

In particular $\{E_t\}$ is a family of orthogonal projectors on the Hilbert space L^2. If X is an L^p-random variable then the stochastic process $X_t = E(X|\mathcal{F}_t)$ is an L^p-process.

In the class of integrable processes a prominent role plays the class of martingales defined as follows. Let $\{X_t\}_{t\in I}$ be an $L^p,\ p \geq 1$, stochastic process with a totally ordered index set I and $\{\mathcal{F}_t\}$ its filtration. We say that $\{X_t\}$ is a *martingale* (respectively, *submartingale* or *supermartingale*) relative to $\{\mathcal{F}_t\}$ if for each $s < t$

$$E(X_t|\mathcal{F}_s) = X_s \ \text{ a.e.} \tag{8.2}$$

(respectively, if equality (8.2) is replaced by '$\leq$' or '$\geq$').

A stochastic process that is a martingale has a number of important properties some of them can be found in Appendix 1. For instance, a discrete L^1 bounded martingale is almost surely convergent. A continuous time martingale, regarded as an L^p valued function, is an analog of a function of bounded variation. This means a possibility to define integrals with respect to martingales, which is the basis of stochastic calculus (see Appendix 1 for details).

Additional assumption on realizations are necessary when dealing with continuous time processes, i.e. when the index set I is an interval. Observe that without an additional assumption on a stochastic process $\{X_t\}_{t\in I}$ such simple expressions as $P\{X_t > 0\}$ or $\int_a^b X_t(\omega)\,dt$ may be meaningless. The reason is that in the above expressions an uncountable number of events is involved and that may lead to non-measurable sets. In order to avoid such situations it is usually assumed that considered processes are separable or measurable. We shall always tacitly assume, when it is necessary, that it is the case. These additional assumptions are not too restrictive, since the processes considered below possesses separable and measurable versions. In Appendix 1 the reader will find a more detailed discussion of this matter.

A stochastic process is called stationary if its probabilistic characteristic do not change in time. Depending on the king of characteristics that we take into account we distinguish two kinds of stationary processes. If all finite dimensional distributions of a process $\{X_t\}$ are time invariant, i.e.

$$P\{(X_{t_1+s},\ldots,X_{t_n+s})\in B\} = P\{(X_{t_1},\ldots,X_{t_n})\in B\}\,,$$

for each $s, t_1, \ldots, t_n \in I$, $n \in \mathbb{N}$, and $B \in \Sigma^n$, then $\{X_t\}$ is said to be *stationary in a narrow sense* or *strictly stationary*.

A process $\{X_t\}$ is said to be *stationary in the wide sense* if it is L^2-stochastic process, i.e. each random variable X_t is square integrable, and if, for each $t \in I$, the mean value of X_t is constant

$$EX_t = m$$

(in the sequel we shall always assume that $m = 0$) and the covariance function

$$R(s,t) \stackrel{\text{df}}{=} EX_s\bar{X}_t - |m|^2$$

depends only on the difference between s and t

$$R(s+u, t+u) = R(s,t)\,,$$

for each $s, t, u \in \mathbb{R}$.

In the first kind of stationarity we do not impose any condition of integrability on the random variables X_t. Thus strict stationarity does not imply stationarity in the wide sense and neither converse implication is true. However in the class of L^2-stochastic processes each strictly stationary process is also stationary in the wide sense.

Strictly stationary stochastic processes arises from dynamical systems with measure preserving transformations. Indeed, consider a dynamical system $(\mathcal{X},\Sigma,\mu;\{S_t\})$, where μ is a normalized measure and each transformation S_t is measure preserving, i.e. $\mu(S_t^{-1}A) = \mu(A)$, for each $A \in \Sigma$. Then let us define the probability space $(\Omega, \mathcal{F}, P)$ by putting $\Omega = \mathcal{X}$, $\mathcal{F} = \Sigma$, and $P = \mu$, and the stochastic process:

$$X_t(\omega) = f(S_t x)\,,$$

where f is some real (or complex) valued measurable function on $\mathcal{X}$. We have

$$\begin{aligned}P\{\omega \in \Omega : \quad & X_{t_1+s}(\omega) \in A_1, \ldots, X_{t_n+s}(\omega) \in A_n\} = \\ & = \mu\{x \in \mathcal{X} : S_{t_1+s}x \in f^{-1}(A_1), \ldots, S_{t_n+s}x \in f^{-1}(A_n)\} \\ & = \mu(S_{t_1+s}^{-1} f^{-1}(A_1) \cap \ldots \cap S_{t_n+s}^{-1} f^{-1}(A_n)) \\ & = \mu(S_s^{-1}(S_{t_1}^{-1} f^{-1}(A_1) \cap \ldots \cap S_{t_n}^{-1} f^{-1}(A_n))) \\ & = \mu(S_{t_1}^{-1} f^{-1}(A_1) \cap \ldots \cap S_{t_n}^{-1} f^{-1}(A_n)) \\ & = \mu\{x \in \mathcal{X} : S_{t_1}x \in f^{-1}(A_1), \ldots, S_{t_n}x \in f^{-1}(A_n)\} \\ & = P\{\omega \in \Omega : X_{t_1}(\omega) \in A_1, \ldots, X_{t_n}(\omega) \in A_n\},\end{aligned}$$

for any choice $t_1, \ldots, t_n, s \in I$ and Borel sets $A_1, \ldots, A_n$. This shows that $\{X_t\}$ is strictly stationary on I. Choosing as f a square integrable function we obtain a stochastic process which is stationary both in the strict and in the narrow sense.

It should be noticed here that a dynamical system has a more complex structure than a stochastic process. Not only each function f defines a stationary stochastic process, but there is also an underlying dynamics $\{S_t\}$ that can be linked with a time operator. Stochastic processes do not posses such underlying dynamics in general, but there are some exceptions. A natural dynamics can be distinguished in stationary processes and some other related stochastic processes, like self-similar ones.

8.1 TIME OPERATORS OF THE STOCHASTIC PROCESSES STATIONARY IN WIDE SENSE

In this section we shall focus our attention on stochastic processes which are stationary in the wide sense, which will be called here simply *stationary*.

Let $\{X_t\}_{t\in I}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, P)$ with values in the state space $\mathcal{X}$. Let the set I of indices, interpreted here as *time*, be either the real line $\mathbb{R}$ or the set of integers $\mathbb{Z}$ and let the state space $\mathcal{X}$ be a subset of an Euclidean space. For simplicity, we assume that $\mathcal{X} = \mathbb{C}$. Let $\{\mathcal{F}_t\}_{t\in I}$ be the natural filtration determined by process $\{X_t\}_{t\in I}$. We assume in addition that the σ-algebra generated by all X_t coincides with $\mathcal{F}$. Denote by E_t the conditional expectation with respect to $\mathcal{F}_t$.

As it was pointed out previously the conditional expectations E_t, regarded as operators on the Hilbert space $L^2 \subset L^1$, are orthogonal projectors and define a self-adjoint operator T

$$Tf = \int_I t\, dE_t f\,. \tag{8.3}$$

Operator T is densely defined on L^2, hence it is also correctly defined on a dense subset of L^1.

The operator T is an analog of time operators associated with K-system or exact systems. One may therefore ask whether T can also be a time operator of the stochastic process $\{X_t\}$. This question acquires meaning if there exists some dynamics (or flow of time) consistent with the filtration determined by $\{X_t\}$. Such dynamics exists for some classes of stochastic processes. The most natural class of stochastic processes that admit a time operator is the class of stochastic processes which are stationary in wide sense.

Assume that the process $\{X_t\}_{t\in I}$ is stationary (in the wide sense). Let $\mathcal{H}(X)$ (resp. $\mathcal{H}_t(X)$) denote the closed sub-space of L^2 spanned by the linear combinations of $X_s,\ s\in I$ (resp. $s\in I,\ s\le t$). $\mathcal{H}(X)$ (resp. $\mathcal{H}_t(X)$) is also a Hilbert space that is in general a proper subspace of L^2 (resp $L^2(\mathcal{F}_t)$). Note that, for each t, the conditional expectation E_t is the orthogonal projection from $\mathcal{H}(X)$ onto $\mathcal{H}_t(X)$.

Define the shift operator

$$V_t X_s \overset{\text{df}}{=} X_{s+t}$$

extending it by linearity on finite linear combinations of $X_{t_1},...,X_{t_n}$. The operator V_t preserves the norm

$$\begin{aligned}
\|V_t(\sum_{j=1}^{n} a_j X_{t_j})\|^2 &= \|\sum_{j=1}^{n} a_j X_{t_j+t}\|^2 \\
&= \sum_{j,k=1}^{n} a_j \bar{a}_k E X_{t_j+t}\bar{X}_{t_k+t} \\
&= \sum_{j,k=1}^{n} a_j \bar{a}_k E X_{t_j}\bar{X}_{t_k} \\
&= \|\sum_{j=1}^{n} a_j X_{t_j}\|^2 .
\end{aligned}$$

Therefore V_t extends to an unitary operator on $\mathcal{H}(X)$.

Proposition 8.1 *The operator T associated with $\{X_t\}$ through* (8.3) *is a time operator with respect to $\{V_t\}$.*

Proof. To see that T and $\{V_t\}$ satisfy

$$TV_t = V_t T + tV_t \tag{8.4}$$

we need the equality

$$V_t E_s = E_{s+t} V_t \tag{8.5}$$

Let Y be arbitrary element from $\mathcal{H}(X)$. Then $E_s Y \in \mathcal{H}_s(X)$ and $V_t E_s Y \in \mathcal{H}_{s+t}(X)$. We can now apply the following simple property of Hilbert spaces:

> Let $\mathcal{H}_0$ be a closed subspace of the Hilbert space $\mathcal{H}$. If $x_0 \in \mathcal{H}_0$ and $x \in \mathcal{H}$, then
>
> $$x - x_0 \perp \mathcal{H}_0 \iff x_0 = P_0 x,$$
>
> where P_0 is the orthogonal projector on $\mathcal{H}_0$.

By the above property $Y - E_s Y \perp \mathcal{H}_s(X)$. Since V_t maps $\mathcal{H}_s$ onto $\mathcal{H}_{s+t}$, we have also have $V_t(Y - E_s Y) \perp \mathcal{H}_{s+t}(X)$. Thus

$$V_t Y - V_t E_s Y \perp \mathcal{H}_{s+t}(X)$$

and applying again the above property we obtain

$$V_t E_s Y = E_{s+t} V_t Y,$$

which proves (8.5).

Now, to prove (8.4) it is enough to use representation (8.3) of T and apply (8.5).

In the sequel we shall assume that the stationary process $\{X_t\}$ is purely nondeterministic, which means that

$$\bigcap_{t \in I} \mathcal{H}_t(X) = \{0\}\,.$$

Purely nondeterministic processes correspond to K-systems or exact systems, i.e. the dynamical systems for which time operators have been already constructed.

It should be noticed that above assumption does not restrict the class of stationary processes for which the time operators can be constructed. Indeed, it follows from the Wold decomposition theorem that any stationary process $\{X(t)\}$ can be represented uniquely as

$$X(t) = X_1(t) + X_2(t)\,,$$

where the process $\{X_1(t)\}$ is purely nondeterministic and $\{X_2(t)\}$ is deterministic, i.e. $\mathcal{H}_t(X) = \mathcal{H}(X)$, for all t. Moreover $\{X_1(t)\}$ and $\{X_2(t)\}$ are L^2-orthogonal. The deterministic process is measurable with respect to the σ-algebra $\bigcap_t \mathcal{F}_t$. Therefore it is not affected be the transition $\mathcal{F}_s \to \mathcal{F}_t$, for $s < t$, and we can say that there is no flow of time for a deterministic process. Consequently, the time operator must be identity on the space $\mathcal{H}(X_2)$, which is the orthogonal complement of $\mathcal{H}(X_1)$.

Until the end of this section we shall confine ourself to discrete time stationary processes. Namely, let us consider L^2-stationary sequences $X_n,\ n \in \mathbb{Z}$ and assume that they are purely nondeterministic. Under this assumption we can characterize the time operator associated with a stationary sequence as follows:

Theorem 8.1 *Let $X = \{X_n\}_{n\in\mathbb{Z}}$ be an L^2-stationary sequence. Then there is an orthonormal basis $\{Y_n\}$ in the Hilbert space $\mathcal{H}(X)$ such that the time operator T associated with the process X has the form*

$$T = \sum_{n\in\mathbb{Z}} n P_n\,, \tag{8.6}$$

where P_n is the projection on the space spanned byY_n.

Proof.

The process $\{X_n\}$ can be represented in the form

$$X_n = \sum_{k=-\infty}^{n} a_{n-k} Y_k \,, \quad n \in \mathbb{Z} \,, \tag{8.7}$$

where $Y_k \in L^2\,,\ k \in \mathbb{Z}$, is a sequence of random variables that are mutually orthonormal, and $\sum_{k=0}^{\infty} |a_k|^2 < \infty$. Moreover we have

$$\mathcal{H}_n(X) = \mathcal{H}_n(Y)\,,$$

which implies that each E_n coincides with the orthogonal projection onto $\mathcal{H}_n(Y)$.

Let $P_n \stackrel{\mathrm{df}}{=} E_n - E_{n-1}$. Since E_n is the orthogonal projection on $\mathcal{H}_n(Y)$, P_n is the orthogonal projection on the space $\mathcal{H}_n(Y) \ominus \mathcal{H}_{n-1}(Y)$, which is, by the assumption that $\{X_n\}$ is purely nondeterministic, one dimensional space generated by Y_n. It follows from the general form of the time operator (see the beginning of Section 2) that he time operator (8.3) associated with $\{X_n\}$ is given by (8.6)

The above theorem connects the time operator with the spectral representation (spectral process) of the stationary process. In this way we may extend the time operator in a natural way beyond the class of stationary processes. Namely to such processes that can be in some way determined by spectral processes.

In this and the next section we shall study in details the spectral properties and extensions of the above defined time operator T. First however, we would like to present the stochastic interpretation of T and its time recalling $\Lambda(T)$. We shall also point out some connection of T and $\Lambda(T)$ with filtering and the generation of new processes.

Recall that for any $Z \in \mathcal{H}(X)$ the value $\langle Z, TZ\rangle$ is interpreted as the average age of Z. It follows from the spectral resolution

$$T = \sum_{n=1}^{\infty} nP_n \tag{8.8}$$

that T attributes the age n to the random variable Y_n. Since $\{Y_n\}$ form a complete orthonormal system in the Hilbert space $\mathcal{H}(X)$, each $Z \in \mathcal{H}(X)$ can be represented as

$$Z = \sum_{n\in\mathbb{Z}} b_n Y_n\,,$$

where $\sum_n |b_n|^2 < \infty$. Consequently each element of $\mathcal{H}(X)$ can be decomposed in terms of eigenvectors of T. Moreover the domain of T consists of all such Z for which $\sum_n n|b_n|^2 < \infty$. We can then say that each Z from the domain of T has well definite age, which is

$$\langle Z, TZ\rangle = \sum_n n|b_n|^2\,.$$

Moreover if

$$\Lambda = \Lambda(T) = \sum_{n\in\mathbb{Z}} \lambda_n P_n \tag{8.9}$$

is a function of the time operator then Λ attributes the age λ_n to Y_n.

It is interesting to interpret Λ-operator in terms of filtering theory. From this point of view the sequence $\{Y_k\}$ is a noise and X_n is the response of a linear homogeneous physically realizable system at instant n to the sequence of impulses $\{Y_k\}$. According to this interpretation at every time instant n there is the same contribution of noise. Λ-operator changes the magnitude of impulses of the noise. As a result the system becomes inhomogeneous with respect to time although still physically realizable.

It should be stressed that the time operator T, and consequently also Λ, has been actually associated with the spectral process (sequence) $\{Y_n\}$. However, a given spectral process may represent many different stationary and non-stationary processes subordinated to $\{Y_n\}$. Each such process can now be regarded as the evolution of some random variable $X \in \mathcal{H}(X)$ transformed through an appropriately chosen Λ-operator. In particular the initial stationary process $\{X_n\}$ can also be seen as a result of a rescaling of the time operator T. Indeed, notice first that it is enough to define a function $\Lambda = \Lambda(T)$ only on the set of integers. Thus taking a sequence $\{\lambda_k\}$, where

$$\lambda_k = \begin{cases} a_{-k}\,, & \text{for } k \leq 0 \\ 0\,, & \text{for } k > 0 \end{cases} \tag{8.10}$$

we get

$$\begin{aligned} X_n &= \sum_{k=-\infty}^{n} a_{n-k} Y_k \\ &= V^n \left(\sum_{k=-\infty}^{0} a_{-k} Y_k \right) \\ &= V^n \left(\sum_{k=-\infty}^{0} \lambda_k Y_k \right) \\ &= V^n \Lambda X_0\,. \end{aligned}$$

Since, on the other hand, $X_n = V^n X_0$, we see that Λ acts on X_0 as the identity operator. However, a different choice of $\{\lambda_k\}$ will lead to different processes. Notice that choosing a random variable $Z_0 \in \mathcal{H}(X)$, a sequence $\{\lambda_k\}$, $\sum_k \lambda_k^2 < \infty$, and putting

$$Z_n = V^n \Lambda Z_0 \tag{8.11}$$

we obtain a stationary process. If either $\lambda_k = 0$, for $k > 0$, or $Z_0 \in \mathcal{H}(X)$ then $\{Z_n\}$ is physically realizable.

In a forthcoming section we shall show that a large variety of stochastic processes can arise from a very simple process just through a time operator rescaling.

Remark. Considered above time operator associated with a stationary process may seem to be completely different than the time operator associated with a K-system. Although T keeps step with the evolution semigroup $\{V_t\}$ the time eigensubspaces are one -dimensional. This is in a sharp contrast with the properties of an "usual" time operator that has eigenfunctions of infinite multiplicity. However, let us notice that formula (8.3) (or 8.6)) actually defines T on the whole space $L^2(\mathcal{F}) \stackrel{\text{df}}{=} L^2(\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is the σ-algebra generated by all X_t , $t \in I$. Suppose that there exists a unitary group $\{U_t\}_{t\in I}$ ($I = \mathbb{R}$ or $\mathbb{Z}$) on $L^2(\mathcal{F})$ that extends $\{U_t\}_{t\in I}$, i.e, $\{U_t\}$ restricted to $\mathcal{H}(X)$ coincides with $\{V_t\}$. If the stationary process is purely nondeterministic and if the unitary group $\{U_t\}$ is such that $L^2(\mathcal{F}_0) \subset U_t L^2(\mathcal{F})$, for $t > 0$, where $\mathcal{F}_0 = \sigma(X_t)_{t<0}$, then, by von Neumann's theorem, the group has a homogeneous Lebesgue spectrum of infinite multiplicity. Therefore T is an ordinary Misra's time operator with respect to the extended group $\{U_t\}$.

8.2 TIME OPERATORS OF STRICTLY STATIONARY PROCESSES - FOCK SPACE

It follows from the previous remark that the problem of extension of the domain of T depends on a possibility of extensions of the semigroup $\{V_t\}$ associated with T. We present below such an extension of the evolution $\{V_t\}$ from $\mathcal{H}(X)$ to the unitary evolution on $L^2(\mathcal{F})$ in the case when the considered process is a Gaussian stationary sequence $\{X_n\}_{n\in\mathbb{Z}}$. In this case the Hilbert space $\mathcal{H}(X)$ is the space of Gaussian random variables. In particular, since the random variables Y_k from the representation (8.7) are mutually orthogonal and normalized they are independent with $\mathcal{N}(0,1)$ distributions, and called the discrete white noise.

The shifts V^m, $V^m X_n = X_{n+m}$, can be extended from $\mathcal{H}(X)$ to the shifts $\widetilde{V}^m$ acting on all functions $f(X_{n_1}, \ldots, X_{n_k})$ of finite subsequences of $\{X_n\}$ by putting

$$\widetilde{V}^m f(X_{n_1}, \ldots, X_{n_k}) = f(X_{n_1+m}, \ldots, X_{n_k+m}) . \tag{8.12}$$

It is easy to see that $\widetilde{V}^m$ is an isometry on functions of the form $f(X_{n_1}, \ldots, X_{n_k})$. Since such functions generate $L^2(\mathcal{F})$, $\widetilde{V}^m$ extends to an isometry on the whole $L^2(\mathcal{F})$. We shall show below that the unitary group $\widetilde{V}^m$ on $L^2(\mathcal{F})$ is associated with the extended time operator $\widetilde{T}$ defined through (8.8). First, however, we shall find the explicit form of the spectral projectors of $\widetilde{T}$.

Let us note that the Hilbert space of square integrable functions $f(X)$ of a Gaussian random variable can be identified with $L^2(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \gamma)$, where $\mathcal{B}_{\mathbb{R}}$ is the Borel σ-algebra on $\mathbb{R}$ and γ the Gaussian measure on $\mathbb{R}$ with the density $(2\pi)^{-1/2}e^{-x^2/2}$. Therefore $L^2(\mathcal{F})$ can be identified with the infinite product space $L^2(\mathbb{R}^\infty, \mathcal{B}_{\mathbb{R}}^\infty, \gamma_\infty)$, where γ_∞ is the corresponding product Gaussian measure on $\mathbb{R}^\infty$. This identification allows us to express spectral projectors of T in terms of the Wick polynomials.

In the sequel we shall use the following:

Proposition 8.2 [Se,Ma] *If $\{Y_k\}$ is an orthonormal basis in $\mathcal{H}(X)$ then the family of all products*

$$H_{n_1}(Y_{m_1}) \cdot \ldots \cdot H_{n_k}(Y_{m_k}) , \tag{8.13}$$

where $H_n(x)$ denotes the n-th Hermite polynomial with the leading coefficient 1, form a complete orthogonal system in $L^2(\mathcal{F})$.

Recall that the n-th Hermite polynomial with the leading coefficient 1 is

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} .$$

In particular $H_0 \equiv 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x, \ldots$. Recall also that $H_n(x)$, $n = 0, 1, \ldots$, form a complete orthogonal basis in $L^2(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \gamma)$, i.e. for any $f \in L^2(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \gamma)$

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} H_n(x) ,$$

where $a_n = \langle f, H_n \rangle$.

Our next task is to find the explicit form of the eigenprojectors E_n of the time operator (8.8) on the space $L^2(\mathcal{F})$, i.e. the conditional expectations $E(\cdot|\mathcal{F}_n)$. Such explicit form can be done for all polynomials

$$p(X_{n_1}, \ldots, X_{n_j}) \tag{8.14}$$

of the random variables $X_{n_1}, \ldots, X_{n_j}$, with $j = 1, 2, \ldots$ and $n_1, \ldots, n_j \in \mathbb{Z}$, which, as follows from the above proposition, form a dense subspace of $L^2(\mathcal{F})$.

Theorem 8.2 *Suppose that $\{X_n\}_{n\in\mathbb{Z}}$ is a Gaussian stationary sequence on the probability space $(\mathcal{F}, P)$ and assume that $\mathcal{F}$ coincides with the smallest σ-algebra generated by all X_n, $n \in \mathbb{Z}$. Let $\{\widetilde{V}^m\}_{m\in\mathbb{Z}}$ be the extension* (8.12) *on $L^2(\mathcal{F})$ of the group of shifts $\{V^m\}_{m\in\mathbb{Z}}$ of $\{X_n\}$. Then the extension $\widetilde{T}$ on $L^2(\mathcal{F})$ of T defined by* (8.8) *is a time operator associated with $\{\widetilde{V}^m\}$. The eigenprojections $\widetilde{P}_n$ of $\widetilde{T}$ are of the form*

$$\widetilde{P}_n = \sum_{k=0}^{\infty} P_n(k) ,$$

where the orthogonal projection $P_n(k)$, $k = 0, 1, \ldots$, when restricted a polynomial of order k, is the Wick Polynomial corresponding to the Gaussian process $\{X_j\}_{j\leq n}$. In particular, if $\{Y_j\}_{j\leq n}$ is an orthonormal system in $L^2(\mathcal{F}_n)$ then for any finite subsequence $Y_{j_1}, \ldots, Y_{j_r}$, $j_r \leq n$, and any homogeneous polynomial of order k:

$$p(Y_{j_1}, \ldots, Y_{j_r}) = \sum_{l_1,\ldots,l_r} a_{l_1,\ldots,l_r} Y_{j_1}^{l_1} \ldots Y_{j_r}^{l_r} \tag{8.15}$$

the orthogonal projection $P_n(k)$ of $p(Y_{j_1}, \ldots, Y_{j_r})$ is of the form:

$$P_n(k)(p(Y_{j_1} \ldots Y_{j_r})) = \sum_{l_1,\ldots,l_r} a_{l_1,\ldots,l_r} H_{l_1}(Y_{j_1}), \ldots, H_{l_r}(Y_{j_r}) \tag{8.16}$$

Proof. Denote by $\mathcal{G}_n$, $n \in \mathbb{Z}$ the Hilbert space $L^2(\mathcal{F}_n)$. It is easy to see that the σ-algebra $\mathcal{F}_n$ coincides with the smallest σ-algebra generated by Y_k, $k \le n$. Moreover the above proposition implies that the products (8.13) with $j_k \le n$ form a complete orthogonal system in $\mathcal{G}_n$. The spaces G_n are the eigenspaces of $\widetilde{T}$ corresponding to the eigenvalue n. Next, denote by $\mathcal{H}_{\le k}(n)$ the Hilbert space spanned by all polynomials $p(Y_{n_1}, \dots, Y_{n_r})$ of the random variables $Y_{n_1}, \dots, Y_{n_r}$ with $r = 1, 2, \dots$ and $n_1, \dots, n_r \le n$. Let $\mathcal{H}_0(n)$ denote the space of constants and let $\mathcal{H}_k(n)$ be the orthogonal complement of $\mathcal{H}_{\le k-1}(n)$ in $\mathcal{H}_{\le k}(n)$, i.e. $\mathcal{H}_{\le k-1}(n) \oplus \mathcal{H}_k(n) = \mathcal{H}_{\le k}(n)$. Therefore

$$\mathcal{G}_n = \bigoplus_{k=0}^{\infty} \mathcal{H}_k(n)\,.$$

Applying the same arguments as in [Ma] we see that the orthogonal projection of the polynomial (8.14) of order k onto$\mathcal{H}_k(n)$ is the Wick polynomial corresponding to the Gaussian process $\{X_t\}_{t \le n}$. The explicit form of these projection is known for any polynomial of an orthogonal system in $\mathcal{G}_n$. In particular, for polynomials of the form (8.15) we obtain (8.16). Since $\{Y_j\}_{j \le k}$ form a complete orthonormal system in $\mathcal{G}_n$ and $\bigcup_{n \in \mathbb{Z}} \mathcal{G}_n = L^2(\mathcal{F})$, the explicit form of the spectral projectors of the time operator $\widetilde{T}$ is known on a dense subspace of $L^2(\mathcal{F})$. Finally we have to show that $\widetilde{T}$ is the time operator with respect to the extended shift $\widetilde{V}$. Indeed, according to Lemma **??** and properties of conditional expectation it is enough to show that

$$\widetilde{V}^m L^2(\mathcal{F}_k) = L^2(\mathcal{F}_{k+m})\,.$$

However, $L^2(\mathcal{F}_k)$ is the closure of the space spanned by all $Y_{j_1}^{l_1} \dots Y_{j_r}^{l_r}$ with $j_1, \dots, j_r \le k$. By the definition (8.12) of $\widetilde{V}$

$$\widetilde{V}^m (Y_{j_1}^{l_1} \dots Y_{j_r}^{l_r}) = Y_{j_1+m}^{l_1} \dots Y_{j_r+m}^{l_r}\,.$$

Since the right hand sides of the latter equality spans $L^2(\mathcal{F}_{k+m})$ this concludes the proof.

Remark. Applying the above theorem we can find the explicit form of the projection of an arbitrary, not necessarily homogeneous, polynomial. Indeed, if $q = q(Y_{n_1}, \dots, Y_{n_r})$ is an arbitrary polynomial of order k then it can be written as a sum $q = p + q_1$, where p is a homogeneous polynomial of order k of the form (8.15) and the order of q_1 is less than k. Since

$$P_n(k)(p + q_1) = P_n(k)(p)\,,$$

we can apply (8.16). If the order of q is greater than k, say $k + r$, then we calculate first $P_n(k+r)(q)$ and next, using the fact that the order of $q - P_n(k+r)(q)$ is less than $k + r$, we calculate

$$P_n(k+r-1)(q - P_n(k+r)(q)) = P_n(k+r-1)(q)\,.$$

We repeat this procedure until we reach the order k.

9

Time operator of diffusion processes

Time operator of diffusion processes

The purpose of this section is to introduce time operators for dissipative systems like the diffusion equation. The idea of the construction of time operators for such systems is based on the intertwining formula introduced by Misra, Prigogine and Courbage in the context of unstable Kolmogorov dynamical systems (see Section 15). Time operators T are canonically conjugate to the group of unitary evolution $\{U_t\}$ acting on the space of square integrable phase functions $f \in L^2(\mathcal{X})$. Let us recall here, for the readers convenience, that a time operator T for the unitary group $\{U_t\}$ on $L^2(\mathcal{X})$ should satisfy the following conditions:

1) T is self-adjoint;
2) The domain D_T of T is dense in $L^2_{\mathcal{X}}$;
3) The unitary group $\{U_t\}$ does not lead out of D_T:

$$U_t(D_T) \subset D_T, \;\; t \in \mathbb{R};$$

4) T satisfies the canonical commutation relation with U_t

$$TV_t = V_t(T + tI). \tag{9.1}$$

The original motivation for introducing the time operators by Misra, Prigogine and Courbage was the fact that they give rise to the definition of entropy operators and to intertwining transformation of the unitary group $\{U_t\}$ with two distinct Markov semigroups W_t^+, $t \in [0, +\infty)$ or $t = 0, 1, 2, \ldots$ and W_t^-, $t \in (-\infty, 0]$ or $t = 0, -1, -2, \ldots$, corresponding to the forward and backward time directions:

$$\begin{aligned} \Lambda_+ U_t &= W_t^+ \Lambda_+, \quad t \in [0, +\infty) \;\text{ or }\; t = 0, 1, 2, \ldots, \\ \Lambda_- U_t &= W_t^- \Lambda_-, \quad t \in (-\infty, 0] \;\text{ or }\; t = 0, -1, -2, \ldots. \end{aligned} \tag{9.2}$$

The strategy for constructing time operator for diffusion equation is motivated by the intertwining formula (9.2), which relates two different evolutions. Then the intertwining transformation transports the time operator from one evolution to the other. The transported time operator, may of course not be self-adjoint. These general results are discussed below. Then we will show that the semigroup associated with the diffusion equation is intertwined with the unilateral shift. Using this fact we construct a non-self-adjoint time operator for the diffusion equation in and obtain the spectral resolution, the age eigenstates and the shift representation of the solution of the diffusion equation. On the basis of the intertwining of the Poincare–Telegraphist equation with the wave equation [???] we will construct also a self-adjoint time operator for the telegraphist equation.

9.1 TIME OPERATORS FOR SEMIGROUPS AND INTERTWINING

We generalize the concept of time operator for an arbitrary semigroup and define the conditions which relate the time operators of intertwined semigroups.

Definition 1. Let $\{W_t : t \geq 0\}$ be a semigroup of continuous linear operators on a Hilbert space $\mathcal{H}$. A linear operator T on $\mathcal{H}$ is called a *time operator* for the semigroup W_t if the following conditions are satisfied:

1) the domain D_T of T is dense in $\mathcal{H}$;

2) the semigroup W_t does not lead out of D_T:

$$W_t(D_T) \subset D_T \ \text{ for } \ t \geq 0.$$

3) T satisfies the canonical commutation relations with W_t:

$$TW_t = W_t(T + t)\,. \tag{9.3}$$

One can see that we just eliminated the condition of self-adjointness from the definition of time operators for unitary groups.

Definition 2. The semigroup V_t^1, $t \geq 0$ of bounded linear operators on the Hilbert space $\mathcal{H}^1$ is intertwined with the semigroup V_t^2, $t \geq 0$ of bounded linear operators on the Hilbert space $\mathcal{H}^2$ if and only if there exists a linear transformation Λ from $\mathcal{H}^1$ to $\mathcal{H}^2$ with the following properties:

1) Λ is one-to-one and the domain D_Λ of Λ is dense in $\mathcal{H}^1$;

2) the image $\Lambda(D_\Lambda)$ is dense in $\mathcal{H}^2$, so the operator Λ^{-1} is densely defined;

3) $V_t^1(D_\Lambda) \subseteq D_\Lambda$ for all $t \geq 0$;

4) the intertwining relation holds on $\Lambda(D_\Lambda) = D_{\Lambda^{-1}}$:

$$V_t^2 = \Lambda V_t^1 \Lambda^{-1}, \quad t \geq 0.$$

We see that if the semigroup V_t^1 is intertwined with the semigroup V_t^2 via Λ, then the semigroup V_t^2 is intertwined with the semigroup V_t^1 via Λ^{-1}.

Indeed properties 1) and 2) for Λ^{-1} follow from 2) and 1) for Λ. Property 3) for Λ^{-1} follows from properties 3) and 4) for Λ. Indeed $V_t^1 f \in D_\Lambda$ for any $f \in D_\Lambda$ from

3). From 4) we have that $\Lambda V_t^1 f = V_t^2 \Lambda f$. Therefore $V_t^2(\Lambda f) \in \Lambda(D_\Lambda) = D_{\Lambda^{-1}}$. Thus V_t^2 does not lead out of $D_{\Lambda^{-1}}$ which is property 3) for Λ^{-1}. Property 4) for Λ^{-1} follows easily from property 4) for Λ and property 3) for Λ^{-1}.

The relation between the time operators of intertwined semigroups is given by the following

Lemma 9.1 *Let V_t^1 and V_t^2, $t \in [0, \infty)$ be semigroups on Hilbert spaces $\mathcal{H}^1$ and $\mathcal{H}^2$ respectively, intertwined via the linear operator $\Lambda : \mathcal{H}^1 \to \mathcal{H}^2$ and T_1 be a time operator for V_t^1. Then the transported operator*

$$T_2 = \Lambda T_1 \Lambda^{-1} \tag{9.4}$$

is a time operator for V_t^2, $t \geq 0$ if the domain

$$D_{T_2} = \{f \in \mathcal{H}^2 | f \in D_{\Lambda^{-1}},\ \Lambda^{-1} f \in D_{T_1},\ T_1 \Lambda^{-1} f \in D_\Lambda\}$$

is dense in $\mathcal{H}^2$.

Proof. Let $f \in D_{T_2}$. We have to prove that $T_2 V_t^2 f$ is well-defined and $T_2 V_t^2 f = V_t^2 T_2 f + t V_t^2 f$. By the definition of T_2 and by the conditions of lemma we have

$$T_2 V_t^2 f = \Lambda T_1 \Lambda^{-1} \Lambda V_t^1 \Lambda^{-1} f = \Lambda T_1 V_t^1 \Lambda^{-1} f.$$

From the inclusion $f \in D_{T_2}$ we obtain $g = V_{T_2} f = \Lambda V_t^1 \Lambda^{-1} f \in Im\,(\Lambda) = D_{\Lambda^{-1}}$, $\Lambda^{-1} g = V_t^1 \Lambda^{-1} f$. But $\Lambda^{-1} f \in D_{T_1}$. Hence, $\Lambda^{-1} g = V_t^1 \Lambda^{-1} f \in D_{T_1}$. By the definition of the time operator $T_1 V_t^1 \Lambda^{-1} f = V_t^1 (T_1 + t) \Lambda^{-1} f$. So

$$T_2 V_t^2 f = \Lambda T_1 V_t^1 \Lambda^{-1} f = \Lambda V_t^1 (T_1 + t) \Lambda^{-1} f =$$

$$= \Lambda V_t^1 T_1 \Lambda^{-1} f + t \Lambda V_t^1 \Lambda^{-1} f = \Lambda V_t^1 T_1 \Lambda^{-1} f + t V_t^2 f.$$

The term $V_t^2 f$ is well-defined. The inclusion $V_t^1 T_1 \Lambda^{-1} f \in D_\Lambda$ follows from $T_1 \Lambda^{-1} f \in D_\Lambda = Im\,\Lambda^{-1}$. Therefore term $\Lambda V_t^1 T_1 \Lambda^{-1} f$ is also well-defined.

Finally,

$$T_2 V_t^2 f = \Lambda V_t^1 T_1 \Lambda^{-1} f + t V_t^2 f = \Lambda V_t^1 \Lambda^{-1} \Lambda T_1 \Lambda^{-1} f + t V_t^2 f =$$

$$= V_t^2 T_2 f + t V_t^2 f = V_t^2 (T_2 + t) f.$$

9.2 INTERTWINING OF THE DIFFUSION EQUATION WITH THE UNILATERAL SHIFT

In order to construct a time operator for the one-dimensional diffusion equation

$$\frac{\partial \varphi_t(x)}{\partial t} = \frac{\partial^2 \varphi_t(x)}{\partial x^2} \tag{9.5}$$

we shall first show that the diffusion equation (9.5) is intertwined with a unilateral shift V_t, $t \in [0, \infty)$ and use the intertwining transformation

$$\Lambda W_t = V_t \Lambda$$

to transport the time operator of the shift to the time operator of (9.5).

Recall that for any $\varphi \in L^2_{\mathbb{R}}$ there exists a unique solution $\varphi_t(x)$ of the equation (9.5) with the initial condition $\varphi_0(x) = \varphi(x)$ such that $\varphi_t \in L^2_{\mathbb{R}}$ for all $t \geq 0$. Denote φ_t by $W_t\varphi$. The operators W_t, $t \in [0, \infty)$ define a continuous semigroup of bounded operators on $L^2_{\mathbb{R}}$ generated by the diffusion equation. The explicit formula for W_t is well known:

$$W_t\varphi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \check{\varphi}(y) e^{-ty^2} e^{-ixy}\, dy\,, \tag{9.6}$$

where $\check{\varphi}$ is the inverse Fourier transform:

$$\check{\varphi}(y) = \Phi^{-1}\varphi(y) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \varphi(x) e^{ixy}\, dx.$$

Obviously the semigroup W_t preserves the spaces of even $\mathcal{H}_e$ and of odd $\mathcal{H}_o$ functions in $L^2_{\mathbb{R}}$:

$$L^2_{\mathbb{R}} = \mathcal{H}_e \oplus \mathcal{H}_o.$$

Let P_e and P_o be the orthoprojections onto $\mathcal{H}_e$ and $\mathcal{H}_o$ respectively.

In the sequel we will need a few technical lemmata. Let $M_e : L^2_{\mathbb{R}_+} \to \mathcal{H}_e$ and $M_o : L^2_{\mathbb{R}_+} \to \mathcal{H}_o$ be the operators defined by the formulae

$$M_e\alpha(x) = \int\limits_0^{\infty} \alpha(u) e^{-ux^2}\, du, \tag{9.7}$$

$$M_o\alpha(x) = x \int\limits_0^{\infty} \alpha(u) e^{-ux^2}\, du. \tag{9.8}$$

Lemma 9.2 *The operators M_o and M_e are densely defined, closed, injective operators with dense image.*

Proof. Let $\alpha \in L^2_{\mathbb{R}_+}$. Then by Cauchy–Bounjakovskii inequality

$$\left| \int\limits_0^{\infty} \alpha(u) e^{-ux^2}\, du \right| \leq \|\alpha\|_{L^2} \sqrt{\int\limits_0^{\infty} e^{-2ux^2}\, du} = \frac{\|\alpha\|_{L^2}}{\sqrt{2}|x|}\,. \tag{9.9}$$

Let $\alpha \in L^1_{\mathbb{R}_+}$. Then

$$\left| \int_0^\infty \alpha(u) e^{-ux^2}\, du \right| \le \|\alpha\|_{L^1} \,. \tag{9.10}$$

Let $\alpha \in L^\infty_{\mathbb{R}_+}$. Then

$$\left| \int_0^\infty \alpha(u) e^{-ux^2}\, du \right| \le \|\alpha\|_{L^\infty} \int_0^\infty e^{-ux^2}\, du = \frac{\|\alpha\|_{L^\infty}}{x^2} \,. \tag{9.11}$$

Now let $\alpha \in L^2_{\mathbb{R}_+} \cap L^1_{\mathbb{R}_+} \cap L^\infty_{\mathbb{R}_+}$. Obviously the latter set is dense in $L^2_{\mathbb{R}_+}$. By (9.9) and (9.10) $|M_e\alpha(x)| \le \min\left\{ \|\alpha\|_{L^1}, \frac{\|\alpha\|_{L^2}}{\sqrt{2}|x|} \right\}$ and therefore $M_e\alpha \in L^2_{\mathbb{R}_+}$. By (9.10) and (9.11) $|M_o\alpha(x)| \le \min\left\{ \frac{\|\alpha\|_{L^2}}{\sqrt{2}}, \|\alpha\|_{L^\infty} \right\}$ and therefore $M_o\alpha \in L^2_{\mathbb{R}_+}$. Hence, the operators M_o and M_e are densely defined. The injectivity of M_o and M_e follows from the injectivity of the Laplace transform, because M_e is the Laplace transform combined with the change of variables from x to x^2 and M_o is the same operator multiplied by x. Let us show that M_o and M_e are closed. From (9.9) it is obvious that these operators are continuous as operators from $L^2_{\mathbb{R}_+}$ to the space $L^0_{\mathbb{R}}$ of measurable functions with the topology of convergence by Lebesgue measure. Therefore the graphs of these operators are closed in the product $L^2_{\mathbb{R}_+} \times L^0_{\mathbb{R}}$ and hence these graphs are closed in the stronger L^2-topology. So the operators M_o and M_e are closed. It remains to prove that the images of M_o and M_e are dense. The proof of these two statements are similar. Let us prove that the image of M_e is dense. Suppose it is not dense. Then there exists a function $\alpha \in L^2_{\mathbb{R}_+}$, $\alpha \ne 0$ such that for all $z \in \mathbb{C}$ with positive real part we have

$$0 = \int_0^\infty \int_0^\infty e^{-z\tau} e^{-\tau x^2}\, d\tau \alpha(x) dx = \int_0^\infty e^{-z\tau} \int_0^\infty \alpha(x) e^{-\tau x^2}\, dx\, d\tau.$$

It means that the Laplace transform of the function $\int_0^\infty \alpha(x) e^{-\tau x^2}\, dx$ vanishes. Therefore $\kappa(\tau) = \int_0^\infty \alpha(x) e^{-\tau x^2}\, dx \equiv 0$. But $\kappa(\tau)$ is the Laplace transform of the function $\frac{\alpha(\sqrt{x})}{2\sqrt{x}}$. Therefore $\alpha = 0$. This contradiction completes the proof.

Let V_t, $t \in [0, \infty)$ be the right shift on $L^2_{[0,\infty)}$:

$$V_t\alpha(\tau) = \begin{cases} 0, \text{ if } \tau \in [0, t), \\ \alpha(\tau - t) \text{ if } \tau \in [t, +\infty) \,. \end{cases} \tag{9.12}$$

Let $\Lambda_e : L^2_{[0,\infty)} \to \mathcal{H}_e$ and $\Lambda_o : L^2_{[0,\infty)} \to \mathcal{H}_o$ be the operators

$$\Lambda_e = \Phi M_e, \qquad \Lambda_o = \Phi M_o\,. \tag{9.13}$$

Lemma 9.3 *The formulae*

$$W_t^o = \Lambda_o V_t \Lambda_o^{-1}, \qquad W_t^e = \Lambda_e V_t \Lambda_e^{-1} \tag{9.14}$$

are true for all $t \geq 0$, where $\Lambda_o = \Phi M_o$, $\Lambda_e = \Phi M_e$.

Proof. First, consider the semigroup $(G_t)_{t\geq 0}$ of continuous linear operators acting on $L^2_{\mathbb{R}}$ given by formula

$$G_t\varphi(x) = e^{-tx^2}\varphi(x)\,, \tag{9.15}$$

By formulae (9.6) and (9.15) we obtain

$$W_t = \Phi G_t \Phi^{-1}\,. \tag{9.16}$$

The operators G_t and Φ^{-1} also preserve the decomposition $L^2_{\mathbb{R}} = \mathcal{H}_o \oplus \mathcal{H}_e$. Let $G_t^o = G_t\big|_{\mathcal{H}_o}$ and $G_t^e = G_t\big|_{\mathcal{H}_e}$. Then from (9.16) we see that the statement of the lemma is equivalent to the equalities

$$G_t^o = M_o V_t M_o^{-1}, \qquad G_t^e = M_e V_t M_e^{-1}\,. \tag{9.17}$$

Now let $g \in D_{M_e^{-1}} = Im\, M_e$. It means that there exists $\alpha \in L^2_{[0,\infty)}$ such that

$$g(x) = M_e\alpha(x) = \int_0^\infty \alpha(\tau)e^{-\tau x^2}\,d\tau.$$

Then

$$\begin{aligned} M_o V_t M_o^{-1} g &= M_o V_t \alpha = \int_0^\infty V_t\alpha(\tau)e^{-\tau x^2}\,d\tau \\ &= \int_0^\infty \alpha(s)e^{-(t+s)x^2}\,ds = e^{-tx^2}\int_0^\infty \alpha(s)e^{-sx^2}\,ds \\ &= e^{-tx^2}g(x) = V_t g(x). \end{aligned}$$

The second equality can be similarly verified.

Thus we showed that the semigroups W_t^e and V_t are intertwined by Λ_e and the semigroups W_t^o and V_t are intertwined by Λ_o.

9.3 THE TIME OPERATOR OF THE DIFFUSION SEMIGROUP

In order to construct a time operator T_W for the semigroup W_t (9.6) we construct first time operators T_W^o and T_W^e for the restrictions $W_t^e = W_t\big|_{\mathcal{H}_e}$ and $W_t^o = W_t\big|_{\mathcal{H}_o}$ W_t onto H_e and H_o. The time operator T_W will be the sum of T_W^e and T_W^o:

$$T_W f = T_W^o P_o f + T_W^e P_e f.$$

We introduced previously time operators for discrete unilateral shifts. We can define also time operators for continuous unilateral shifts.

Lemma 9.4 *A self-adjoint time operator T for the semigroup V_t (9.12) is given by the canonical formula*

$$T\alpha(\tau) = \tau\alpha(\tau)\,.$$

Proof. The domain D_T includes all functions from $L^2_{[0,\infty)}$ with compact support. Therefore D_T is dense. The inclusion $V_t(D_T) \subset D_T$ is trivial. It remains to check the canonical commutation relation for T. Let $\alpha \in D_T$. Then

$$\begin{aligned} TV_t\alpha(\tau) &= \begin{cases} 0, \text{ if } \tau \in [0,t], \\ \tau\alpha(\tau - t) \text{ if } \tau \in (t, +\infty) \end{cases} \\ &= \begin{cases} 0, \text{ if } \tau \in [0,t], \\ (\tau - t)\alpha(\tau - t) \text{ if } \tau \in (t, +\infty) \end{cases} + t \begin{cases} 0, \text{ if } \tau \in [0,t], \\ \alpha(\tau - t) \text{ if } \tau \in (t, +\infty) \end{cases} \\ &= V_tT\alpha + tV_t\alpha. \end{aligned}$$

The explicit formula of the time operator for W_t is given by:

Theorem 9.1 *The operators $T^o_W = \Lambda_o T \Lambda_o^{-1}$ and $T^e_W = \Lambda_e T \Lambda_e^{-1}$ are time operators for the semigroups W^o_t and W^e_t respectively. The explicit formulae for these operators are*

$$T^e_W g(x) = \frac{1}{2} \int\limits_{|x|}^{\infty} y g(y)\, dy\,, \tag{9.18}$$

$$T^o_W g(x) = \frac{x}{2} \int\limits_{|x|}^{\infty} g(y)\, dy\,. \tag{9.19}$$

Proof. That the operators T^o_W and T^e_W are time operators for the semigroups W^o_t and W^e_t follows directly from lemmas 1–4. It suffices to prove (9.18) and (9.19) for some dense linear subspaces of $\mathcal{H}^e$ and $\mathcal{H}^o$. We shall prove them for the dense subspaces of $\mathcal{H}_e$ and $\mathcal{H}_o$ consisting respectively of the functions of the form

$$g(x) = \int\limits_0^{\infty} \alpha(\tau) e^{-\tau x^2}\, d\tau\,, \tag{9.20}$$

$$h(x) = x \int\limits_0^{\infty} \alpha(\tau) e^{-\tau x^2}\, d\tau\,, \tag{9.21}$$

where the function $\alpha \in L^2_{[0,\infty)}$ has compact support in $(0, \infty)$.

For even functions (9.20) we have:

$$\begin{aligned}
\Phi^{-1}g(x) &= \int_0^\infty \frac{\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{3/2}} e^{-\tau x^2}\, d\tau, \\
M_e^{-1}\Phi^{-1}g(x) &= \frac{\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{3/2}}, \\
TM_e^{-1}\Phi^{-1}g(x) &= \frac{\alpha\left(\frac{1}{4\tau}\right)}{2(2\tau)^{1/2}}, \\
M_eTM_e^{-1}\Phi^{-1}g(x) &= \int_0^\infty \frac{\alpha\left(\frac{1}{4\tau}\right)}{2(2\tau)^{1/2}} e^{-\tau x^2}\, d\tau, \\
T_W^e g(x) = \Phi M_eTM_e^{-1}\Phi^{-1}g(x) &= \int_0^\infty \frac{\alpha(\tau)}{4\tau} e^{-\tau x^2}\, d\tau.
\end{aligned}$$

Now

$$(T_W^e g)'(x) = \int_0^\infty \frac{\alpha(\tau)}{4\tau}(-2\tau x)e^{-\tau x^2}\, d\tau = -\frac{x}{2}g(x).$$

Therefore

$$T_W^e g(x) = \frac{1}{2}\int_{|x|}^{+\infty} yg(y).$$

For odd functions (9.21) we have

$$\begin{aligned}
\Phi^{-1}h(x) &= ix\int_0^\infty \frac{\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{1/2}} e^{-\tau x^2}\, d\tau, \\
M_o^{-1}\Phi^{-1}h(x) &= \frac{i\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{1/2}}, \\
TM_o^{-1}\Phi^{-1}h(x) &= i\sqrt{\frac{\tau}{2}}\alpha\left(\frac{1}{4\tau}\right), \\
M_oTM_o^{-1}\Phi^{-1}h(x) &= ix\int_0^\infty \sqrt{\frac{\tau}{2}}\alpha\left(\frac{1}{4\tau}\right) e^{-\tau x^2}\, d\tau, \\
T_W^o h(x) = \Phi M_oTM_o^{-1}\Phi^{-1}h(x) &= x\int_0^\infty \frac{\alpha(\tau)}{4\tau} e^{-\tau x^2}\, d\tau.
\end{aligned}$$

Now

$$\left(\frac{T_W^o h}{x}\right) = \int_0^\infty \frac{\alpha(\tau)}{4\tau}(-2\tau x)e^{-\tau x^2}\, d\tau = -\frac{1}{2}h(x).$$

Therefore

$$T_W^o h(x) = \frac{x}{2}\int_{|x|}^{+\infty} h(y).$$

The theorem is proved.

Corollary 9.1 *Combining the formulae (9.18) and (9.19) for T_W^e and T_W^e we obtain the time operator $T_W = T_W^o P_o + T_W^e P_e$ for the entire semigroup W_t:*

$$T_W f(x) = \int_{|x|}^\infty \frac{(y+x)f(y) + (y-x)f(-y)}{4}\, dy. \tag{9.22}$$

9.4 THE SPECTRAL RESOLUTION OF THE TIME OPERATOR

The spectral resolution of the time operator (9.22) is obtained from the spectral resolutions of the restrictions (9.18) and (9.19) onto the spaces of even and odd functions.

Theorem 9.2 *The time operators T_W^o and T_W^e (9.18) and (9.19) have spectral resolutions:*

$$\begin{aligned} T_W^e &= \int_0^\infty \tau \, dP^e(\tau), \\ T_W^o &= \int_0^\infty \tau \, dP^o(\tau). \end{aligned}$$

where $P^e(\tau)$ and $P^o(\tau)$ are (nonorthogonal) projections given by the formulae

$$P^e(\tau)g(x) = \int_{1/4\tau}^\infty \alpha(\tau')e^{-\tau' x^2} \, d\tau', \quad \text{where } g(x) = \int_0^\infty \alpha(\tau')e^{-\tau' x^2} \, d\tau',$$

$$P^o(\tau)h(x) = x \int_{1/4\tau}^\infty \alpha(\tau')e^{-\tau' x^2} \, d\tau', \quad \text{where } h(x) = x \int_0^\infty \alpha(\tau')e^{-\tau' x^2} \, d\tau'.$$

Proof. The spectral resolution of the self-adjoint time operator of the right shift V_t is:

$$T\alpha(\tau') = \int_0^\infty \tau \, dP(\tau)\alpha(\tau'),$$

where $P(\tau)\alpha(\tau') = \alpha(\tau')\theta(\tau' - \tau)$, θ is the Heaviside function. By (9.3) for $g \in D_{T_W^e}$ we obtain

$$T_W^e g = \Lambda_e T \Lambda_e^{-1} g = \Lambda \int_0^\infty \tau \, dP(\tau)\Lambda_e^{-1} g = \int_0^\infty \tau \, d\Lambda_e P(\tau)\Lambda_e^{-1} g$$

Let $P_e(\tau) = \Lambda_e P(\tau)\Lambda_e^{-1}$. Then we obtain the spectral resolution formula

$$T_W^e = \int_0^\infty \tau \, dP^e(\tau).$$

Analogously,

$$T_W^o = \int_0^\infty \tau \, dP^o(\tau),$$

where $P_o(\tau) = \Lambda_o P(\tau)\Lambda_o^{-1}$. It remains to calculate the explicit formulae for $P^e(\tau)$ and $P^o(\tau)$. It suffices to check these formulae for some dense linear subspaces of $\mathcal{H}^e$

and $\mathcal{H}^o$. Therefore it suffices to prove the first formula for the functions of the form (9.20). A straightforward calculation gives

$$\Phi^{-1}g(x) = \int_0^\infty \frac{\alpha\left(\frac{1}{4\tau'}\right)}{(2\tau')^{3/2}} e^{-\tau' x^2}\, d\tau',$$

$$M_e^{-1}\Phi^{-1}g(x) = \frac{\alpha\left(\frac{1}{4\tau'}\right)}{(2\tau')^{3/2}},$$

$$P(\tau)M_e^{-1}\Phi^{-1}g(x) = \begin{cases} \dfrac{\alpha\left(\frac{1}{4\tau'}\right)}{(2\tau')^{3/2}} \text{ if } \tau' \leq \tau, \\ 0 \text{ if } \tau' > \tau, \end{cases}$$

$$M_e P(\tau)M_e^{-1}\Phi^{-1}g(x) = \int_0^\tau \frac{\alpha\left(\frac{1}{4\tau'}\right)}{(2\tau')^{3/2}} e^{-\tau' x^2}\, d\tau',$$

$$P^e(\tau)g(x) = \Phi M_e P(\tau) M_e^{-1}\Phi^{-1}g(x) = \int_{\frac{1}{4\tau}}^\infty \alpha(\tau')e^{-\tau' x^2}\, d\tau'.$$

The first formula is proved. The second can be proved analogously.

9.5 AGE EIGENFUNCTIONS AND SHIFT REPRESENTATION OF THE SOLUTION OF THE DIFFUSION EQUATION

We see directly from the expressions (9.18) and (9.19) that the functions

$$\chi_e^\tau(x) = \frac{1}{\sqrt{\tau}} e^{-\frac{x^2}{4\tau}}\,, \tag{9.23}$$

$$\chi_0^\tau(x) = \frac{x}{\sqrt{\tau}} e^{-\frac{x^2}{4\tau}} \tag{9.24}$$

are age eigenfunctions of T_W:

$$T_W \chi_e^\tau(x) = \tau \chi_e^\tau(x)$$

$$T_W \chi_o^\tau(x) = \tau \chi_o^\tau(x).$$

The Time operator T_W has spectral multiplicity two. We plot the age eigenfunctions (9.23),(9.24) in figs. 1 and 2.

Fig.1

Fig.2

The semigroup W_t acts as a shift on the age eigenfunctions:

$$W_t \chi_e^\tau(x) = \chi_e^{\tau+t}(x) \ \ and \ \ W_t \chi_o^\tau(x) = \chi_o^{\tau+t}(x).$$

If $\varphi \in L^2(\mathbb{R})$ is represented in terms of the age eigenfunctions as

$$\varphi(x) = \int_0^\infty \frac{\varphi_\tau^e}{\sqrt{\tau}} e^{-\frac{x^2}{4\tau}}\, d\tau + \int_0^\infty \frac{x\varphi_\tau^o}{\sqrt{\tau}} e^{-\frac{x^2}{4\tau}}\, d\tau = \int_0^\infty \frac{\varphi_\tau^e + x\varphi_\tau^o}{\sqrt{\tau}} e^{-\frac{x^2}{4\tau}}\, d\tau$$

then the semigroup W_t of the diffusion equation acts as follows:

$$W_t\varphi(x) = \int_0^\infty \frac{\varphi_\tau^e + x\varphi_\tau^o}{\sqrt{\tau + t}} e^{-\frac{x^2}{4(\tau+t)}}\, d\tau\,. \tag{9.25}$$

The coordinates φ_τ^e and φ_τ^o of φ can be expressed through the inverse Laplace transform of the even and odd parts of φ up to the change of variables $x \mapsto x^2$.

Formula (9.25) provides a new representation of the solution of the diffusion equation.

We remark before ending this section that the age eigenfunctions (9.23) and (9.24) can be also obtained from the age eigenstates of the time operator T (9.18) of the unilateral shift V_t (9.12) via the transformations Λ_e and Λ_o (9.13) which can be continuously extended to Schwartz distributions:

$$\chi_e^\tau = \Lambda_e \delta_\tau,$$

$$\chi_o^\tau = \Lambda_o \delta_\tau.$$

δ_τ, $\tau \in [0, \infty)$ are the Dirac δ-functions concentrated at τ, being the eigenfunctions of T (9.18).

9.6 TIME OPERATOR OF THE TELEGRAPHIST EQUATION

The operator

$$T \begin{pmatrix} \psi(x) \\ \dot{\psi}(x) \end{pmatrix} = - \begin{pmatrix} \sum_{a=1}^{3} x^a \frac{\partial}{\partial x^a} \Delta^{-1} \dot{\psi}(x) \\ \psi(x) + \sum_{a=1}^{3} x^a \frac{\partial}{\partial x^a} \psi(x) \end{pmatrix} \tag{9.26}$$

was shown [???] to be a self-adjoint time operator for the group V_t, associated to the wave equation

$$\partial_t \begin{pmatrix} \psi(x) \\ \dot{\psi}(x) \end{pmatrix} = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} \psi(x) \\ \dot{\psi}(x) \end{pmatrix} \tag{9.27}$$

in the Hilbert space of square-integrable initial data with finite energy.

It was also shown in [???] that the transformation

$$\Lambda(T) = \int e^{-\gamma\tau} \, dP_\tau$$

where P_τ is the spectral resolution of the time operator (9.26):

$$T = \int \tau \, dP_\tau$$

intertwines the wave equation (9.27) with the Poincare–Telegraphist equation, for $t \geq 0$:

$$\frac{\partial^2 u}{\partial t^2} = \Delta u - 2\gamma \frac{\partial u}{\partial t} - \gamma^2 u \, . \tag{9.28}$$

Since the operators Λ and T commute, the operator (9.26) is also a time operator for the semigroup generated by the Poincare–Telegraphist equation, which is a dissipative hyperbolic equation describing diffusion propagating with finite velocity [Kelly].

9.7 NONLOCAL ENTROPY OPERATOR FOR THE DIFFUSION EQUATION

The norm $||W_t\varphi||^2$ for every initial state φ, defines a local entropy for the diffusion process. However, Misra, Prigogine, and Courbage defined a nonlocal Entropy operator as a decreasing operator-function of the Time operator [Mi,Pr,MPC]. We may define also in the same way a nonlocal Entropy operator for the diffusion equation:

$$\mathbf{M} = \mathbf{M}(T_W) = \int_{-\infty}^{\infty} \mathbf{M}(\tau) d(P^e(\tau) \oplus P^o(\tau)) \ . \tag{9.29}$$

For the simple profile function

$$\mathbf{M}(\tau) = e^{-2\gamma\tau} \ , \quad \gamma > 0 \ . \tag{9.30}$$

we have:

Proposition 9.1 *The Entropy operator (9.29), (9.30) acts on the functions $f(x)$ in the domain of T_W as:*

$$\mathbf{M}f(x) = \frac{i}{4}\sqrt{\frac{\gamma}{2}} \int_{\lambda - i\infty}^{\lambda + i\infty} dz \frac{H_1^{(1)}(i\sqrt{2\gamma z})}{\sqrt{z(x^2 - z)}} \Big[(\sqrt{x^2 - z} + 2x) f(\sqrt{x^2 - z}) + $$

$$+ (\sqrt{x^2 - z} - 2x) f(-\sqrt{x^2 - z}) \Big] \ , \tag{9.31}$$

where $\lambda > 0$ is an arbitrary positive constant, $f(\pm\sqrt{x^2 - z})$ stands for the analytic continuation of the function $f(x)$ to the complex plane, the branch of the square root is *chosen so that real part of $\sqrt{z}$ is nonnegative and $H_1^{(1)}(\zeta) = -\frac{1}{\pi}\int_{-\infty}^{\infty} e^{i\zeta\cosh t - t}dt$ is the Hankel function of the first type.*

Proof. Following the proof of Theorem 1, we consider separately the action of the entropy operator on even and odd functions. For the even function $g(x)$ (9.20) we have:

$$M_e^{-1}\Phi^{-1}g(x) = \frac{\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{3/2}}$$

$$M_e e^{-2\gamma\tau} M_e^{-1}\Phi^{-1}g(x) = \int_0^\infty \frac{\alpha\left(\frac{1}{4\tau}\right)}{(2\tau)^{3/2}} e^{-2\gamma\tau} e^{-\tau x^2} d\tau$$

$$\mathbf{M}^e g(x) = \Phi M_e e^{-2\gamma\tau} M_e^{-1}\Phi^{-1}g(x) = \int_0^\infty \alpha(\tau) e^{-\frac{\gamma}{2\tau}} e^{-\tau x^2} dx \ .$$

The latter integral is the Laplace transform of the product of the functions $\alpha(\tau)$ and $e^{-\frac{\gamma}{2\tau}}$ with respect to the variable $z = x^2$,

$$\mathbf{M}^e g(x) = \mathbf{M}^e g(\sqrt{z}) = \mathcal{L}[\alpha(\tau)e^{-\frac{\gamma}{2\tau}}] \ .$$

Using Mellin's formula we get

$$\mathbf{M}^e g(x) = \frac{1}{2i\pi}\int_{\lambda-i\infty}^{\lambda+i\infty} \mathcal{L}[\alpha(\tau)](x^2 - z)\mathcal{L}[e^{-\frac{\gamma}{2\tau}}](z)dz \ ,$$

where $\lambda > 0$ is an arbitrary positive constant. From Eq.(9.20) we have

$$\mathcal{L}[\alpha](z) = \int_0^\infty \alpha(\tau)e^{-\tau z}d\tau = g(\sqrt{z}) \ .$$

From [GradRyz] we have:

$$\mathcal{L}[e^{-\frac{\gamma}{2\tau}}](z) = \int_0^\infty e^{-\frac{\gamma}{2\tau} - \tau z} dz = -\pi\sqrt{\frac{\gamma}{2}}\frac{1}{\sqrt{z}} H_1^{(1)}(i\sqrt{2\gamma z}) \ ,$$

Therefore

$$\mathbf{M}^e g(x) = \frac{i}{2}\sqrt{\frac{\gamma}{2}}\int_{\lambda-i\infty}^{\lambda^+ i\infty} g(\sqrt{x^2 - z}) H_1^{(1)}(i\sqrt{2\gamma z})\frac{dz}{\sqrt{z}} \ . \qquad (9.32)$$

For the odd functions $h(x)$ (9.21) we have

$$M_o^{-1}\Phi^{-1}h(x) = \frac{i\left(\frac{1}{4\tau}\right)}{\sqrt{2\tau}}$$

$$M_o e^{-2\gamma\tau} M_o^{-1}\Phi^{-1}h(x) = ix\int_0^\infty \alpha\left(\frac{1}{4\tau}\right)\frac{1}{\sqrt{2\tau}} e^{-2\gamma\tau} e^{-\tau x^2} d\tau$$

$$\mathbf{M}^o h(x) = \Phi M_o e^{-2\gamma\tau} M_o^{-1}\Phi^{-1}h(x) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty \alpha\left(\frac{1}{4\tau}\right) \frac{1}{\sqrt{2\tau}} e^{-2\gamma\tau} \int_{-\infty}^\infty iye^{-\tau y^2 - ixy} dy d\tau =$$

$$= 2x \int_0^\infty \alpha(\tau) e^{-\frac{\gamma}{2\tau}} e^{-\tau x^2} d\tau \ .$$

From Eq.(9.21) we have

$$\mathbf{M}^o h(x) = 2x\mathcal{L}[\alpha(\tau)e^{-\frac{\gamma}{2\tau}}] = \frac{1}{i\pi} \int_{\lambda - i\infty}^{\lambda + i\infty} \mathcal{L}[\alpha](x^2 - z)\mathcal{L}[e^{-\frac{\gamma}{2\tau}}](z) dz =$$

$$= x\sqrt{\frac{\gamma}{2}} \int_{\lambda - i\infty}^{\lambda + i\infty} \frac{h(x^2 - z)}{\sqrt{x^2 - z}} H_1^{(1)}(i\sqrt{2\gamma z}) \frac{dz}{\sqrt{z}} \ . \tag{9.33}$$

Separating the even and odd components of a function $f(x)$

$$f(x) = P_e f(x) + P_o f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

we calculate the Entropy operator

$$\mathbf{M} f(x) = \mathbf{M}^e P_e f(x) + \mathbf{M}_o P_o f(x) \ .$$

From Eqs.(9.32) and (9.33) we obtain Eq.(9.31).

10

Time operator of self-similar processes

Let $X = \{X_t\}_{t\geq 0}$ be a real valued stochastic process. X is *self-similar* if there exist a constant $H > 0$ such that $\{X_{at}\} \stackrel{d}{=} \{a^H X_t\}$, for each $a > 0$, where $\stackrel{d}{=}$ denotes the equality of all finite dimensional distributions. We call this X an *H-ss process*. The constant H is called the index of self-similarity or Hurst exponent.

The Brownian motion is a well known example of a self-similar process with the index $H = 1/2$. Another example of a self-similar process with the index $1/2 < H < \infty$ is the symmetric-α-stable (SαS) Lévy motion, where $H = 1/\alpha$ [ST].

Recall that the process $X = \{X_t\}_{t\geq 0}$ is the SαS Lévy motion, $0 < \alpha \leq 2$, if it has independent and stationary increments, $X(0) = 0$ a.e. and for each $n = 1, 2, \ldots$, $a_1, \ldots, a_n \in \mathbb{R}$ the random variable $a_1 X_1 + \ldots + a_n X_n$ is symmetric α-stable, i.e. with the characteristic function $t \mapsto e^{-\theta^\alpha |t|^\alpha}$, for some $\theta > 0$. Recall also that stationarity of increments is in the strong sense. This means that

$$\{X_{t+s} - X_s\} \stackrel{d}{=} \{X_t - X_0\},$$

for all s.

The basic tool that allows to associate a time operator with a self-similar process is the following proposition

Proposition 10.1 (Lamperti – see [ST]) *If the process $\{X_t\}_{t>0}$ is H-ss then the process*

$$Y_t \stackrel{\text{df}}{=} e^{-tH} X_{e^t}, \quad -\infty < t < \infty, \tag{10.1}$$

is stationary. Conversely if $\{Y_t\}_{t\in\mathbb{R}}$ is stationary, then

$$X_t \stackrel{\text{df}}{=} t^H Y_{\ln t}, \quad t > 0, \tag{10.2}$$

is H-ss.

The term ‘stationary’ used in this section means also stationary in the strong sense, i.e. for each $n \in \mathbb{N}$, $s, t_1, \ldots, t_N \in \mathbb{R}$, the distribution of the random vector

$$(X_{t_1+s}, \ldots, X_{t_n+s})$$

does not depends on s.

Since the above definition does not impose any condition on the moments of X, stationarity in the strong sense does not imply the stationarity in the wide sense (see Section 9), and neither conversely. However, if a stationary in the strong sense process has finite second moments, then it is also stationary in the wide sense and we can apply to it the results of Section 9.

For a given H-self-similar process $X = \{X_t\}$ we proceed with the construction of the time operator as follows. We correspond to X the stationary process Y through (10.1). If the second moments of X are finite then the process Y is also stationary in the wide sense and we can associate with Y through

$$Tf = \int t \, dE_t f$$

the operator T acting on the Hilbert space $\mathcal{H}(Y)$ spanned by Y_t, $t \in \mathbb{R}$, where the projectors E_t are the conditional expectations with respect to the σ-algebras $\mathcal{F}_t \stackrel{\mathrm{df}}{=} \sigma(Y_s)_{s \le t}$. T is a time operator with respect to the group $\{V_t\}_{t \in \mathbb{R}}$ of shifts of Y, $V_t Y_s = Y_{s+t}$.

Because of (10.2) $\mathcal{H}(Y) = \mathcal{H}(X)$, where $\mathcal{H}(X)$ is spanned by X_t, $t > 0$. We define on the space $\mathcal{H}(X)$ the operator

$$\widetilde{T} = \int_0^\infty s \, d\widetilde{E}_s \, , \tag{10.3}$$

where $\widetilde{E}_s \stackrel{\mathrm{df}}{=} E_{\ln s}$. It is easy to see that $\widetilde{T}$ is the exponential function of the operator T, $\widetilde{T} = \exp(T)$. Since T is a time operator with respect to the group $\{V_t\}_{t \in \mathbb{R}}$, it satisfies the relation

$$TV_t = V_t(T + tI) \, . \tag{10.4}$$

The group of shifts $\{V_t\}$ has the property that it generates the whole stationary process from just one random variable Y_0. We shall introduce below a similar group of transformation generating the self-similar process X and show an analog of the relation (10.4) with respect to the operator $\widetilde{T}$. Define

$$\widetilde{V}_t = t^H V_{\ln t} \, , \ \ t > 0 \, . \tag{10.5}$$

$\{\widetilde{V}_t\}_{t>0}$ is a multiplicative group on $\mathcal{H}(X)$. Moreover it generates the selfsimilar process X form the random variable $X(1)$. Indeed, observe that

$$\widetilde{V}_s X(t) = s^H t^H V_{\ln s} Y(\ln t) = (st)^H Y(\ln(st)) = X(st) = \widetilde{V}_t X(s) \, ,$$

for $s, t > 0$, which implies that $\{\widetilde{V}_t\}_{t>0}$ possesses group properties on $\mathcal{H}(X)$ with respect to multiplication on $\mathbb{R}_+$. In particular, putting $s = 1$ we obtain

$$X(t) = \widetilde{V}_t X(1)\,, \quad \text{for } t > 0\,.$$

It remains to show that $\widetilde{T}$ is a time operator with respect to $\{\widetilde{V}_t\}_{t>0}$. First, however let us find the analog of the relation (10.4) in the case of multiplicative semigroups. It is natural to expect that the multiplicative analog of (10.4), i.e. that the multiplicative group $\{\widetilde{V}_t\}$ keeps step with t is

$$\widetilde{T}\widetilde{V}_t = \widetilde{V}_t(\widetilde{T}\cdot tI)$$

or, equivalently,

$$\widetilde{T}\widetilde{V}_t = t\widetilde{V}_t\widetilde{T}\,. \tag{10.6}$$

Indeed, since $\widetilde{T} = \exp(T)$ we have

$$\widetilde{T}\widetilde{V}_t = \left(\int_{-\infty}^{\infty} e^s\, dE_s\right) t^H V_{\ln t} = t^H V_{\ln t}\int_{-\infty}^{\infty} e^{s+\ln t} dE_s = \widetilde{V}_t t\int_{-\infty}^{\infty} e^s dE_s = \widetilde{V}_t\widetilde{T}\,,$$

which proves (10.6).

Let us also notice the following relation between the operator T and the group $\{\widetilde{V}_t\}$:

$$T\widetilde{V}_t = \widetilde{V}_t T + \ln t\widetilde{V}_t\,,$$

and between $\widetilde{T}$ and the group $\{V_t\}$:

$$\widetilde{T}V_t = e^t V_t\widetilde{T}\,.$$

The following example illustrates how an application of Λ-operator can lead from short to long memory processes and conversely. Assume that the orthonormal basis $\{Y_n\}_{n\in\mathbb{Z}}$ in $\mathcal{H}(X)$ consists of independent identically distributed normal $\mathcal{N}(0,1)$ random variables. Let a_n, $n = 0, 1, 2, \ldots$ be positive numbers such that $\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} a_{n+k}a_k < \infty$. Since, in particular, $\{a_n\}$ is square summable the random variable

$$Z_0 \stackrel{\text{df}}{=} \sum_{k=-\infty}^{0} a_{-k}Y_k$$

is an element of $\mathcal{H}_0(X)$. Therefore the shifts of Z_0 define a gaussian process

$$Z_n \stackrel{\text{df}}{=} V^n Z_0$$

with the correlation function

$$C(n) \stackrel{\text{df}}{=} \langle Z_0, Z_n\rangle = \sum_{k=-\infty}^{n} a_{-k}a_{n-k}$$

(we put $a_k = 0$, for $k < 0$).

The above assumptions concerning $\{a_n\}$ imply that the correlation function is summable: $\sum_{n\in\mathbb{Z}} |C(n)| < \infty$. One can show easily that the central limit theorem for $\{Z_n\}$ holds, i.e. the sum $\frac{1}{\sqrt{N}} \sum_{n=1}^{N} Z_n$ converge to the normal $\mathcal{N}(0, \sigma^2)$ distribution with $\sigma^2 = \sum_{n\in\mathbb{Z}} C(n)$. This is, however, no longer true if we apply Λ-transformation to Z_0. Indeed, consider the process

$$\widetilde{Z}_n \stackrel{\mathrm{df}}{=} V^n \Lambda Z_0 .$$

Its correlation function $\widetilde{C}(n)$ has the form

$$\widetilde{C}(n) = \sum_{k=-\infty}^{n} \lambda_k \lambda_{n+k} a_{-k} a_{n-k} .$$

Taking $\{\lambda_k\}$ in such a way that the asymptotic behavior of $\widetilde{C}(n)$ is:

$$\widetilde{C}(n) \sim K/n^{\alpha} , \quad \text{as } n \to \infty ,$$

where K is a constant and $0 < \alpha < 1$, one can prove that a non-central limit theorem holds. Namely (see [ST], Th.7.2.11) that the averaged sums:

$$\frac{1}{N^H} \sum_{n=1}^{[Nt]} Z_n , \quad t \in [0,1] ,$$

converge (in the sense of finite dimensional distributions) to a fractional Brownian motion $B_H(t)$ with the self-similarity parameter (Hurst exponent) $H = 1 - \alpha/2$.

We have shown above how to associate the time operator to a stochastic process by describing its spectral resolution in terms of the filtration of the process. Although the connection of the filtration of a stochastic process with the flow of time is very natural, the explicit form of the filtration and corresponding eigenprojectors is rarely known in practice. Usually the knowledge of a stochastic process is limited. In the case of stationary (in the wide sense) process it is usually known only its covariance function or spectral density. In fact, the most relevant result, from the point of view of applications of stationary processes, concerns these two functions. Therefore, for practical purposes it is necessary to derive time operator basing only on such limited knowledge. We shall show below how to derive the eigenprojectors of the time operator associated with a stationary process knowing only the covariance function. In order to do this we have to recall the basic facts from the spectral analysis of stationary processes.

Denote by R the covariance function of the stationary process $\{X(n)\}_{n\in\mathbb{Z}}$:

$$R(n) \stackrel{\mathrm{df}}{=} \mathrm{Cov}[X(k+n), X(k)] = E(X(k+n) - m)\overline{(X(k) - m)} .$$

Assuming that $m = 0$ we have

$$R(n) = EX(k+n)\overline{X(k)} .$$

By Herglotz's Theorem [AG] there is a unique measure μ on the σ-algebra $\mathcal{B}_{[-\pi,\pi]}$ of the Borel subsets of $[-\pi,\pi]$ such that

$$R(n) = \int_{-\pi}^{\pi} e^{in\pi} \mu(d\lambda)\,.$$

The measure μ is the control measure of a unique orthogonal valued measure $M : \mathcal{B}_{[-\pi,\pi]} \to \mathcal{H}(X)$ such that

$$X(n) = \int_{-\pi}^{\pi} e^{in\lambda} M(d\lambda)\,.$$

This means that for each $Y \in \mathcal{H}(X)$ there is a function $f \in L^2(\mu)$ such that Y has the following spectral representation:

$$Y = \int_{-\pi}^{\pi} f(\lambda) M(d\lambda) \tag{10.7}$$

with

$$\left\| \int f\,dM \right\|^2 = \int |f|^2 d\mu\,.$$

In particular $\|M(A)\|^2 = \mu(A)$, where the norm $\|\cdot\|$ denote the L^2 norm. Note that in the spectral representation (10.7) the action of the shift V on $\mathcal{H}(X)$, $V^k X(n) = X(n+k)$, can be written explicitly

$$V^k Y = V^k \int_{-\pi}^{\pi} f(\lambda) M(d\lambda) = \int_{-\pi}^{\pi} e^{in\lambda} f(\lambda) M(d\lambda)\,.$$

11

Time operator of Markov processes

The term *Markov process* means a family of random variables on a probability space. An equivalent object is a *Markov semigroup* or simply a *Markov operator* if the time is discrete.

Our goal is to correspond to a given Markov process or, equivalently, to a Markov semigroup a time operator.

If the Markov process is given as a family of random variables $\{X_t\}$ an operator T can be associated with the family $\{\mathcal{F}_t\}$ as above. However, the natural evolution semigroup associated with $\{X_t\}$ is, the Markov semigroup $\{M_t\}$ which acts on a different space then $L^2(\mathcal{F}, P)$ (details below). Because of this we cannot directly relate T with the evolution semigroup, i.e. validate the relation

$$TM_t = M_t T + tM_t\,, \quad \text{for all } t\,. \tag{11.1}$$

We shall show below that such association is possible by the means of dilation theory and replacement of the original Markov process by the corresponding canonical process.

The original goal of dilation theory is to find a way to study general bounded operators through isometries or unitary operators (see [SzNF] for an excellent treatment of the subject). This is moreover a natural framework to study the question whether a Markov process may arise as a projection of deterministic dynamical system. The latter question is closely related to the long standing problem of the relationship between deterministic laws of dynamics and probabilistic description of physical processes, known as the *problem of irreversibility*.

In contrast to the usual coarse graining approach B. Misra, I. Prigogine and M. Courbage [MPC,Pr] formulated the problem of irreversibility taking as a fundamental physical principle the law of entropy increase. In this theory the problem of reconciling the dynamical evolution with the irreversible thermodynamical evolution is viewed in terms of establishing a non-unitary equivalence between the unitary dynamical group and probabilistic Markov processes. More explicitly, let $\{U_t\}$ be a unitary group of evolution, which is induced by point dynamics $\{S_t\}$ of the phase space, acting on the Hilbert space $\mathcal{H}$ spanned by the square integrable phase space functions. One considers now the possibility of relating the group $\{U_t\}$ with a Markovian semigroup $\{W_t\}$ through a non-unitary positivity preserving intertwining transformation Λ in such a way that

$$\Lambda U_t = W_t \Lambda\,, \quad \text{for } t \geq 0\,. \tag{11.2}$$

If the transformation Λ is invertible then the intertwining between the unitary dynamics and the Markov evolution involves no loss of information (non-unitary "equivalence"). If, on the other hand, Λ is an orthogonal projection on a subspace of $\mathcal{H}$, then the relation between $\{U_t\}$ and $\{W_t\}$ can be seen as a coarse-graining compatible with dynamics.

The conservative systems for which invertible transformation to dissipative systems have been constructed are qualified by the existence of internal time. An internal time operator for the unitary evolution evolution $\{U_t\}$ is a self-adjoint operator T satisfying the canonical commutation relation:

$$U_{-t} T U_t = T + tI\,, \quad \text{for each } t\,. \tag{11.3}$$

Internal time operators for unitary dynamics were introduced [Mi] in the context of unstable dynamical systems of the Kolmogorov type. Misra, Prigogine and Courbage showed that the unitary evolution $\{U_t\}$ of a Kolmogorov system can be intertwined (11.2) either by a similarity

$$W_t = \Lambda U_t \Lambda^{-1}$$

or by a coarse-graining projection $\Lambda = P$,

$$W_t = P U_t P\,,$$

with the Markov semigroup $\{W_t\}$.

Since 1980's the Misra, Prigogine and Courbage theory of irreversibility has been developed further in several directions: classical dynamical systems [Co, EP, Ll, MP1, SW, ZS], relativistic systems [AMr], or quantum systems [LM], to mention just the main direction. Of course, the above list of references is far from being complete. It is therefore natural that also new problems have been raised. One of them is the problem of implementability of the resulting Markov semigroup, i.e. the question whether the semigroup $\{W_t\}$ is implementable by a (noninvertible) point transformation of the phase space. Another problem is to characterize the class of dynamical systems, both classical and quantum, which can be intertwined with dissipative systems. Yet another problem, which is in a sense converse to the latter, is to characterize the Markov semigroups that can be intertwined with unitary groups or which can arise as their projections.

The first question was raised because the formulation of invertible dynamics in terms of point transformations of the phase space is equivalent to the formulation in terms of unitary measure preserving groups. The question of implementability of $\{W_t\}$ is then about whether the irreversibility can be observed on the level of trajectories. Although all unitary and positivity preserving transformations on L^2 (isometries on $L^p, p \neq 2$) are implementable [GGM,La] the semigroup $\{W_t\}$ related with $\{U_t\}$ through a similarity transformation Λ is, in general, not implementable [GMC,AG,SAT] (see also [AMS] for quantum dynamics).

Concerning the second problem, it is known that K-systems, both classical and quantum, admit time operators thus can be intertwined with dissipative systems. It is also known that if a unitary dynamics admits a time operator then it has the ergodic property of mixing.

The third problem has been so far only partially solved. One of the reasons of the difficulties to resolve this problem is the above mentioned negative answer to the problem of implementability. A unitary group of positivity preserving operators which acts on the Hilbert space spanned by phase space functions is implemented by some transformations of the phase space. But the induced Markov semigroup is in general not related to the underlying phase space and its dynamics. It should be noticed that in all known constructions of the similarity between the unitary and Markov semigroups the knowledge of the underlying dynamics plays a crucial role. It is in fact crucial from the point of view of the possibility of constructing a time operator which relates both semigroups. Therefore the third question is actually about the ability of relating a Markov semigroup with some phase space dynamics and with the existence of a time operator. We shall show how to resolve this problem by relating Markov semigroups with canonical Markov processes acting on larger spaces. Namely that the original Markov semigroup can be dilated to some unitary group which also admits a time operator. Thus the concept of time operator goes beyond the unitary dynamics where it was initially introduced (see also [AStime,ASS,APSS]). In view of the above mention answer on the first problem, that Markov semigroups are not connected with the underlying dynamics on the phase space, the dilation approach to the problem of associating time operators with Markov semigroups, appears to be not only justified but the only possible, which guarantees a general solution.

11.1 MARKOV PROCESSES AND MARKOV SEMIGROUPS

The purpose of this section is to recall the correspondence between Markov processes and semigroups. Let $(\Omega, \mathcal{F}, P)$ be a probability space and X_t an $\mathcal{X}$-valued stochastic process ($\mathcal{X} \subset \mathbb{R}^n$) on Ω. X_t is a *Markov process* if for any time instants $t_1 < \ldots < t_n < t_{n+1}$, states $x_1, \ldots, x_n, x_{n+1} \in \mathcal{X}$ and a Borel set $B \subset \mathcal{X}$ we have

$$P\{X_{t_{n+1}} \in B | X_{t_1} = x_1, \ldots, X_{t_n} = x_n\} = P\{X_{t_{n+1}} \in B | X_{t_n} = x_n\}\,. \quad (11.4)$$

Because of (11.4) all finite dimensional distributions

$$P\{X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\}\,,$$

where $B_1, ..., B_n$ are Borel subset of $\mathcal{X}$, are determined by the transition probabilities $Q_{s,t}$:

$$Q_{s,t}(x, B) = P\{X_t \in B | X_s = x\}\,,$$

for $s < t$. Assuming that the transition probabilities are *stationary*, i.e. $Q_{s,t}(x, B)$ depend only on the difference $t - s$ and that the set of indices t is the interval $[0, \infty)$, we can consider instead of $Q_{s,t}(x, B)$ the transition probabilities:

$$Q_t(x, B) = P\{X_t \in B | X_0 = x\}, \tag{11.5}$$

for all $t > 0$. Recall that the transition probabilities have the following properties:

1) For each t and fixed B, $Q_t(x, B)$ is measurable as a function of x

2) For each t and fixed x, $Q_t(x, .)$ is a probability measure on Borel subsets of $\mathcal{X}$

3) For each s, t

$$Q_{s+t}(x, B) = \int\limits_{\mathcal{X}} Q_t(y, B) Q_s(x, dy)$$

(Chapman-Kolmogorov equation).

B by a theorem Ionescu Tulcea [Ne] an arbitrary family of transition probabilities, i.e. any family of functions $Q_t(x, B)$ satisfying conditions 1)-3), determines a Markov process. On the other hand suppose that the set $\mathcal{X}$ is equipped with the measure structure, i.e. with a σ-algebra Σ of subsets of $\mathcal{X}$ and a measure μ. Then each transition probability $Q_t(x, B)$ determines an operator M_t on $L^2_{\mathcal{X}}$ by putting

$$M_t(\mathbb{1}_B) = \int_{\mathcal{X}} Q_t(x, B)\mu(dx)\,,$$

and then extending M_t to a bounded and positivity preserving linear operator on $L^2_{\mathcal{X}}$. The family $\{M_t\}_{t\geq 0}$, which forms a semigroup of Markov operators on $L^2_{\mathcal{X}}$, is called a *Markov semigroup*.

The term *Markov operator* means, in general (see [Fog]), an arbitrary positivity preserving contraction on L^1. Here, however, by the Markov operator we mean a linear operator M on the space $L^2_{\mathcal{X}}$ (or on any of the spaces $L^p_{\mathcal{X}}$, for $p \geq 1$) with the following properties:

(a) M is a contraction

$$\|Mf\| \leq \|f\|$$

(b) M preserves positivity

$$Mf \geq 0 \quad \text{if} \quad f \geq 0$$

(c) M preserves probability normalization

$$\int_{\mathcal{X}} Mf d\mu = \int_{\mathcal{X}} f d\mu\,.$$

If, in addition, the measure μ on $(\mathcal{X}, \Sigma)$ is normalized then

(d) M preserve constants

$$M1 = 1\,.$$

The properties (a)–(d) characterize Markov process. Namely (see [Ne]), each operator M_t with properties (a)-(d) determines the transition probability

$$Q_t(x, B) \stackrel{\mathrm{df}}{=} M_t^* \mathbb{1}_B(x) \tag{11.6}$$

where M_t^* is the M_t adjoint on the space $L^2_{\mathcal{X}}$, and the transition probabilities define in turns a stochastic process, called the *canonical Markov process*.

11.2 CANONICAL PROCESS

Let us consider the infinite product $\widetilde{\mathcal{X}} = \prod_{t \in I} \mathcal{X}_t$, where I is the set of indices of the stochastic process $\{X_t\}$ and, for each $t \in I, \mathcal{X}_t = \mathcal{X}$, where $\mathcal{X}$ is the space of states of the process $\{X_t\}$. Let us correspond to each finite set of indices $\{t_1, ..., t_n\}$ the σ-algebra $\mathcal{C}_{t_1,...,t_n}$ generated by the cylinders

$$\mathcal{C}^{t_1,...,t_n}_{B_1,...,B_n} \stackrel{\mathrm{df}}{=} \{\tilde{x} \in \widetilde{\mathcal{X}} : \tilde{x}(t_1) \in B_1, ..., \tilde{x}(t_n) \in B_n\}, \tag{11.7}$$

where $B_1, ..., B_n$ runs through all Borel subsets of $\mathcal{X}$. By $\mathcal{C}$ we shall denote the σ-algebra generated by cylinders (11.7) for all possible choices of finite subsets $\{t_1, ..., t_n\}$. The stochastic process $\{X_t\}$ determines a probability measure μ on $(\widetilde{\mathcal{X}}, \mathcal{C})$ such that the restriction $\mu_{t_1,...,t_n}$ of μ to $\mathcal{C}_{t_1,...,t_n}$ coincides with the finite dimensional distributions:

$$\Pr\{X_{t_1} \in B_1, ..., X_{t_n} \in B_n\},$$

where $B_1, ..., B_n$ are Borel subset of $\mathcal{X}$. It is easy to see that the family of measures $\mu_{t_1,...,t_n}$ is consistent, i.e. if $\{s_1, ..., s_n\} \subset \{t_1, ..., t_n\}$ then the restriction of $\mu_{t_1,...,t_n}$ to $\mathcal{C}_{s_1,...,s_k}$ coincides with $\mu_{s_1,...,s_k}$. Conversely, the well known Kolmogorov's extension theorem says that any given consistent family of probability measures $\mu_{t_1,...,t_n}$ determines the unique probability measure μ on $(\widetilde{\mathcal{X}}, \mathcal{C})$, which restricted to $\mathcal{C}_{t_1,...,t_n}$ coincides with $\mu_{t_1,...,t_n}$. With the probability space $(\widetilde{\mathcal{X}}, \mathcal{C}, \mu)$ it is associated a *canonical process*

$$X_t(\tilde{x}) \stackrel{\mathrm{df}}{=} \tilde{x}(t),$$

for each $\tilde{x} \in \widetilde{\mathcal{X}}$ and $t \in I$, i.e. the random variable X_t is the projection on the t-coordinate. In the theory of stochastic processes it is often assumed that only the knowledge of all finite dimensional distributions is relevant. This is also the case in this section. Therefore, as it follows from the Kolmogorov extension theorem, the original stochastic process and the corresponding canonical process carry the same information. For this reason we shall assume in the sequel that the probability space $(\Omega, \mathcal{F}, P)$ is $(\widetilde{\mathcal{X}}, \mathcal{C}, \mu)$ and the considered process is the canonical process. This assumption will allow to relate to the same probability space the stochastic process and the corresponding Markov semigroup that we discuss below.

11.3 TIME OPERATORS ASSOCIATED WITH MARKOV PROCESSES

The problem of the association of time operators with Markov semigroups can be seen a part of a more general problem of the association of time operators with stochastic processes. The most natural way to associate a time operator with a stochastic process $\{X_t\}$ would be to associate it with the filtration, i.e. the increasing family of σ-algebras $\{\mathcal{F}_t\}$, where $\mathcal{F}_t = \sigma\{X_s : s \leq t\}$, through (8.3) (in such case E_t would be the conditional expectations $E(\cdot|\mathcal{F}_t)$). Such correspondence can be established for any stochastic process. However, the time operator must be, by the definition, related to the time evolution of the stochastic processes expressed in terms of a semigroup of operators acting on a Hilbert space. Such dynamical semigroups (groups) can be specified for some classes of stochastic processes. If, for example, a stochastic process is stationary in the wide sense then it defines a dynamical group of unitary operators (shifts of the time) acting on the Hilbert space generated by the realizations of the stochastic process. As we shall see below Markov semigroups can also be extended (dilated) to unitary groups associated with stationary processes and, consequently, a time operator can be associated with such stationary extension.

Recall that a semigroup of isometries $\{\tilde{M}_t\}$ on a Banach space $\tilde{X}$ is a dilation of the contractive semigroup $\{M_t\}$ on the Banach space X if there exist two linear operators $I : X \longrightarrow \tilde{X}$ and $E : \tilde{X} \longrightarrow X$ such that

$$M_t = E\tilde{M}_t I \ , \quad \text{for each } t \, . \tag{11.8}$$

In other words the following diagram commutes

$$\begin{array}{ccc} X & \xrightarrow{M_t} & X \\ I \downarrow & & \uparrow E \\ \tilde{X} & \xrightarrow{\tilde{M}_t} & \tilde{X} \end{array}$$

Usually, a dilation is realized in such a way that X is isomorphic to a subspace of $\tilde{X}$, I is the canonical injection of X into $\tilde{X}$, and E is the projection of $\tilde{X}$ onto X. If X is a Hilbert space and $\tilde{M}_t$ are unitary operator we may speak about unitary dilations of contractions.

Let us assume for the sake of the clarity of presentation that the time is discrete, $t = 0, 1, 2, \ldots$. This assumption allows to consider a single Markov operator M instead of the whole Markov semigroup because $M_t = M^t$. The case of continuous time can be however treated in the same way. Our aim is to construct a positive dilation $\tilde{M}$ of M which is implemented by a measure preserving transformation S of some probability space $(\Omega, \mathcal{F}, P)$.

Dilations of contractive semigroups such as the constructed in Ref. [SzNF] cannot be applied in our case. We are interested in more specific dilations of the Markov semigroup $\{M^n\}$, where M satisfies (a)–(d), acting on the Hilbert space $X = L^2_{\mathcal{X}}$ to a group acting on the larger space $\tilde{X} = L^2_{\Omega}$ which corresponds to a stationary process. In particular we are looking for such dilation that both $\tilde{M}^n, \ n = 0, \pm 1, \pm 2, ...$ and E are positivity preserving.

Faced with the difficulty to obtain meaningful positive dilations of Markov semigroups using the Sz-Nagy dilation theory, we constructed [AGS] a positive dilation based on the natural extensions of dynamical systems through canonical Markov processes. There the positivity is clearly related to that of the inducing semigroup. It is now our task to apply the dilation approach to an arbitrary Markov Semigroup in order to associate it with a time operator. Let us consider first the case of a single Markov operator M on $L^2_{\mathcal{X}}$ that satisfies (a)–(d).

Suppose then it is given a Markov operator M on the space $L^2_{\mathcal{X}}$. Recall that M corresponds to a transition probability $Q(x, A)$ through (11.6). Define the product space

$$\Omega = \prod_{n=-\infty}^{\infty} \mathcal{X}_n \ , \ \text{ where } \mathcal{X}_n = \mathcal{X} \ ,$$

and the σ-algebra $\mathcal{F}$ generated by the cylindrical sets

$$C_{A_1,...,A_k} = ... \times \Omega \times A_1 \times A_2 \times ... \times A_k \times \Omega \times ... \ , \tag{11.9}$$

where $k \in \mathbf{N}, \ A_i \in \mathcal{A}$.

On $(\mathcal{X}, \mathcal{F})$ we define the cylindrical measure P

$$P(C_{A_1,...,A_k}) = \int_{A_1} \int_{A_2} ... \int_{A_k} Q(x_{k-1}, dx_k)...Q(x_1, dx_2) \ \mu(dx_1) \ , \tag{11.10}$$

and show [AGS] that P is correctly defined and normalized. Thus using the Kolmogorov theorem [Ne] we extend P to a probability measure on the whole σ-algebra $\mathcal{F}$.

Denote by S the left shift on Ω, i.e.

$$S\omega = S(x_n)_{-\infty}^{\infty} \overset{\text{def}}{=} (x_{n+1})_{-\infty}^{\infty} \ , \ \omega \in \Omega \ . \tag{11.11}$$

It can be shown [AGS] that P is invariant with respect to S, which amounts to showing that to check that if either $A_0 = \Omega$ or $A_{k+1} = \Omega$ then

$$P(C_{A_0,...,A_k}) = P(C_{A_1,...,A_{k+1}}) = P(C_{A_1,...,A_k}) \ . \tag{11.12}$$

Now let $\tilde{M}$ be the operator on L^2_{Ω} associated with S as follows:

$$\tilde{M}\tilde{f}(\omega) = \tilde{f}(S^{-1}\omega) \ , \quad \tilde{f} \in L^2_{\Omega} \ . \tag{11.13}$$

The operator $\tilde{M}$ is a dilation of M. Precisely, we show that

$$M^n = E\tilde{M}^n I \ , \ \ n \in \mathbf{N} \ , \tag{11.14}$$

where I is the canonical injection L^2_X into L^2_Ω:

$$L^2_{\mathcal{X}} \ni f \longmapsto If = \tilde{f} \in L^2_\Omega$$

defined as

$$(If)(..., x_{-1}, x_0, x_1, ...) \stackrel{\text{def}}{=} f(x_0) \tag{11.15}$$

and E is the conditional expectation with respect to the σ-algebra generated by the cylinders

$$C^0_A = \{\omega = (x_n)^\infty_{-\infty} \; : \; x_0 \in A\} \; , \quad A \in \mathcal{A} \, . \tag{11.16}$$

Consequently, for any $f \in L^2_{\mathcal{X}}$, the function $E\tilde{M^n}If$ depends only on the 0–coordinate and can be uniquely identified with a function on Ω. The complete proof of the above constructed dilation the reader can find in Ref. [AGS].

In order to construct the time operator $\tilde{T}$ associated with the dilation $\{\tilde{M}^n\}$ we can proceed as in the case of K-flows. We distinguish the σ-algebra $\mathcal{F}_0$ generated by those cylinders $C^{0,\ldots,k}_{A_0,\ldots,A_k}$ $(k = 0, 1, ... \; , \; A_0, ..., A_k \in \Sigma)$ for which the set A_0 is placed on the 0^{th} coordinate and note that the σ-algebra $S^{-n}\mathcal{F}_0$, $n \in$ **Z**, is the σ-algebra generated by $C^{n,\ldots,n+k}_{A_0,\ldots,A_k}$ with A_0 placed on the n^{th} coordinate. Then we see that

(i) $S\mathcal{F}_0 \supset \mathcal{F}_0$

(ii) $\sigma(\bigcup_{n=-\infty}^{\infty} S^n\mathcal{F}_0) = \mathcal{F}$.

However, the the third condition of K-flows:

(iii) $\bigcap_{n=-\infty}^{\infty} S^n\mathcal{F}_0$ is the trivial σ-algebra

is, in general not satisfied. The condition (iii) is satisfied, i.e. the dilated evolution is a K-flow, provided the Markov operator has additionally the following property [AGS]:

(e) M strongly converges to equilibrium

$$\|M^n f - 1\|_{L^p} \longrightarrow 0, \text{ as } n \to \infty \text{ for each probability density } f \in L^p \, .$$

The dynamical system with the Frobenius-Perron operator satisfying (e) is called the *exact system*.

The above consideration leads to the conclusion that Markov operators satisfying conditions (a)–(e) can be dilated to K-flows. The corresponding time operator can be then constructed by the means of the conditional expectations

$$E_n = E(\cdot|\mathcal{F}_n)$$

by putting

$$T = \sum_{n=-\infty}^{\infty} n(E_n - E_{n-1}) .$$

The case of continuous Markov semigroup $\{M_t\}_{t\geq 0}$ is treated analogously. As before we assume that each M_t corresponds through (11.5) to the transition probability $Q_t(x, A)$. Now the dilation space Ω will be the product

$$\Omega = \prod_{t\in\mathbf{R}} \mathcal{X}_t \,, \quad \text{where } \mathcal{X}_t = \Omega \,.$$

The σ-algebra $\mathcal{F}$ is generated by the cylinders

$$C^{t_0,\ldots,t_k}_{A_0,\ldots,A_k} \,, \quad \text{where } t_0, ..., t_k \in \mathbf{R} \,, \quad A_0, ..., A_k \in \Sigma \,.$$

The probability measure P on $(\Omega, \mathcal{F})$ and the distinguish σ-algebra $\mathcal{F}_0$ are also defined analogously (see [AGS] for details). Then we may define the group of shifts transformations of Ω

$$S_t\omega = x(\cdot + t) \,, \quad \text{for each function } \omega = x(\cdot) \in \Omega \,.$$

Finally, we define the operators

$$\tilde{M}_t f(\omega) = f(S_{-t}\omega) \,, \quad f \in L^2_\Omega \,,$$

and check in essentially the same way as before that $\tilde{M}_t$ are the dilations of M_t. The time operator T is now of the form (8.3) (or stochastic integral)

$$T = \int_{-\infty}^{\infty} t dE_t$$

where E_t are the conditional expectations $E_t(\cdot|S_t\mathcal{F}_0)$. In this way we have defined the time operator T on the Hilbert space L^2_Ω, but it can be defined as linear operators on any space $L^p(\Omega, \mathcal{F}, P)$, where $1 \leq p \leq \infty$.

It can be checked directly that T satisfies the relation

$$T\tilde{M}_t = \tilde{M}_t T + t\tilde{M}_t \,. \tag{11.17}$$

However, since T is the time operator associated with a K-system $(\Omega, \mathcal{F}, P; S)$ (or a K-flow) the relation (11.17) also follows the previous results [MPC,GMC,SW].

We may now ask the question whether the above constructed time operator can be projected on the space $L^2_\mathcal{X}$ giving rise to the time operator associated with the Markov semigroup $\{M_t\}$, i.e. whether the operator ETI satisfies (11.1) with respect to M_t. The answer to this question is however negative. To see this it is enough to notice that for any $f \in L^2_\mathcal{X}$ the injection If is the function depending only on the "zero-coordinate" but T places zero for the projection on this coordinate thus $ETIf = 0$. Therefore the time operator of the dilated semigroup is, like the resulting K-flow, generically associated with the extended space.

12

Time operator and approximation

The first connection, although indirect, of the time operator with the approximation theory has been obtained through wavelets [AnGu,AStime]. An arbitrary wavelet multiresolution analysis can be viewed as a K-system determining a time operator whose age eigenspaces are the wavelet detail subspaces. Conversely, in the case of the time operator for the Renyi map the eigenspaces of the time operator can be expanded from the unit interval to the real line giving the multiresolution analysis corresponding to the Haar wavelet. However, the connections of time operator with wavelets are much deeper than the above mentioned. As we shall see later time operator is in fact a straightforward generalization of multiresolution analysis.

In order to connect time operator with approximation it is necessary to go beyond Hilbert spaces. Recall that one of the most important vector spaces from the point of view of application is the Banach space $\mathcal{C}_{[a,b]}$ of continuous functions on an interval $[a, b]$. The space of continuous functions plays also a major role in the study of trajectories of stochastic processes.

Time operator can be, in principle, defined on a Banach space in the same way as on a Hilbert space. However its explicit construction is in general a non-trivial task. Having given a nested family of closed subspaces of a Hilbert space we can always construct a self-adjoint operator with spectral projectors onto those subspaces. This is not true in an arbitrary Banach space. The reason is that it is not always possible to construct an analog of orthogonal projectors on closed subspaces. Moreover, even if a self-adjoint operator with a given family of spectral projectors can be defined there appear additional problems associated with convergence of such expansion and with possible rescalings of the time operator.

For some dynamical systems associated with maps the time operator can be extended from the Hilbert space L^2 to the Banach space L^p. This can be achived by replacing the methods of spectral theory [MPC,GMC], by more efficient martingales methods. For example, for K-flows it is possible to extend the time operator from L^2 to L^1 including to its domain absolutely continuous measures on the the phase space [SuL1,Su]. Martingales methods can not be, however, applied for the space of continuous functions.

In this section we discuss connections of time operator with wavelets, especially those restricted to the interval $[0,1]$, and the corresponding multiresolutions analysis. We establish a link between the Shannon sampling theorem and the eigenprojectors of the time operator associated with the Shannon wavelet. We construct the time operator associated with the Faber-Schauder system on the space $\mathcal{C}_{[0,1]}$ and study its properties. Such time operator corresponds to the interpolation of continuous functions by polygonal lines. We give the explicit form of the eigenprojectors of this time operator and characterize the functions from its domain in terms of their modulus of continuity.

12.1 TIME OPERATOR IN FUNCTION SPACES

Let V and T be two linear operators on the Banach space $\mathcal{B}$ such that V is bounded and T is densely defined. We shall say that T is an *(internal) time operator* on $\mathcal{B}$ associated with V if V preserves the domain of T, i.e. $V(D(T)) \subset D(T)$, and

$$TV^k = V^kT + kV^k\,, \ \text{ for } k \in \mathcal{I}\,, \tag{12.1}$$

where $\mathcal{I}$ is either the set $\mathbb{Z}$ of integers or the set $\mathbb{N}$ of natural numbers. This corresponds to the case when V is inververtible or not.

The above definition of time operator is a straightforward generalization of time operator on Hilbert spaces [Pr,ASaSh]. The operator V is interpreted as a generalized dynamics. In the sequel we shall only consider time operators associated with a particular class of operators V that is specified below.

Consider a Banach space $\mathcal{B}$ that can be decomposed as an infinite direct sum of closed subspaces

$$\mathcal{B} = \bigoplus_{n\in\mathcal{I}} \mathcal{B}_n \tag{12.2}$$

in the sense that each $x \in \mathcal{B}$ has a unique representation

$$x = \sum_{n\in\mathcal{I}} x_n\,, \tag{12.3}$$

where $x_n \in \mathcal{B}_n$ and the series (12.3) converges in $\mathcal{B}$. A linear operator V on the Banach space $\mathcal{B}$ of the form (12.2) will be called a *generalized shift* with respect to $\{\mathcal{B}_n\}$ if V is bounded and satisfies

$$V\mathcal{B}_n \subset \mathcal{B}_{n+1}\,, \ \text{ for } n \in \mathcal{I}\,. \tag{12.4}$$

We do not assume that V is an isometry neither that it maps $\mathcal{B}_n$ onto $\mathcal{B}_{n+1}$.

Let P_n be the projection from $\mathcal{B}$ onto $\mathcal{B}_n$, i.e. P_n is a linear operator on $\mathcal{B}$ that corresponds to each $x \in \mathcal{B}$ its n-th component x_n in the representation (12.3). The family $\{P_n\}_{n\in\mathcal{I}}$ is a resolution of identity, $x = \sum_{n\in\mathcal{I}} P_n x$, for $x \in \mathcal{B}$, and determines a time operator. Namely we have

Proposition 12.1 *Assume that the Banach space $\mathcal{B}$ has the direct sum decomposition* (12.2) *and let $\{P_n\}_{n\in\mathcal{I}}$ be the corresponding family of projectors. Then the operator*

$$T = \sum_{n\in\mathcal{I}} nP_n \,, \tag{12.5}$$

defined for all $x \in \mathcal{B}$ for which the above series converges, is a time operator corresponding to any generalized shift V with respect to $\{\mathcal{B}_n\}_{n\in\mathcal{I}}$.

Proof. We shall show first that $V(D(T)) \subset D(T)$. Let $x = \sum_n x_n$ belong to the domain $D(T)$ of T. This means that the series $\sum_n nP_n x = \sum_n nx_n$ converges in $\mathcal{B}$. On the other hand, the series $\sum_n x_n$ is also convergent. Since V is bounded, by the assumption, both series $\sum_n Vx_n$ and $\sum_n nVx_n$ converge. Thus $\sum_n (n+1)Vx_n = \sum_n nVx_n + \sum_n Vx_n$ is also convergent, which shows that $Vx \in D(T)$.

In order to show the identity (12.5) notice first that

$$VP_n = P_{n+1}V\,, \ \ \text{for each } n \in \mathcal{I}\,.$$

Indeed, if $x \in \mathcal{B}$, $x = \sum_n x_n$, then $VP_n x = Vx_n$. Conversely, $P_{n+1}Vx = P_{n+1}\sum_k Vx_k = Vx_n$, since Vx_n belongs to $\mathcal{B}_{n+1}$.

By the induction we can show that

$$V^k P_n = P_{n+k}V^k\,, \ \ \text{for all } k\,, n \in \mathcal{I}\,. \tag{12.6}$$

Finally applying (12.6) and using the fact that the operators V^k are bounded and preserve the domain of T we have

$$TV^k x = \sum_n nP_n V^k x = V^k \sum_n nP_{n-k}x = V^k \sum_n (n+k)P_n x = V^k T + kV^k x\,.$$

Having the direct sum decomposition (12.2) of the Banach space $\mathcal{B}$ we can define a nested family of subspaces of $\mathcal{B}$ that corresponds to multiresolution analysis in wavelets or to filtration in stochastic processes. Define

$$\mathcal{B}_{\le n} \stackrel{\text{df}}{=} \bigoplus_{j\le n} \mathcal{B}_j$$

and denote by E_n the projection from $\mathcal{B}$ onto $\mathcal{B}_{\le n}$, i.e. $E_n = \sum_{j\le n} P_j$. E_n is the projection onto the past till the time instant n. The following elementary lemma relate the projectors E_n with the dynamics V.

Lemma 12.1 *Suppose that V is a generalized shift on $\mathcal{B} = \bigoplus_{n\in\mathcal{I}} \mathcal{B}_n$. Then we have*

$$V^k E_n = E_{n+k}V^k\,, \ \ \text{for all } k\,, n \in \mathcal{I}\,. \tag{12.7}$$

Proof. It is enough to show that (12.6) $\Rightarrow$ (12.7), which is elementary in the case $\mathcal{I} = \mathbb{Z}$. In the case $\mathcal{I} = \mathbb{N}$ we have

$$\begin{aligned} V^k E_n &= V^k P_1 + \ldots + V^k P_n \\ &= (P_{1+k} + \ldots + P_{n+k}) V^k \\ &= (E_{n+k} - E_k) V^k \\ &= E_{n+k} V^k \,. \end{aligned}$$

The latter equality follows from the direct sum decomposition (12.2) and the fact that $E_k V^k = 0$, for each $k \geq 1$.

We begin the construction of time operators associated with approximations in L^p spaces and in $\mathcal{C}_{[0,1]}$ with an application of the results from Section 7 concerning the time operator for the Renyi map, which is the simplest chaotic system and the prototype of exact endomorphisms [LM].

Recall that the 2-adic Renyi map is defined on the unit interval $[0, 1)$ by the formula

$$Sx = 2x \pmod 1 \,.$$

The Lebesgue measure is invariant with respect to S. The Koopman operator V

$$Vf(x) = f(Sx) = \begin{cases} f(2x)\,, & \text{for } x \in [0, \frac{1}{2}) \\ \\ f(2x-1)\,, & \text{for } x \in [\frac{1}{2}, 1)\,, \end{cases} \tag{12.8}$$

determines the evolution semigroup $\{V^n\}_{n \geq 0}$ on $L^p_{[0,1]}$, $p \geq 1$.

The eigenfunctions of the time operator are constructed as follows. First we define the functions

$$\varphi_0(x) = \mathbb{1}_{[0,1)}(2x) - \mathbb{1}_{[0,1)}(2x-1)$$

and

$$\varphi_k(x) = V^k \varphi_0(x)\,, \quad \text{for } k = 1, 2, \ldots \,.$$

For each natural number k there exist unique integers $n \geq 0$ and $\varepsilon_j = 0, 1$, $j = 0, \ldots, n-1$, $\varepsilon_n = 1$, such that

$$2^n \leq k < 2^{n+1} \quad \text{and} \quad k = \varepsilon_0 2^0 + \ldots + \varepsilon_n 2^n \,.$$

We put $w_0 = 1$ and, for $k \geq 1$,

$$w_k(x) = \varphi_0^{\varepsilon_0}(x) \ldots \varphi_n^{\varepsilon_n}(x)\,. \tag{12.9}$$

Thus $w_1 = \varphi_0$, $w_2 = \varphi_1$, $w_3 = \varphi_0 \varphi_1$, and so on.

We have modified here the ordering of the eigenfunctions in order to get a direct connection with the Walsh-Paley system. Inded, observe that extending functions w_k periodically on $\mathbb{R}$ and taking into account the above introduced ordering we obtain the Walsh-Paley system [SWS].

The functions $w_0, \ldots, w_{2^n - 1}$ form a basis in the vector space of all functions that are measurable with respect to the σ-algebra $\mathcal{A}_n$ generated by the dyadic division of

$[0,1]$ on 2^n parts. The block $w_{2^n}, \ldots, w_{2^{n+1}-1}$ is the contribution (details) that is necessary in order to obtain all $\mathcal{A}_{n+1}$ measurable functions.

It is well known [SWS] that $w_0, w_2, \ldots$ form a Schauder basis in each of the spaces $L^p_{[0,1]}$, $1 < p < \infty$. This means that each function $f \in L^p_{[0,1]}$ has a unique expansion

$$f = w_0 + \sum_{j=1}^{\infty} a_j w_j = w_0 + \sum_{n=1}^{\infty} \sum_{k=2^{n-1}}^{2^n-1} a_k w_k$$

convergent in the L^p-norm. In particular $w_0, w_1, \ldots$ form a complete orthonormal system in the Hilbert space $L^2_{[0,1]}$. Therefore each L^p space with $1 < p < \infty$ has the following direct sum decomposition

$$L^p_{[0,1]} = \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \ldots ,$$

where W_0 is the space of constant functions and $\mathcal{W}_n$, $n = 1, 2, \ldots$, is the linear space spanned by $w_{2^{n-1}}, \ldots, w_{2^n-1}$.

Denote by P_n the projection onto $\mathcal{W}_n$

$$P_n = \sum_{k=2^{n-1}}^{2^n-1} \langle \cdot , w_k \rangle w_k, \tag{12.10}$$

and put $\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \ldots$. We have

Proposition 12.2 ([AStime]) *The operator T defined on $\mathcal{W}$ as*

$$Tf = \sum_{n=1}^{\infty} n P_n f$$

is a time operator with respect to the semigroup $\{V^n\}_{n=1}^{\infty}$ generated by the Koopman operator of the Renyi map. Each number $n = 1, 2, \ldots$ is an eigenvalue of T and the functions $w_{2^{n-1}}, \ldots, w_{2^n-1}$ are the corresponding eigenvectors.

In order to prove the above theorem one should notice that the Koopman operator is a multiplication operator and $V(\varphi_j^{\varepsilon_j}) = \varphi_j^{\varepsilon_{j+1}}$. Hence V transports any Walsh function w_k, which is of the form (12.9), from $\mathcal{W}_n$ into a Walsh function from $\mathcal{W}_{n+1}$. Then it is enough to check that the assumptions of Lemma 1 are fulfilled.

The time operator constructed in Proposition 12.1 is associated with approximations of p-integrable functions by step functions. The structure of T when restricted to the L^2-space coincides with the multiresolution analysis associated with the Haar wavelet.

We remind the reader that a *multiresolution analysis* (MRA) of $L^2_{\mathbb{R}}$ is a sequence $\{\mathcal{V}_n\}_{n\in\mathbb{Z}}$ of closed subspaces of $L^2_{\mathbb{R}}$ such that

(i) $\{0\} \subset \ldots \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \ldots \subset L^2_{\mathbb{R}}$,

(ii) $\bigcap_{n\in\mathbb{Z}} \mathcal{V}_n = \{0\}$

(iii) $\overline{\bigcup_{n\in\mathbb{Z}} \mathcal{V}_n} = L^2(\mathbb{R})$,

(iv) $f(\cdot) \in \mathcal{V}_0 \Longleftrightarrow f(2^n \cdot) \in \mathcal{V}_n$

(v) There is a function $\phi \in L^2_{\mathbb{R}}$ (the *scaling function*) whose integer translates $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ form an orthonormal basis in $\mathcal{V}_0$.

Sometimes the condition(v) in the definition of MRA is replaced by a weaker condition

(v') There is a function $\phi \in L^2_{\mathbb{R}}$ such that the set $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ form a Riesz basis in $\mathcal{V}_0$, i.e. it is dense in $\mathcal{V}_0$ and there exist positive constants A and B such that

$$A \sum_{k\in\mathbb{Z}} c_k^2 \leq \| \sum_{k\in\mathbb{Z}} c_k \phi(\cdot - k)\|^2_{L^2(\mathbb{R})} \leq B \sum_{k\in\mathbb{Z}} c_k^2 \,,$$

for each $\{c_k\}_{k\in\mathbb{Z}} \subset \mathbb{R}$ such that $\sum_{k\in\mathbb{Z}} c_k^2 < \infty$.

The replacement of the condition (v) by (v') is also necessary for an extension of the concept of MRA on Banach spaces where the notion of orthogonality is in general meaningless.

We have shown in Section 7, restricting the Haar wavelet to the interval $[0,1]$ by the periodization method (see [Da]), that the eigenspaces $\mathcal{W}_1, W_2, \ldots$ of the time operator of the Renyi map coincide with the corresponding wavelet spaces. The ladder of spaces $\mathcal{W}_0 \subset \mathcal{W}_0 \oplus \mathcal{W}_1 \subset \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \mathcal{W}_2 \subset \ldots$ forms the multiresolution spaces of the Haar wavelet restricted to $[0,1]$. Moreover the condition (iv), when applied to a function f defined on the interval [0,1] and then extended periodically on $\mathbb{R}$ is nothing but the action of the Koopman operator of the Renyi map.

Generalizing the above considerations we can interpret the multiresolution analysis associated with a given wavelet as a spectral decomposition of the time operator associated with the operator $f(x) \mapsto f(2x)$ (the Koopman operator of the Renyi map when restricted by periodization to $[0,1]$).

This means that the time operator method is a generalization of MRA on the case when the scaling condition (iv) is replaced by the “keeping pace condition” (12.1) to be satisfied by an arbitrary bounded operator.

We would like to comment on the multiplicity of time operators that can be associated with wavelets. Using wavelets to analyze a signal we do not restrict a'priori its wave length, which can be arbitrary large. Therefore the MRA gives rise to a time operator with a uniform (infinite) multiplicity [AnGu]. However, using a MRA for the approximation of functions on $[0,1]$ we encounter the natural bound for the wave length, which is 1. Such is the case of the Haar wavelet on $[0,1]$, where the multiplicity of the eigenspace $\mathcal{W}_n$ is 2^{n-1}.

12.2 TIME OPERATOR AND SHANNON THEOREM

We present now an interesting connection of time operator and Shannon's sampling theorem. Recall first that the function

$$\phi(x) = \mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x} \tag{12.11}$$

is called the Shannon scaling function. The function $\phi(x)$ is the Fourier transform of $\mathbb{1}_{[-\pi,\pi]}$ – the indicator of the set $[-\pi, \pi]$. Using the scaling function we define in the usual way the nested family $\ldots \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \ldots$ of subspaces of $L^2_{\mathbb{R}}$, i.e. $\mathcal{V}_n$ is spanned by all the translations $\phi(2^n x - k)$, $k \in \mathbb{Z}$. In this way we obtain the multiresolution analysis associated with ϕ [Da]. Denoting by $\mathcal{W}_n$ the orthogonal complement of $\mathcal{V}_n$ in $\mathcal{V}_{n+1}$, i.e. $\mathcal{V}_{n+1} = \mathcal{W}_n \oplus \mathcal{V}_n$, we obtain the following direct sum decomposition

$$L^2_{\mathbb{R}} = \bigoplus_{n \in \mathbb{Z}} \mathcal{W}_n \, .$$

Therefore the operator $T = \sum_{n \in \mathbb{Z}} n P_n$, where P_n is the projection onto $\mathcal{W}_n$ is a time operator with respect to any bounded operator V, which satisfies the conditions of Lemma 1. In particular, V can be the shift $Vf(x) = f(2x)$.

Knowing that for each n the set $\{\psi(2^n x - k),\ k \in \mathbb{Z}\}$, where

$$\psi(x) = (\pi x)^{-1}(\sin 2\pi x - sin\pi x) \, ,$$

forms an orthonormal basis in $\mathcal{W}_n$ we obtain the explicit form of the eigenprojectors P_n

$$P_n f(x) = \sum_{k \in \mathbb{Z}} \langle f, \psi(2^n \cdot -k) \rangle \psi(2^n x - k) \, .$$

Using the time operator language we can say that the space $\mathcal{V}_n$ consists of the elements from $L^2_{\mathbb{R}}$ that are no older than n. Moreover the Shannon theorem allows to identify these elements.

To see this recall the Shannon theorem, which says that if a function $f \in L^2_{\mathbb{R}}$ is such that its Fourier transform is supported on the interval $[-2^n\pi, 2^n\pi]$, $n \in \mathbb{Z}$, then

$$f(x) = \sum_{k \in \mathbb{Z}} f\left(\frac{k}{2^n}\right) \frac{\sin \pi(2^n x - k)}{\pi(2^n x - k)} = \sum_{k \in \mathbb{Z}} f\left(\frac{k}{2^n}\right) \phi(2^n x - k) \, . \tag{12.12}$$

Since the family $\{\phi(2^n x - k),\ k \in \mathbb{Z}\}$ forms an orthonormal basis in $\mathcal{V}_n$, the equation (12.12) says that the projection

$$E_n = \sum_{j \leq n} P_n$$

on the past prior n leaves f invariant. Therefore the age of f is at most n.

In terms of the signal processing we say that f is band limited with the band width Ω if its Fourier transform has a compact support contained in the interval $[-\Omega, \Omega]$.

Therefore the projection on the past prior n is equivalent to apply the low pass filter for damping the frequencies higher than 2^n. Correspondingly, the projection on the future is equivalent to apply the high pass filter. By the Shannon theorem *the age of (a signal) $f \in L^2_{\mathbb{R}}$ is equivalent to the sampling that is necessary to recover* f. The transition from the time instant n to $n+1$ means the sampling of f with doubled accuracy. The natural flow of time, $-\infty < n < \infty$, is reflected in the signal processing as the transition from large to fine scales (resolutions).

12.3 TIME OPERATOR ASSOCIATED WITH THE HAAR BASIS IN $L^2_{[0,1]}$

The importance of the Walsh basis in the construction of the time operator associated with the Renyi map is due to the fact that the spectral decomposition of T is particularly simple. The Walsh functions are the eigenvectors of T and the action of the Koopman operator can be described explicitly as a shift from one Walsh function to another. On the other hand the Walsh basis is not so convenient when dealing with continuous functions. For example, although the Walsh functions are linear combinations of the Haar functions it is well known that the Haar expansion of a continuous function on the interval $[0,1]$ converges uniformly, while its Walsh series may be even pointwise divergent. For this reason it would be more convenient, when dealing with continuous functions, to represent T in the basis of Haar functions.

In Section 7 we represented the time operator T associated with the Renyi map in the Haar basis. Below we generalize this result showing that T as represented in the Haar basis is in fact a time operator with respect to a wide class of dynamical semigroups. We shall focus our attention on T as an operator on the space of continuous functions. Accordingly, we discuss the conditions under which a continuous function belongs to the domain of T and is image through T is also a continuous function. We shall also show in this section how a time rescaling affects the dynamics giving an estimation of the correlation function associated with the underlying dynamical semigroup. We shall see that although for a wide class of functions the decay is exponential, the exponent depends on the "degree of smoothness" of the considered function. The additional advantage of such choice is that the Haar basis can be transported to the space $C_{[0,1]}$ giving rise to the time operator associated with the Faber-Schauder basis.

Recall that the Haar functions χ_j on the interval $[0,1)$ bare defined as follows:

$$\chi_1 \equiv 1\,, \quad \chi_{2^n+k}(x) = 2^{\frac{n}{2}}\,\mathbb{1}_{\left[\frac{2k-2}{2^{n+1}},\frac{2k-1}{2^{n+1}}\right)}(x) - 2^{\frac{n}{2}}\,\mathbb{1}_{\left[\frac{2k-1}{2^{n+1}},\frac{2k}{2^{n+1}}\right)}(x)\,,$$

for $n = 0, 1, \ldots, k = 1, \ldots, 2^n$.

For a given n the Haar functions χ_{2^n+k}, $k = 1, \ldots, 2^n$, are the eigenfunctions of T corresponding to the same eigenvalue n (see Section 7). Each function $f \in L^2_{[0,1]}$ has the expansion in the Haar basis, which can be written in one of the following equivalent forms

$$f = \sum_{j=1}^{\infty} a_j \chi_j = a_1 \chi_1 + \sum_{m=0}^{\infty} \sum_{k=1}^{2^m} a_{2^m+k} \chi_{2^m+k} = a_1 \chi_1 + \sum_{n=1}^{\infty} \sum_{k=2^{n-1}+1}^{2^n} a_k \chi_k \,, \tag{12.13}$$

with $a_j = \int_0^1 f(x)\chi_j(x)dx$.

The Haar functions form an orthonormal basis in $L^2_{[0,1]}$ and the linear space generated by $\{\chi_k\}_{k=2^{n-1}+1}^{2^n}$ coincides with $\mathcal{W}_n$. Also the orthogonal projection onto the space generated by $\{\chi_k\}_{k=2^{n-1}+1}^{2^n}$ coincides with the orthogonal projection P_n as defined by (12.10) for the Walsh basis. Therefore introduced in Section 2 time operator T on $\mathcal{W}$ assumes the form

$$Tf = \sum_{n=1}^{\infty} n \sum_{k=2^{n-1}+1}^{2^n} a_k \chi_k \,, \tag{12.14}$$

where $f \in L^2_{[0,1]}$ with $\int_0^1 f(x)\,dx = 0$ and a_k are as in (12.13). The Koopman operator V of the Renyi map does not transport a Haar function corresponding to the eigenvalue n onto a single Haar function but onto a linear combination of two Haar functions corresponding to the eigenvalue $n+1$ (see Section 7). Nevertheless one can check that the assumptions of Lemma 1 are satisfied so that the commutation relation (12.1) still holds. This however follows from the following more general result:

Theorem 12.1 *The operator T defined on $\mathcal{W} = L^2_{[0,1]} \ominus \{1\}$ as*

$$Tf = \sum_{n=1}^{\infty} nP_n \,,$$

where

$$P_n f = \sum_{2^{n-1}+1}^{2^n} \left[\int_0^1 f(x)\chi_k(x)dx \right] \chi_k$$

is a time operator with respect to any semigroup $\{V^n\}_{n\geq 0}$, where V is a bounded operator on $\mathcal{W}$ such that $V(\mathcal{W}_n) \subset \mathcal{W}_{n+1}$, for each $n = 1, 2, \ldots$.

Proof. Recall that we have the direct sum decomposition

$$\mathcal{W} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \ldots$$

and that each P_n is the orthogonal projector onto $\mathcal{W}_n$. We shall show that

$$VP_n = P_{n+1}V \,,$$

for each $n \in \mathbb{N}$. Indeed, since each $f \in \mathcal{W}$ has a unique expansion

$$f = \sum_{n=1}^{\infty} \sum_{k=2^{n-1}+1}^{2^n} a_k \chi_k \,. \tag{12.15}$$

then

$$VP_nf = V\left(\sum_{k=2^{n-1}+1}^{2^n} a_k\chi_k\right) = \sum_{k=2^{n-1}+1}^{2^n} a_kV\chi_k\,.$$

On the other hand, among the elements $\{V\chi_j\}$ only those with $2^{n-1} < j \le 2^n$ are elements of $\mathcal{W}_{n+1}$. Therefore

$$P_{n+1}Vf = \sum_{j=2^{n-1}+1}^{2^n} a_jV\chi_j\,,$$

which implies (12.23). Consequently the assumptions of Lemma 12.1 are satisfied and $\widetilde{T}$ is a time operator with respect to $\{V^n\}_{n\ge 0}$.

In the sequel we shall also consider time operator on other spaces than $L^2_{[0,1]}$. If it is not stated otherwise, then the letter T will stand for the time operator as defined by (12.14).

One of the advantages of the choice of Haar functions to determine T is that they are more suitable when dealing with continuous functions. As we have shown in Section 3 (Theorem 3) any function $f \in \mathcal{C}_{[0,1]}$ such that its modulus of continuity satisfies the property

$$\int_0^1 \omega_f(x)\frac{\log x}{x}dx > -\infty$$

belongs to the domain of T and $Tf \in \mathcal{C}_{[0,1]}$.

If a function f is a Lipschitz function, i.e. there are constants $K > 0$ and $0 < \alpha \le 1$ such that $|f(x) - f(y)| \le K|x - y|^\alpha$, for each $x, y \in [0,1]$, then $\omega_f(x) \le K|x|^\alpha$. It is therefore easy to see that a Lipschitz function with an exponent α, $0 < \alpha \le 1$, satisfies the assumption of Theorem 1 and, consequently, belongs to the domain of T expanded in the Haar basis.

It is worth to note that if the time operator T is expanded in terms of the Walsh basis then the sufficient condition that a continuous function belongs to its domain is more restrictive. In particular, in the class of Lipschitz functions only those with the exponent $\alpha > \frac{1}{2}$ belong to the domain of T. The proof of this fact, which is based on some estimations for the Walsh-Fourier coefficients, will be presented elsewhere.

Note that the family of Haar functions forms also a Schauder basis in the Banach space $L^p_{[0,1]}$, $1 \le p < \infty$ and this basis is unconditional if $p > 1$. This means that every function $f \in L^p_{[0,1]}$ has representation (12.13) convergent in L^p-norm (unconditionally convergent if $p > 1$). The Walsh functions also form a Schauder basis in L^p, $1 < p < \infty$, but not in L^1 (see [CiKw] and references therein).

One of the basic tools in the study of dynamical systems with the use of time operator is time scaling, which corresponds to filtering in signal processing. Scaling is defined as replacing the time operator T by some of its operator function $\Lambda(T)$, where $\Lambda(\cdot)$ is a real valued function. The application of Λ on T may affect its domain. Modifying slightly Theorem 2 in Section 7 it is also possible to give sufficient conditions under which a continuous function f belong to the domain of $\Lambda(T)$ and $\Lambda(T)f$ is also continuous.

Let us show now how scaling affect the dynamics. We consider the action semigroup $\{V^N\}$ on the functions transformed through $\Lambda(T)$. The relevant value that we would like to evaluate is the correlation function

$$R_f(N) \stackrel{\mathrm{df}}{=} (V^N \Lambda(T)(f), \Lambda(T)(f)) \,,$$

where $(\cdot,\cdot)$ denotes the scalar product in $L^2_{[0,1]}$.

Theorem 12.2 *Let V be the Koopman operator of some map of the interval $[0,1]$ such that the corresponding semigroup $\{V^N\}$ on $L^2_{[0,1]}$ satisfies the assumptions of Theorem 1. Then for every function $f \in \mathcal{C}_{[0,1]}$ which belongs to the domain of $\Lambda(T)$ we have*

$$|R_f(N)| \le \frac{1}{4} \sum_{n=1}^{\infty} |\Lambda(n)\Lambda(n+N)| \omega\left(\frac{1}{2^n}\right) \omega\left(\frac{1}{2^{n+N}}\right) .$$

Proof. In order to calculate the correlation function let us represent Λ as follows

$$\Lambda(f) = \sum_{n=0}^{\infty} \Lambda(n+1) \sum_{k=1}^{2^n} a_{2^n+k} \chi_{2^n+k} \,.$$

Let N and n be fixed. By the assumption each $V^N \chi_{2^n+k}$, $k = 1, \ldots, 2^n$, is a linear combination of some basis elements $\chi_{2^{n+N}+k}$:

$$V^N \chi_{2^n+k} = \sum_{j=1}^{n_k} \alpha_{2^{n+N}+l_j(k)} \chi_{2^{n+N}+l_j(k)} \,, \quad k = 1, \ldots, 2^n \,, \tag{12.16}$$

for some choice of indices $l_1(k), \ldots, l_{n_k}(k)$.

Since V is also a Koopman operator, V^N is a multiplicative map, i.e. $Vfg = VfVg$. Because, for a given n, the functions χ_{2^n+k}, $k = 1, \ldots, 2^n$, have disjoint supports, then also $V^N \chi_{2^n+k}$ must have disjoint supports. This implies that in the representation (12.16) all $l_j(k)$ are different and

$$n_1 + \ldots + n_{2^n} \le 2^{n+N} \,. \tag{12.17}$$

We have

$$R_f(N) = \left(\sum_{n=0}^{\infty} \Lambda(n+1) \sum_{k=1}^{2^n} a_{2^n+k} V^N \chi_{2^n+k}, \sum_{n=0}^{\infty} \Lambda(n+1) \sum_{k=1}^{2^n} a_{2^n+k} \chi_{2^n+k} \right)$$

$$= \sum_{n=0}^{\infty} \Lambda(n+1)\Lambda(n+N+1) \left(\sum_{k=1}^{2^n} a_{2^n+k} \sum_{j=1}^{n_k} \alpha_{2^{n+N}+l_j(k)} \chi_{2^{n+N}+l_j(k)}, \right.$$

$$\left. \sum_{k=1}^{2^{n+N+1}} a_{2^{n+N}+k} \chi_{2^{n+N}+k} \right)$$

$$= \sum_{n=0}^{\infty} \Lambda(n+1)\Lambda(n+N+1) \sum_{k=1}^{2^n} a_{2^n+k} \sum_{j=1}^{n_k} a_{2^{n+N}+l_j(k)} \alpha_{2^{n+N}+l_j(k)} \,.$$

Thus

$$|R_f(N)| \le \sum_{n=0}^{\infty} |\Lambda(n+1)\Lambda(n+N+1)| \sum_{k=1}^{2^n} \frac{1}{2 \cdot 2^{\frac{n}{2}}} \omega\left(\frac{1}{2^{n+1}}\right)$$

$$\cdot \sum_{j=1}^{n_k} \frac{1}{2 \cdot 2^{\frac{n+N}{2}}} \omega\left(\frac{1}{2^{n+N+1}}\right) \left|\alpha_{2^{n+N}+l_j(k)}\right|$$

$$= \frac{1}{4} \cdot \frac{1}{2^{\frac{N}{2}}} \sum_{n=0}^{\infty} |\Lambda(n+1)\Lambda(n+N+1)| \frac{1}{2^n} \omega\left(\frac{1}{2^{n+1}}\right) \omega\left(\frac{1}{2^{n+N+1}}\right)$$

$$\cdot \sum_{k=1}^{2^n} \sum_{j=1}^{n_k} \left|\alpha_{2^{n+N}+l_j(k)}\right| .$$

Note that

$$\left\|V^N \chi_{2^n+k}\right\|_{L^2}^2 = \left\| \sum_{j=1}^{n_k} \alpha_{2^{n+N}+l_j(k)} \chi_{2^{n+N}+l_j(k)} \right\|^2 = \sum_{j=1}^{n_k} \left|\alpha_{2^{n+N}+l_j(k)}\right|^2 ,$$

and since $\|V\| \le 1$, we have

$$\sum_{j=1}^{n_k} \left|\alpha_{2^{n+N}+l_j(k)}\right|^2 \le \left\|V^N\right\|^2 \|\chi_{2^{n+N}+k}\|_{L^2}^2 \le 1 \,.$$

Therefore, applying Hölder's inequality we get

$$\sum_{j=1}^{n_k} \left|\alpha_{2^{n+N}+l_j(k)^{\frac{1}{2}}}\right| \le \sqrt{n_k} \left(\sum_{j=1}^{n_k} \left|\alpha_{2^{n+N}+l_j(k)}\right|^2 \right) \le \sqrt{n_k} \,.$$

Applying Hölder's inequality once more together with (12.17) we obtain

$$\sum_{k=1}^{2^n}\sum_{j=1}^{n_k}\left|\alpha_{2^{n+N}+l_j(k)}\right| \le \sum_{k=1}^{2^n}\sqrt{n_k} \le 2^{\frac{n}{2}}\left(\sum_{k=1}^{2^n} n_k\right)^{\frac{1}{2}} \le 2^{\frac{n}{2}}\cdot 2^{\frac{n+N}{2}} .$$

Consequently

$$|R_f(N)| \le \frac{1}{4}\cdot\frac{1}{2^{\frac{N}{2}}}\sum_{n=0}^{\infty}|\Lambda(n+1)\Lambda(n+N+1)|\frac{1}{2^n}\omega\left(\frac{1}{2^{n+1}}\right)\omega\left(\frac{1}{2^{n+N+1}}\right)2^n 2^{\frac{N}{2}} ,$$

and rearranging the summation we finally get

$$|R_f(N)| \le \frac{1}{4}\sum_{n=1}^{\infty}|\Lambda(n)\Lambda(n+N)|\omega\left(\frac{1}{2^n}\right)\omega\left(\frac{1}{2^{n+N}}\right) ,$$

which ends the proof.

Under the same assumptions as in Theorem 12.2 we have

Corollary 12.1 *If f is a Lipschitz function with the exponent p, $0 < p \le 1$, then*

$$|R_f(N)| \le \frac{1}{4}\cdot\frac{1}{2^{pN}}\sum_{n=1}^{\infty}|\Lambda(n)\Lambda(n+N)|\frac{1}{2^{2pn}} . \tag{12.18}$$

Corollary 12.2 *If $\Lambda(x)$ is a bounded function then the series on the right hand side of (12.18) is convergent and*

$$|R_f(N)| \le K\left(\frac{1}{2^p}\right)^N = Ke^{-(p\ln 2)N} ,$$

for some $K > 0$.

It should be noticed that for a bounded Λ any continuous function is in the domain of $\Lambda(T)$. In particular, taking $\Lambda(x) \equiv 1$ we see that the decay of the correlations of $(V^N f, f)$ is at least exponential with the exponent depending on the degree of smoothness of f. If f is differentiable then the exponent is $\ln 2$.

Example Let V be the Koopman operator of the Renyi map. Since

$$V\chi_{2^n+k} = \frac{1}{\sqrt{2}}\left(\chi_{2^{n+1}+k} + \chi_{2^{n+1}+2^n+k}\right) ,$$

V satisfies the assumptions of Theorem 1.

12.4 TIME OPERATOR ASSOCIATED WITH THE FABER-SCHAUDER BASIS IN $\mathcal{C}_{[0,1]}$

Although each continuous function can be expanded in terms of the Haar basis the Haar functions lay outside the space $\mathcal{C}_{[0,1]}$. We shall show that the Haar basis in $L^1_{[0,1]}$ can be transported to $\mathcal{C}_{[0,1]}$ by the means of integration giving rise to a new basis in $\mathcal{C}_{[0,1]}$. Namely, let us define the operator of indefinite integration $J : L^1_{[0,1]} \to \mathcal{C}_{[0,1]}$:

$$(Jf)(t) \stackrel{\text{df}}{=} \int_0^t f(s)ds\,, \quad \text{for } f \in L^1_{[0,1]}\,.$$

The range of J consists of absolutely continuous functions. Since the series (12.13) is also uniformly convergent [KaSt] we can apply J to both sides getting

$$\begin{aligned}(Jf)(t) &= \sum_{j=1}^{\infty} \left[\int_0^1 f(s)\chi_j(s)ds\right](J\chi_j)(t) \\ &= \sum_{j=1}^{\infty} \left[\int_0^1 \chi_j(s)d(Jf)(s)\right]\varphi_j(t)\,,\end{aligned} \tag{12.19}$$

where $\varphi_j(t) \stackrel{\text{df}}{=} (J\chi_j)(s)\,,\ \ j = 1, 2, \ldots$.

Actually the representation (12.19) extends on all functions $g \in \mathcal{C}_{[0,1]}$. To be more precise the family $\{\varphi_j\}_{j=1}^{\infty}$ together with the constant function $\varphi_0 \equiv 1$ form a Schauder basis in the Banach space $\mathcal{C}_{[0,1]}$. We have [Ci]

$$g(t) = g(0)\varphi_0 + \sum_{j=1}^{\infty} \left[\int_0^1 \chi_j(s)\,dg(s)\right]\varphi_j(t)\,, \tag{12.20}$$

where the series converges uniformly in $[0, 1]$.

In a similar way, applying J to both sides of (12.14), we can transport the time operator T to $\mathcal{C}_{[0,1]}$. To be more precise, let $\widetilde{\mathcal{C}}$ be the space of all functions $g \in \mathcal{C}_{[0,1]}$ such that $g(0) = g(1) = 0$ and let $\mathcal{C}_n$, $n = 1, 2, \ldots$, be the subspace of $\widetilde{\mathcal{C}}$ spanned by φ_k, $2^{n-1} < k \leq 2^n$. Define the operator $\widetilde{P}_n : \widetilde{\mathcal{C}} \to \mathcal{C}_n$ putting

$$\widetilde{P}_n g(t) = \sum_{k=2^{n-1}+1}^{2^n} \int_0^1 \chi_k(s)dg(s)\varphi_k(t)\,. \tag{12.21}$$

Theorem 12.3 *The operator $\widetilde{T}$ defined on $\widetilde{\mathcal{C}}$ as*

$$\widetilde{T} = \sum_{n=1}^{\infty} n\widetilde{P}_n\,,$$

is a time operator with respect to any semigroup $\{V^n\}_{n\geq 0}$, where V is a bounded operator on $\widetilde{\mathcal{C}}$ such that $V(\mathcal{C}_n) \subset \mathcal{C}_{n+1}$, for each $n = 1, 2, \ldots$. The explicit form of

$\widetilde{P}_n$ *is*

$$\widetilde{P}_n g(t) = 2^{\frac{n-1}{2}} \sum_{k=1}^{2^{n-1}} \left[2g\left(\frac{2k-1}{2^n}\right) - g\left(\frac{k-1}{2^{n-1}}\right) - g\left(\frac{k}{2^{n-1}}\right)\right] \varphi_{2^{n-1}+k}(t) \tag{12.22}$$

Proof. Since the functions $\varphi_0, \varphi_1, \dots$ form a Schauder basis in $\mathcal{C}_{[0,1]}$ we have the direct sum decomposition

$$\widetilde{\mathcal{C}} = \mathcal{C}_1 \oplus \mathcal{C}_2 \oplus \dots$$

and each $\widetilde{P}_n$ is the projector onto $\mathcal{C}_n$. It is easy to see that

$$V\widetilde{P}_n = \widetilde{P}_{n+1}V \, , \tag{12.23}$$

for each $n \in \mathbb{N}$. Indeed, since each $g \in \widetilde{\mathcal{C}}$ has a unique expansion

$$g = \sum_{n=1}^{\infty} \sum_{k=2^{n-1}+1}^{2^n} b_k \varphi_k \, . \tag{12.24}$$

then

$$VP_n g = V\left(\sum_{k=2^{n-1}+1}^{2^n} b_k \varphi_k\right) = \sum_{k=2^{n-1}+1}^{2^n} b_k V\varphi_k \, .$$

On the other hand, among the elements $\{V\varphi_j\}$ only those with $2^{n-1} < j \le 2^n$ are elements of $\mathcal{C}_{n+1}$. Therefore

$$P_{n+1}Vg = \sum_{j=2^{n-1}+1}^{2^n} b_j V\varphi_j \, ,$$

which implies (12.23). Consequently the assumptions of Lemma **??** are satisfied and $\widetilde{T}$ is a time operator with respect to $\{V^n\}_{n\ge 0}$. The explicit form (12.22) of $\widetilde{P}_n$ follow directly from (12.21).

Constructed in this way time operator arises as the integral transformation of the time operator T for the Renyi map expanded in terms of the Haar basis, i.e. for $g = Jf$ we have

$$\widetilde{T}g = Tf \, .$$

Moreover, since $\widetilde{T}$ is a time operator with respect to any bounded operator V, which maps each Schauder function $\varphi_{2^{n-1}+k}$, $k = 1, \dots, 2^{n-1}$, onto a linear combination of the functions $\varphi_{2^n+k'}$, $k' = 1, \dots, 2^n$, it can be also associated with the Koopman operator V of the Renyi map (12.8) acting on the space $\widetilde{\mathcal{C}}$. Indeed, V is, of course, bounded on $\widetilde{\mathcal{C}}$ and one can check easily that

$$V\varphi_{2^{n-1}+k} = \sqrt{2}(\varphi_{2^n+k} + \varphi_{2^n+2^{n-1}+k}) \, ,$$

for $n = 1, 2 \ldots$, $k = 1, \ldots, 2^n$.

Observe also that the Koopman operator V of the Renyi map together with the integral operator T satisfy the following commutation relation

$$VJf = 2JVf \,, \tag{12.25}$$

valid for each $f \in L^1$ such that $\int_0^1 f(s)d(s) = 0$.

Using the time operator terminology we can say now that the function $g \in \widetilde{\mathcal{C}}$, has the age n if its representation (12.20) consists of the n-th block, i.e. those with the indices $k = 2^{n-1} + 1, \ldots, 2^n$. Therefore the flow of time means step by step interpolation of g by polygonal lines. The polygonal line $l_n(x)$, $n = 1, 2, \ldots$, corresponding to the dyadic division of the interval $[0, 1]$ on 2^n parts is

$$l_n(x) = \sum_{k=1}^{2^n} \left\{ \left[g\left(\frac{k}{2^n}\right) - g\left(\frac{k-1}{2^n}\right) \right] (2^n x - k + 1) + g\left(\frac{k-1}{2^n}\right) \right\} \mathbb{1}_{\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]}(x) \,.$$

Since the time operators considered here are defined on infinite dimensional Banach spaces, it is easy to see that their domains are always proper subsets of the underlined spaces. Therefore it arises the problem of characterization of the domain of a time operator. It is also important for applications of time operator techniques, especially for filtering, to characterize the domain of a function of a given time operator.

It is easy to see that if the eigenfunctions of a time operator T defined on a Banach space $\mathcal{B}$ form an unconditional Schauder basis then for any bounded function $\Lambda : \mathbb{N} \to \mathbb{R}$ the domain of $\Lambda(T)$ coincides with $\mathcal{B}$. However the space $\mathcal{C}_{[0,1]}$ does not have an unconditional basis. Therefore not for each bounded function Λ the operator $\Lambda(\widetilde{T})$ is correctly defined on $\mathcal{C}_{[0,1]}$. The next theorem provides sufficient conditions for a function g to be in the domain of $\widetilde{T}$, as well as to be in the domain of $\Lambda(\widetilde{T})$ in terms of the modulus of continuity.

Theorem 12.4 *Let Λ be a real valued function defined on $\mathbb{N}$. Any function $g \in \widetilde{\mathcal{C}}$ such that its modulus of continuity ω_g satisfies the property*

$$\sum_{n=1}^{\infty} |\Lambda(n)| \omega_g(2^{-n}) < \infty \tag{12.26}$$

belongs to the domain of $\Lambda(\widetilde{T})$ and the series $\sum_n \Lambda(n)\widetilde{P}_n g(t)$ is uniformly and absolutely convergent. In particular, if the modulus of continuity ω_g satisfies

$$\int_0^1 \omega_g(t) \frac{\log t}{t} dt > -\infty \tag{12.27}$$

then g belongs to the domain of $\widetilde{T}$ and the series

$$\sum_{n=1}^{\infty} n\widetilde{P}_n g(t) \tag{12.28}$$

is uniformly and absolutely convergent.

Proof. Let $g \in \widetilde{C}$ be such that its modulus of continuity satisfies (12.26). Let $b_{n,k} \stackrel{\text{df}}{=} \int_0^1 \chi_{2^{n-1}+k}(s)dg(s)$. We have

$$\begin{aligned} |b_{n,k}| &= \left|\int_0^1 \chi_{2^{n-1}+k}(s)dg(s)\right| \\ &\le 2^{\frac{n-1}{2}} \left|g\left(\tfrac{k-1}{2^{n-1}}\right) - g\left(\tfrac{2k-1}{2^n}\right)\right| + 2^{\frac{n-1}{2}} \left|g\left(\tfrac{2k-1}{2^n}\right) - g\left(\tfrac{k}{2^{n-1}}\right)\right| \\ &\le 2^{\frac{n+1}{2}} \omega_g\left(\tfrac{1}{2^n}\right), \end{aligned}$$

for $n = 1, 2, \ldots,\ k = 1, \ldots, 2^{n-1}$. Therefore

$$\sum_{n=1}^{\infty} |\Lambda(n)\widetilde{P}_n g(t)| \le \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} |\Lambda(n) b_{n,k}| \varphi_{n,k}(t) \le 2 \sum_{n=1}^{\infty} |\Lambda(n)| \omega_g\left(\frac{1}{2^n}\right),$$

which proves the first part of the theorem. To show the second part observe that putting $\Lambda(x) = x$ we have

$$\sum_{n=1}^{\infty} n|\widetilde{P}_n g(t)| \le \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} n|b_{n,k}|\varphi_{n,k}(t) \le \sum_{n=1}^{\infty} n\omega_g\left(\frac{1}{2^n}\right).$$

On the other hand, since the function $\omega_g(t)$ is non-decreasing on $[0, 1]$ and $-\frac{\log t}{t}$ is decreasing, we have

$$\begin{aligned} -\int_0^1 \omega_g(t)\frac{\log t}{t}dt &= \sum_{n=1}^{\infty} \int_{\frac{1}{2^n}}^{\frac{1}{2^{n-1}}} \omega_g(t)\left(-\frac{\log t}{t}\right)dt \\ &\ge \sum_{n=1}^{\infty} \omega_g\left(\frac{1}{2^{n-1}}\right)\left(-\frac{\log \frac{1}{2^n}}{\frac{1}{2^n}}\right)\left(\frac{1}{2^{n-1}} - \frac{1}{2^n}\right) \\ &= \log 2 \sum_{n=1}^{\infty} n\omega_g\left(\frac{1}{2^{n-1}}\right) \\ &\ge 4^{-1} \log 2 \sum_{n=1}^{\infty} n\omega_g\left(\frac{1}{2^n}\right). \end{aligned} \tag{12.29}$$

The last inequality is a consequence of the property $\omega_g(x+y) \le \omega_g(x)+\omega_g(y)$ valid for $x, y, x+y \in [0, 1]$. Since the left hand side of (12.29) is finite by the assumption, the series (12.28) is uniformly and absolutely convergent. This concludes the proof.

Corollary 12.3 *If* Λ *is a bounded function on* $\mathbb{N}$*, then each* $g \in \widetilde{C}$ *such that*

$$\sum_{n=1}^{\infty} \omega_g(2^{-n}) < \infty$$

belongs to the domain of $\Lambda(\widetilde{T})$.

Corollary 12.4 *If* g *is a Lipschitz function with an exponent* $0 < p \leq 1$, *then* g *belongs to the domain of* $\widetilde{T}$.

Proof. It follows from the definition of ω_g that if g satisfies $|g(x) - g(y)| \leq |x - y|^p$ then $\omega_g(t) \leq Kt^p$. Since $\int_0^1 t^{p-1} \log t dt > -\infty$, for $p > 0$, the condition (12.27) is satisfied.

We have already mentioned about the importance of time rescalings realized through the Λ operators defined as functions of the time operator. Somewhat different is the role of the integration transformation J, which satisfies together with T the commutation relation (12.25). Applying the transformation J on a functional basis makes approximations "smoother". We have seen already that applying J on the orthonormal Haar basis we obtain the time operator associated with approximations by continuous functions, i.e. with the Faber-Schauder basis in $\mathcal{C}_{[0,1]}$. Similarly, starting from a time operator associated with an orthogonal basis of continuous functions we can get a time operator associated with approximation in the space $\mathcal{C}^{(1)}_{[0,1]}$ of differentiable functions. As an example let us consider the Franklin system ϕ_n, $n = 0, 1, \dots$ in $\mathcal{C}_{[0,1]}$. Recall that the functions ϕ_n are obtained through the Schmidt orthonormalization of the Faber-Schauder functions [Cif,SWS]. The Franklin system is a Schauder basis in $\mathcal{C}_{[0,1]}$. We can therefore apply Proposition 1 constructing, as in Theorem 12.3, the time operator associated with a given partition of $\{\phi_n\}$ on blocks. On the other hand the system

$$1, J\phi_0, J\phi_1, \dots \tag{12.30}$$

is also a Schauder basis in $\mathcal{C}_{[0,1]}$ and each $f \in \mathcal{C}_{[0,1]}$ has the expansion

$$f(t) = f(0) + \sum_{n=0}^{\infty} a_n J\phi_n \,,$$

where $a_n = \int_0^1 \phi_n(s) df(s)$. This implies that (12.30) is also a Schauder basis in $\mathcal{C}^{(1)}_{[0,1]}$ endowed with the norm $\|f\| \stackrel{\mathrm{df}}{=} \max_{0 \leq s \leq 1} |f(s)| + \max_{0 \leq s \leq 1} |f'(s)|$. Repeating again the proof of Theorem 12.3 we obtain a time operator in $\mathcal{C}^{(1)}_{[0,1]}$ associated with this basis, which nothing but the composition of the integral transformation with time operator constructed formerly on $\mathcal{C}_{[0,1]}$.

13

Time operator and quantum theory

13.1 SELF-ADJOINT OPERATORS, UNITARY GROUPS AND SPECTRAL RESOLUTION

If A is a self-adjoint operator then there exists a unique family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ of projections such that

i) $E_{\lambda_1} \leq E_{\lambda_2}$ if $\lambda_1 \leq \lambda_2$

ii) s-$\lim_{\lambda\to-\infty} E_\lambda = 0$, s-$\lim_{\lambda\to+\infty} E_\lambda = I$

iii) $f \in D_A$ if and only if $\int_{\mathbb{R}} \lambda^2 d(f, E_\lambda f) < \infty$

iv) for $f \in D_A$ and any g

$$(g, Af) = \int \lambda \, d(g, E_\lambda f)$$

The family $\{E_\lambda\}$ satisfying (i-iv) is called the spectral family of A. We shall write $A = \int \lambda \, d\lambda$ to mean (iii) and (iv). Conversely every family $\{E_\lambda\}$ of projections satisfying (i-iii) defines a unique self-adjoint operator A given by $A = \int \lambda \, d\lambda$.

A given one-parameter unitary group $\{U_t\}_{t\in\mathbb{R}}$ is associated with an unique spectral family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ of a self-adjoint operator A such that

$$U_t = \int e^{-i\lambda t} dE_\lambda \, . \tag{13.1}$$

The operator $A = \int \lambda \, dE_\lambda$ is called the (self-adjoint) generator of the unitary group $\{U_t\}$. Thus $U_t = e^{-iAt}$ in the sense of functional calculus. The operator A can also be defined as follows:

$$f \in D_A \text{ iff } \left.\frac{d}{dt}(U_t f)\right|_{t=0} \text{ exists in the strong topology}$$

and

$$-iAf = \left.\frac{d}{dt}(U_t f)\right|_{t=0}, \text{ for } f \in D_A .$$

Every self-adjoint operator A defines a unitary group $\{U_t\}$ by (13.1), where $\{E_t\}$ is the spectral family of A.

13.2 DIFFERENT DEFINITIONS OF TIME OPERATOR AND THEIR INTERRELATIONS

Let consider the following definitions of time operator:

(a) A self-adjoint operator T is said to be a time operator for the unitary group $\{U_t\}$ of time evolution if

$$\begin{aligned} &\text{(i)} \quad U_t D_T = D_T \\ &\text{(ii)} \quad U_t^* T U_t = T + tI \text{ on } D_T . \end{aligned}$$

(b) A self-adjoint operator T with the spectral family $\{E_t\}$ is said to be a time operator for the unitary group $\{U_t\}$ if it is satisfied the *imprimitivity condition*:

$$U_t^* E_\lambda U_t = E_{\lambda - t}, \text{ for each } \lambda, t \in \mathbb{R} .$$

(c) (*Weyl commutation relation*) The self-adjoint generator T of a unitary group $\{V_s\}$, $V_s = e^{-isT}$, is a time operator for a unitary group $\{U_t\}$ if U_t and V_s satisfy the Weyl commutation relation

$$U_t V_s = e^{ist} V_s U_t, \text{ for each } s, t \in \mathbb{R}$$

(d) (*Canonical commutation relation*) Let L be the self-adjoint generator of a evolution group $\{U_t\}$, $U_t = e^{-iLt}$. Then a self-adjoint operator T is said to be a time operator of $\{U_t\}$ if there is a dense domain D on which both LT and TL are defined and

$$i[L, T] = I \text{ on } D .$$

We shall show below that the definitions (a), (b) and (c) are equivalent and imply (d). We shall also discuss the conditions under which the canonical commutation relation implies the first three definitions of time operator.

(a)⇒(b):
Let $\{E_\lambda\}$ be the spectral family of a time operator T in the sense of (a). The the spectral family of $U_t^*TU_t$ is the family $\{U_t^*E_\lambda U_t\}$:

$$U_t^*TU_t = \int \lambda \, d(U_t^*E_\lambda U_t)\,. \tag{13.2}$$

On the other hand $U_t^*TU_t = T + tI$, thus

$$T + tI = \int (\lambda + t)\, dE_\lambda = \int \lambda \, dE_{\lambda - t}\,. \tag{13.3}$$

From the uniqueness of the spectral resolution of a self-adjoint operator it follows that

$$U_t^*E_\lambda U_t = E_{\lambda - t}\,.$$

(b)⇒(a):
Suppose that the spectral family $\{E_\lambda\}$ of T satisfies the imprimitivity condition for $\{U_t\}$. We have to show that

$$U_tD_T = D_T \;\text{ and }\; U_t^*TU_t = T + tI \;\text{ on } D_T\,.$$

Now, let $f \in D_T$. Then $U_tf \in D_T$, for all real t. Indeed

$$\begin{aligned}
\int \lambda^2 \, d(U_tf, E_\lambda U_tf) &= \int \lambda^2 \, d(f, U_t^*E_\lambda U_tf) \\
&= \int \lambda^2 \, d(f, E_{\lambda - t}f) \\
&= \int (\lambda + t)^2 d(f, E_\lambda f) \\
&= \int \lambda^2 d\|E_\lambda f\|^2 + t^2 \int d\|E_\lambda f\|^2 + 2t \int \lambda \, d\|E_\lambda f\|^2\,.
\end{aligned}$$

First and the third terms are finite because $f \in D_T$ and second term equals t^2 from the properties of any resolution of identity. Thus $U_tD_T \subset D_T$ for all t. Since $U_t^* = U_{-t}$, we also have $U_t^*D_T \subset D_T$. Therefore $D_T = U_t(U_t^*D_T) \subset U_tD_T$, which proves the equality $U_tD_T = D_T$.

To verify that $U_t^*TU_t = T + tI$ on D_T it is enough to consider the scalar product $(g, U_t^*TU_tf)$ for $f \in D_T$ and any g. We have

$$\begin{aligned}
(g, U_t^*TU_tf) &= \int \lambda \, d(g, U_t^*E_\lambda U_tf) \\
&= \int \lambda \, d(g, E_{\lambda - t}f) \\
&= \int (\lambda + t) \, d(g, E_\lambda f) \\
&= \int \lambda \, d(g, E_\lambda f) + t \int d(g, E_\lambda f) \\
&= (g, Tf) + t(g, f)\,,
\end{aligned}$$

which proves (ii).

(b)⇒(c):
Let $\{E_\lambda\}$ be the spectral projectors of T and consider the unitary operators $V_s = \int e^{-i\lambda s} dE_\lambda$. We have

$$\begin{aligned} U_t^* V_s U_t &= \int e^{-i\lambda s} d(U_t^* E_\lambda U_t) \\ &= \int e^{-i\lambda s} dE_{\lambda - t} \\ &= \int e^{-i(\lambda+t)s} dE_\lambda \\ &= e^{-its} V_s \,. \end{aligned}$$

Thus $V_s U_t = e^{-its} U_t V_s$, for all t. Taking the Hermitian conjugate we get $U_t V_s = e^{its} V_s U_t$, for all t.

(c)⇒(b):
Assume that $U_t V_s = e^{its} V_s U_t$, for all s, t, hence

$$U_t V_s U_t^* = e^{its} V_s \,. \tag{13.4}$$

The left hand side of (13.4) has, for a fixed t the representation

$$U_t V_s U_t^* = \int e^{-i\lambda s} d(U_t E_\lambda U_t^*) \,.$$

On the other hand the unitary group $e^{its} V_s$ is given by

$$e^{its} V_s = \int e^{its} e^{-\lambda s} dE_\lambda = \int e^{-is(\lambda - t)} dE_\lambda = \int e^{-is\lambda} dE_{\lambda + t} \,.$$

By the uniqueness of spectral family and (13.4) we have $U_t E_\lambda U_t^*$, for all t.

We have proved that definitions (a), (b) and (c) of time operator are equivalent. Note also that the Weyl commutation relation can be written

$$V_t^* U_t V_s = e^{ist} U_t \,. \tag{13.5}$$

If $\{F_\lambda\}$ is the spectral family of the self-adjoint generator L of the evolution group $\{U_t\}$, $U_t = \int e^{-i\lambda t} dF_\lambda$, then it follows from (13.5) that

$$V_s^* F_\lambda V_s = F_{\lambda + s} \,, \quad \text{for all } s \,, \tag{13.6}$$

where $V_s = e^{-isT}$. Thus the spectral family of L satisfies the imprimitivity condition with respect to the unitary group $\{V_s\}$, $V_s^* = e^{isT}$.

Because the spectral family $\{F_\lambda\}$ of L satisfies the imprimitivity condition (13.6) with respect to V_s we can again verify as before that $V_s D_L = D_L$ and $V_s^* L V_s = L - sI$ on D_L, for all $s \in \mathbb{R}$. There is thus a "duality" between L and T.

(c)⇒(d):
Let L be a self-adjoint generator of the group $\{U_t\}$. Then the Weil commutation relation assumes the form

$$e^{-itL}e^{-isT} = e^{ist}e^{-isT}e^{-itL}\,, \quad \text{for each } s,t \in \mathbb{R}\,. \tag{13.7}$$

Multiplying (13.7) by e^{-s} we get

$$e^{-itL}e^{s(-iT-I)} = e^{s(-iT+itI-I)}e^{-itL}\,. \tag{13.8}$$

Then the integration (13.8) with respect to s on $(0,\infty)$ gives

$$e^{-itL}(-iT-I)^{-1} = (-iT+itT-I)^{-1}\,e^{-itL}\,. \tag{13.9}$$

After taking the derivative of (13.9) in the point $t=0$ we have

$$-iL(-iT-I)^{-1} = -i(-iT-I)^{-2} + (-i)(-iT-I)^{-1}L$$

or equivalently

$$L(T-iI)^{-1} = (T-iI)^{-1}L + i(T-iI)^{-2}\,. \tag{13.10}$$

Adding to both sides of (13.10) $-i(T-iI)^{-1}$ and then multiplying from the left and from the right by $(L-iI)^{-1}$ we get

$$(T-iI)^{-1}(L-iI)^{-1} = (L-iI)^{-1}(T-iI)^{-1} + i(L-iI)^{-1}(T-iI)^{-2}(L-iI)^{-1}\,. \tag{13.11}$$

Let y be an element of $\mathcal{H}$ and put

$$x = (T-iI)^{-1}(L-iI)^{-1}y\,. \tag{13.12}$$

Then $x \in D_{(L-iI)(T-iI)}$ and

$$y = (L-iI)(T-iI)x\,. \tag{13.13}$$

It follows from (13.11) that

$$x = (L-iI)^{-1}(T-iI)^{-1}(y+ix)\,. \tag{13.14}$$

This implies that $x \in D_{(T-iI)(L-iI)}$ and

$$(T-iI)(L-iI)x = (L-iI)(T-iI)x + ix\,. \tag{13.15}$$

Then, summarizing, we obtain that $x \in D_{LT} \cap D_{TL}$ and

$$(TL-LT)x = ix\,. \tag{13.16}$$

It is obvious that any $x \in D_{LT} \cap D_{TL}$ can be expressed in the form (13.12) (by letting y as in (13.13)). Thus (13.16) holds for each $x \in D_{LT} \cap D_{TL}$ or equivalently

$$[L,T]x = -ix\,, \quad \text{for each } x \in D_{LT} \cap D_{TL}\,, \tag{13.17}$$

which concludes the proof of (c)⇒(d).

Notice that we have actually proved that

$$(T - iI)(L - iI)(D_{LT} \cap D_{TL}) = (L - iI)(T - iI)(D_{LT} \cap D_{TL}) = \mathcal{H}\,.$$

This means, in particular, that the image of a dense subset Ω of $D_{LT} \cap D_{TL}$ (here $\Omega = D_{LT} \cap D_{TL}$) through both $(T - iI)(L - iI)$ and $(L - iI)(T - iI)$ is dense in $\mathcal{H}$. It can be proved [FGN,Pu] that the latter condition is also sufficient for the canonical commutation relation to imply Weyl commutation relation. In fact in order to prove that (d)⇒(c) it is enough to assume that there exists a dense set $\Omega \subset D_{LT} \cap D_{TL}$ for which either $(T - iI)(L - iI)\Omega$ or $(L - iI)(T - iI)\Omega$ is dense in $\mathcal{H}$.

13.3 SPECTRUM OF L AND T

If $U_t = e^{-iLt}$ and T is defined by any of the equivalent conditions (a), (b) or (c) important restrictions on the spectrum of L and T are implied. First, from the condition (a) it follows that

$$(U_t f T, U_t) = (f, Tf) + t(f, f)\,, \quad \text{for any } f \in D_T\,. \tag{13.18}$$

Therefore, by taking t in (13.18) to be suitably large, positive or negative number, expectation value of T in the state $U_t f$ may be made equal to any positive or negative value. Thus the spectrum of T can not be bounded from below or above. Similar conclusion follows for L because of "duality"between T and L. Imprimitivity condition implies more. It implies that the spectrum of T (and L) are whole of real line and spectral multiplicity is uniform. With the aid of Plesner's theorem (Plesner 1929) the imprimitivity condition implies that the spectra of T and L are absolutely continuous.

13.4 INCOMPATIBILITY BETWEEN THE SEMIBOUNDENESS OF THE GENERATOR H OF THE EVOLUTION GROUP AND THE EXISTENCE OF A TIME OPERATOR CANONICAL CONJUGATE TO H

In standard formulation of quantum mechanics the (pure) states of the system are represented by (unit) vectors ψ of a Hilbert space $\mathcal{H}$. Time evolution group is given by $\{e^{-iHt}\}$, where H is the Hamiltonian that represents the energy observable and hence must be bounded from below. Previous remarks about the spectrum of the generator of the evolution group admitting a time operator in the sense of (a), (b) or (c) show that $\{e^{-iHt}\}$ can not admit a time operator in this sense because H is semibounded. The time-energy uncertainty relation can however be derived from the existence of time operator T which is canonically conjugate to L:

$$i[H, T] = I\,, \quad \text{on } D\,. \tag{13.19}$$

As noted before this canonical commutation relation does not imply condition (a), (b) or (c) on D. Since a main motivation for introducing a time operator in quantum mechanics is to give a theoretical foundation to time-energy uncertainty relation (Pauli) we shall give a simple proof that the existence a time operator T in the weaker sense that it satisfies (13.19) on a suitable domain D is incompatible with the semiboundeness of H.

For this purpose we shall not suppose that D is dense in $\mathcal{H}$. But we shall suppose that there is a vector $\psi \in D$ such that $e^{itT}\psi \in D$, for each $t \geq 0$.

Consider now the function $t \longmapsto (e^{iTt}\psi, He^{iTt}\psi)$, for some $\psi \neq 0$, such that $e^{iTt}\psi \in D$ for all t. This function is differentiable with respect to t with the derivative:

$$\begin{aligned} \tfrac{d}{dt}\left(e^{iTt}\psi, He^{iTt}\psi\right) &= -i\left(e^{iTt}\psi, THe^{iTt}\psi\right) + i\left(e^{iTt}\psi, HTe^{iTt}\psi\right) \\ &= \left(e^{iTt}\psi, He^{iTt}\psi\right) \\ &= \left(\psi, [H,T]\,\psi\right) \equiv \|\psi\|^2 \,. \end{aligned}$$

Integrating both sides from 0 to $t > 0$ we get

$$\left(e^{iTt}\psi, He^{iTt}\psi\right) - (\psi, H\psi) = t\|\psi\|^2$$

or

$$(\psi, H\psi) = -t\|\psi\|^2 + \left(e^{iTt}\psi, He^{iTt}\psi\right) \,.$$

Since $H \geq 0$, $\left(e^{iTt}\psi, He^{iTt}\psi\right) \geq 0$. Thus $(\psi, H\psi) \leq -t\|\psi\|^2$, which is contradiction because $(\psi, H\psi)$ is a constant finite number.

Thus in standard formulation of quantum mechanics there can not exist a time operator even in the weaker sense – that is canonically conjugate to the generator of time evolution group.

13.5 LIOUVILLE-VON NEUMANN FORMULATION OF QUANTUM MECHANICS

The previous argument shows that a time operator can not be defined in the usual formulation of quantum mechanics in any of the senses (a), (b), (c) or (d). To define a time operator we need to go to an extended formulation of quantum dynamics in which the generator of time-evolution group is not necessarily semibounded. This is the well known Liouville-von Neumann formation given in terms of the evolution of more general states than the pure states.

First some general remarks about Hilbert-Schmidt (H.S.) trace class (called also nuclear) operators. A bounded operator A in Hilbert space $\mathcal{H}$ is called a H.S. operator if

$$\sum_{k=1}^{\infty} \|A\phi_k\|^2 \equiv \|A\|_2^2 < \infty \,.$$

It can be shown that $\|A\|_2$ (called H.S. norm of A) is independent of the choice of orthonormal basis $\{\phi_k\}$. The class of H.S. operators is a linear space. If A is H.S.

operator so is A^* and $\|A^*\|_2 = \|A\|_2$. If B is a bounded operator and A H.S. then both BA and AB are again H.S.. We can define an inner product in the space $\mathcal{B}_2(\mathcal{H})$ of H.S. by putting for $A, B \in \mathcal{B}_2(\mathcal{H})$

$$(A, B) = \sum_{k=1}^{\infty} (\psi_k, A^* B \psi_k)\,,$$

where $\{\psi_k\}$ is an orthonormal basis in $\mathcal{H}$. It can be shown that the above defined inner product does not depend on the choice of orthonormal basis $\{\psi_k\}$. We can write

$$(A, B) = \mathrm{Tr}\,(A^* B)\,, \quad \text{for } A, B \in \mathcal{B}_2(\mathcal{H})\,.$$

With this inner product $\mathcal{B}_2(\mathcal{H})$ is a Hilbert space. If S_1 and S_2 are two bounded operator in $\mathcal{H}$ they define a bounded operator, denoted by $S_1 \times S_2$, on $\mathcal{B}_2(\mathcal{H})$ as follows:

$$(S_1 \times S_2)A \equiv S_1 A S_2\,, \quad \text{for } A \in \mathcal{B}_2(\mathcal{H})\,.$$

If S_1 or S_2 are not bounded then $S_1 \times S_2$ does not define a bounded operator but defines an unbounded operator in $\mathcal{B}_2(\mathcal{H})$ whose domain consists of all $A \in \mathcal{B}_2(\mathcal{H})$ is again in $\mathcal{B}_2(\mathcal{H})$. Operators in $\mathcal{B}_2(\mathcal{H})$ of the form $S_1 \times S_@$ are called *factorizable* operators of $\mathcal{B}_2(\mathcal{H})$. There exist, of course, operators in $\mathcal{B}_2(\mathcal{H})$ which are not factorizable. An operator A is said to be of trace class if $A = S_1 S_2$, where both S_1 and S_2 are H.S.. If A is trace class, $\mathrm{Tr}\,A = \sum_k (\phi_k, A\phi_k) < \infty$, where $\{\psi_k\}$ is an orthonormal basis, then it is independent of basis $\{\psi_k\}$. Trace operators form a linear manifold of $\mathcal{B}_2(\mathcal{H})$. If A is trace class and B is bounded then both AB and BA are trace class and $\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$. If A is non negative (and hence also self-adjoint) trace class operator then there exists an orthonormal family $\{\phi_k\} \subset \mathcal{H}$ and positive numbers λ_k, $k = 1, 2, \ldots$, with $\sum_k \lambda_k < \infty$ such that

$$A = \sum_{k=1}^{\infty} \lambda_k |\phi_k\rangle\langle\phi_k|\,. \tag{13.20}$$

Here we have used Dirac's notation $|\phi_k\rangle\langle\phi_k|$ for the projection onto the one dimensional subspace of $\mathcal{H}$ generated by ϕ_k.

If A is of the form (13.20) then $\mathrm{Tr}\,A = \sum_k \lambda_k$. If A is a positive trace class operator then the square root $A^{1/2}$ of A is a Hilbert Schmidt operator and $\mathrm{Tr}\,A = (A^{1/2}, A^{1/2})$. Positive trace class operators ρ with $\mathrm{Tr}\,\rho = 1$ will be called *density operators*.

Elementary observables (or *propositions*) with only two outcomes: 1 or 0 (*yes* and *no*) are represented by projection operators of the Hilbert space of the system. Any other observable (self-adjoint operator) can be expressed in terms of the propositions through the spectral theorem. For a given state of quantum system measurement of a proposition E is not generally a definite outcome. The outcome 'yes' will occur with certain probability (or expectation) $p(E)$. A quantum state is thus specified by the probability $p(E)$ for every projection E. In other words, a quantum state is given by the mapping $E \longmapsto p(E)$, for all projections E. Obviously

(i) $0 \le p(E) \le 1$, for all E

(ii) $p(I) = 1$, where I is the identity operator

(iii) $p(0) = 0$

(iv) If $\{E_i\}$ is a countable family of mutually orthogonal projections then they represent mutually orthogonal propositions and $\sum_i E_i$ represents the proposition whose measurement yields the outcome 1 (or 'yes') if and only if the outcome of measurement of one of the proposition E_i is 1. Thus we will require that for mutually orthogonal propositions holds $p(\sum_i E_i) = \sum_i p(E_i)$.

Condition (iv) implies that if $E_1 > E_2$ then $p(E_1) \geq p(E_2)$.

It has been shown (Gleason 1957, Jost 1976) that every mapping $E \longmapsto p(E)$ from the class of projections E satisfying conditions (i)–(iv) corresponds to a density operator ρ such that $p(E) = \mathrm{Tr}\,(\rho E)$. Quantum states are thus represented by density operators ρ. We can obviously write the probability (or expectation) of E in the state ρ as

$$\mathrm{Tr}\,(\rho E) = \left(\rho^{1/2}, E\rho^{1/2}\right) .$$

If A is an observable, i.e. a self-adjoint operator with the spectral resolution $A = \int \lambda\, dE_\lambda$ then the expectation $\langle A \rangle_\rho$ of A in the state ρ is given by

$$\langle A \rangle_\rho = \int_\lambda d\left(\rho^{1/2}, E_\lambda \rho^{1/2}\right) = \left(\rho^{1/2}, A\rho^{1/2}\right) = \mathrm{Tr}\,(A\rho)\,,$$

if $A\rho^{1/2}$ is defined and H.S..

A density operator ρ is a one dimensional projection operator $|\phi\rangle\langle\phi|$ (with $\|\phi\| = 1$) if and only if $\rho^2 = \rho$. density operators ρ with $\rho^2 = \rho$ thus represent pure states ϕ. More generally, as said before, a density ρ is of the form

$$\rho = \sum_{k=1}^{\infty} \lambda_k |\phi_k\rangle\langle\phi_k|\,,$$

with $\lambda_k \geq 0$, $\sum_k \lambda_k = 1$ and $\{\phi_k\}$ an orthonormal family of vectors. Such a state ρ is thus a ***mixture*** of mutually orthogonal pure states ϕ_k with relative weights λ_k. The expectation of an observable A fro the pure state $\rho = |\phi\rangle\langle\phi|$ is $\mathrm{Tr}\,(A\rho) = \mathrm{Tr}\,(|A\phi\rangle\langle\phi|) = (\phi, A\phi)$ as required.

Regarding time evolution of density operators we start from the time evolution of pure states $\rho = |\phi\rangle\langle\rho|$. Under time evolution ϕ evolves in time t to $e^{-iHt}\phi$. Therefore ρ evolves to $|e^{-iHt}\phi\rangle\langle e^{-iHt}\phi| = e^{-iHt}\,(|\phi\rangle\langle\phi|)\,e^{iHt}$. A more general mixture $\rho = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ will thus evolve to $\sum_k |e^{-iHt}\phi_k\rangle\langle e^{-iHt}\phi_k| = e^{-iHt}\rho e^{iHt}$.

The class of density operators is not even a linear space. To take advantage of operator theory in Hilbert space we shall extend the time evolution to all Hilbert-Schmidt operators $\mathcal{B}_2(\mathcal{H})$. If A is any H.S. operator we shall define time evolution of A by

$$A \longmapsto e^{-iHt} A e^{iHt}$$

in time t. This extension is consistent with the evolution of density operators. Note that the evolution of H.S. operator preserves the inner product in $\mathcal{B}_2(\mathcal{H})$. In fact

$$\begin{aligned}\left(e^{-iHt}Ae^{iHt}, e^{-iHt}Be^{iHt}\right) &= \mathrm{Tr}\left[\left(e^{-iHt}Ae^{iHt}\right)^* \left(e^{-iHt}Be^{iHt}\right)\right] \\ &= \mathrm{Tr}\left(e^{-iHt}A^*Be^{iHt}\right) \\ &= \mathrm{Tr}\left(A^*B\right) \\ &= (A,B)\,,\end{aligned}$$

for any $A, B \in \mathcal{B}_2(\mathcal{H})$. Therefore the group $\{U_t\}$ defined by

$$U_tA = e^{-iHt}Ae^{iHt}$$

is an unitary group in $\mathcal{B}_2(\mathcal{H})$. This is the time evolution group in the Liouville-von Neumann formulation of quantum mechanics.

Since $\{U_t\}$ is an unitary group in the Hilbert space $\mathcal{B}_2(\mathcal{H})$ it has a self-adjoint operator $L \in \mathcal{B}_2(\mathcal{H})$, called the *Liouvillian*) as the generator of $\{U_t\}$, i.e. $U_t = e^{-iLt}$. From the definition of U_t it follows that

$$LA = [H, A]\,,$$

for all $A \in \mathcal{B}_2(\mathcal{H})$ for which the right hand side is also in $\mathcal{B}_2(\mathcal{H})$. Obviously $L = H\times I - I\times$ and it is not a factorizable operator of $\mathcal{B}_2(\mathcal{H})$, but a linear combination of them.

Unlike the Hamiltonian H of $\mathcal{H}$, the spectrum of L is no longer required to be bounded from below. In fact, we shall find that if H has absolutely continuous spectrum extending over the entire interval $[0,\infty)$ then the spectrum of L is absolutely continuous and of uniform (in fact infinite) multiplicity extending over the entire real line. In this situation we can find a time operator T on $\mathcal{B}_2(\mathcal{H})$ with respect to $\{e^{-iLt}\}$ in the sense of (a), (b) or (c). As said before this time operator T will satisfy the canonical commutation relation $i[L,T] = I$ on a dense domain of $\mathcal{B}_2(\mathcal{H})$.

Before discussing the construction of time operator T in this case let us show that the canonical commutation relation between L and T implies a time-energy uncertainty relation.

13.6 DERIVATION OF TIME ENERGY UNCERTAINTY RELATION

We shall derive now time energy uncertainty relation from the commutation relation between the Liouvillian and the time operator. Since expectation value $\langle A\rangle_\rho$ of usual observables A (i.e. self-adjoint operators in $\mathcal{H}$) in the state ρ is given by

$$\langle A\rangle_\rho = \mathrm{Tr}\,(A\rho) = \left(\rho^{1/2}, A\rho^{1/2}\right)$$

we shall define the expectation value of non factorizable operators such as L and T by the same formula:

$$\langle L\rangle_\rho = \left(\rho^{1/2}, L\rho^{1/2}\right) \;\text{ and }\; \langle T\rangle_\rho = \left(\rho^{1/2}, T\rho^{1/2}\right)\,.$$

From the canonical commutation relation between L and T we first an uncertainty relation between L and T in the standard manner.

Now, mean square deviations $\langle \Delta L\rangle_\rho^2$, $\langle \Delta T\rangle_\rho^2$ of L and T respectively in the state ρ, for which $\rho^{1/2}$ is in the domain of L, L^2, T and T^2 is given by

$$\langle \Delta T\rangle_\rho^2 = \left(\rho^{1/2}, T^2\rho^{1/2}\right) - \left(\rho^{1/2}, T\rho^{1/2}\right) = \left(\rho^{1/2}, T'^2\rho^{1/2}\right) ,$$

where $T'^2 = T - (\rho^{1/2}, T\rho^{1/2})I$. Similarly

$$\langle \Delta L\rangle_\rho^2 = \left(\rho^{1/2}, L^2\rho^{1/2}\right)$$

because $(\rho^{1/2}, L\rho^{1/2}) = 0$, as $L\rho^{1/2} = H\rho^{1/2} - \rho^{1/2}H$. Since T' differs from T by a multiple of identity $i[L, T'] = i[L, T] = I$.

Now let $S = T' - i\lambda L$ (λ real). Then $S^* = T' + iL$ in the appropriate domain as T' and L are. We have

$$\left(S\rho^{1/2}, S\rho^{1/2}\right) \geq 0 .$$

On the other hand

$$\begin{aligned}\left(S\rho^{1/2}, S\rho^{1/2}\right) &= \left(\rho^{1/2}, S^* S\rho^{1/2}\right)\\ &= \left(\rho^{1/2}, (T' + i\lambda L)\,(T' - i\lambda L)\,\rho^{1/2}\right)\\ &= \left(\rho^{1/2}, T'^2\rho^{1/2}\right) + \lambda\left(\rho^{1/2}, i[L, T']\rho^{1/2}\right) + \lambda^2\left(\rho^{1/2}, L^2\rho^{1/2}\right) .\end{aligned}$$

Since $i[L, T'] = i[L, T] = I$ and $(\rho^{1/2}, \rho^{1/2}) = \mathrm{Tr}\,\rho = 1$, we get

$$\left(S\rho^{1/2}, S\rho^{1/2}\right) = \langle \Delta T\rangle_\rho^2 + \lambda + \lambda^2\langle \Delta T\rangle_\rho^2 \geq 0 .$$

positiveness of this real quadratic form in λ implies

$$\langle \Delta T\rangle_\rho^2\langle \Delta L\rangle_\rho^2 \geq \frac{1}{4}$$

or

$$\langle \Delta T\rangle_\rho\langle \Delta L\rangle_\rho \geq \frac{1}{2} .$$

To derive the time energy uncertainty relation $\langle \Delta L\rangle_\rho$ and $\langle \Delta H\rangle_\rho \equiv (\mathrm{Tr}\,(\rho H^2) - (\mathrm{Tr}\,\rho H)^2)^{1/2}$.

Recall that $L = H \times I - I \times H$. Since $H \times I$ and $I \times H$ commute with each other and $(H \times I)\cdot(I \times H) = H \times H$, $L^2 = H^2 \times I - 2H \times H + I \times H^2$, we have

$$\begin{aligned}\langle \Delta L\rangle_\rho^2 &= \left(\rho^{1/2}, [H^2 \times I - 2H \times H + I \times H^2]\rho^{1/2}\right)\\ &= \mathrm{Tr}\,\left(H^2\rho\right) - 2\left(H\rho^{1/2}, H\rho^{1/2}\right) + \mathrm{Tr}\,\left(H^2\rho\right)\end{aligned} .$$

In the above equality we used the fact that both $H\rho^{1/2}$ and $H^2\rho^{1/2}$ are H.S.. It is only for such states ρ that $\langle \Delta L\rangle_\rho^2$ is finite and uncertainty relation is nontrivial.

Thus

$$\langle \Delta L\rangle_\rho^2 = 2\left[\mathrm{Tr}\,(H^2\rho) - [\mathrm{Tr}\,(H\rho)]^2 + (\rho^{1/2}, H\rho^{1/2})^{1/2} - (H\rho^{1/2}, H\rho^{1/2})\right]$$

($[\mathrm{Tr}\,(H\rho)]^2$ being equal to $(\rho^{1/2}, H\rho^{1/2})^{1/2}$ when $H(\rho^{1/2}$ is H.S.). Hence

$$\langle\Delta L\rangle_\rho^2 = 2\langle\Delta H\rangle_\rho^2 + \left[(\rho^{1/2}, H\rho^{1/2})^2 - (H\rho^{1/2}, H\rho^{1/2})\right] .$$

From Cauchy-Schwartz inequality

$$(\rho^{1/2}, H\rho^{1/2})^2 \leq (H\rho^{1/2}, H\rho^{1/2})(\rho^{1/2}, \rho^{1/2}) = (H\rho^{1/2}, H\rho^{1/2}) \,,$$

since $(\rho^{1/2}, \rho^{1/2}) = \mathrm{Tr}\,\rho = 1$. This implies

$$(\rho^{1/2}, H\rho^{1/2})^2 - (H\rho^{1/2}, H\rho^{1/2}) \leq 0$$

and consequently

$$\langle\Delta L\rangle_\rho^2 \leq 2\langle\Delta H\rangle_\rho^2 \text{ or } \langle\Delta L\rangle_\rho \leq \sqrt{2}\langle\Delta L\rangle_\rho \,,$$

The uncertainty $\langle\Delta L\rangle_\rho\langle\Delta T\rangle_\rho \geq 1/2$ then yields the time-energy uncertainty relation

$$\langle\Delta H\rangle_\rho\langle\Delta T\rangle_\rho \geq \frac{1}{2\sqrt{2}} \,.$$

Note that defining the evolution of ρ by $\mathcal{U}_t = e^{-iHt} \times e^{iHt}$ we have put $\hbar = 1$, hence the Plank constant $\hbar$ does not appear in the time-energy uncertainty relation.

Although it is interesting that root mean square deviation $\langle\Delta H\rangle_\rho$ of the time operator T in the state ρ and root mean square deviation $\langle\Delta H\rangle_\rho$ of the energy operator H satisfy an uncertainty relation, the physical interpretation of this uncertainty is not entirely clear. Usually $(\rho^{1/2}, T\rho^{1/2})$ is interpreted as the average *age* of the state ρ. Therefore $\langle\Delta T\rangle_\rho$is the root mean square deviation of the outcomes of measurements of age of the system in the state ρ. But, up to now no operational meaning of such measurement is given. It would be more satisfactory if $\langle\Delta T\rangle_\rho$ could be interpreted as root mean square deviation of "time of decay" or time of transitions of an unstable state. To decide if such an interpretation can be given to time operator, it would be necessary to study the time operator in a model of unstable system, such as the Friedrichs-Lee model. This, however, has not yet been done. Moreover, the discussions of "arrival time" (G.R. Allock, *Ann. Phys. N.Y.* **53**, 251 (1969)) and quantum Zeno effect (Misra, Sudarshan) indicate that such an interpretation may not be immediate. But let us recall that Einstein felt the need for a concept concerning time instant of decay in the theoretical description (ed. Paul A. Schilpp, *Einstein: Philosopher-Scientist*, vol II, pp.665-688). Perhaps, the concept of time operator or some suitable modification of it will satisfy this need.

13.7 CONSTRUCTION OF SPECTRAL PROJECTIONS OF THE TIME OPERATOR T

The Hamiltonian H acting on the space $\mathcal{H}$, the Hilbert space of pure states, has absolutely continuous spectrum over the entire halfline $[0, \infty)$. Let us first suppose that the spectrum of H is of simple spectral multiplicity. Then we can consider the spectral representation of H. This is a unitary mapping from $\mathcal{H}$ onto $L^2(\mathbb{R}_+, d\mu)$ such that if under this mapping

$$\mathcal{H} \ni \psi \longmapsto \psi(\cdot) \in L^2(\mathbb{R}_+, d\mu)$$

then H is represented by the multiplication operator by μ, i.e.

$$(H\psi)(\mu) = \mu\psi(\mu) .$$

An arbitrary operator $\rho \in \mathcal{B}_2(\mathcal{H})$ goes under this mapping to an integral operator on in $L^2(\mathbb{R}_+, d\mu)$ represented by the kernel $\rho(\mu, \mu')$ such that

$$\int_{\mathbb{R}_+ \times \mathbb{R}_+} |\rho(\mu, \mu')|^2 \, d\mu \, d\mu' < \infty .$$

This correspondence between $\rho \in \mathcal{B}_2(\mathcal{H})$ and kernel $\rho(\mu, \mu')$ is a unitary mapping from $\mathcal{B}_2(\mathcal{H})$ onto $L^2(\mathbb{R}_+ \times \mathbb{R}_+, d\mu \, d\mu')$. Under this mapping

$$e^{-iHt} \times e^{iHt} \equiv e^{-iLt} :\longmapsto e^{-i(\mu-\mu')t}\rho(\mu, \mu')$$

and

$$L : \rho \longmapsto (\mu - \mu')\rho(\mu, \mu') .$$

We shall construct the spectral projections of time operator T with respect to $\{e^{iLt}\}$ in a related representation of $\mathcal{B}_2(\mathcal{H})$.

Let us consider a change of variables $(\mu, \mu') \mapsto (\lambda, \nu)$ given by $\lambda = \mu - \mu'$, $\nu = 1/2(\mu+\mu') - 1/2|\mu-\mu'|$ and the inverse transformation $\mu = \nu + 1/2|\lambda| + 1/2\lambda$, $\mu' = \nu + 1/2|\lambda| - 1/2\lambda$. As μ and μ' range over $\mathbb{R}_+$, λ ranges over all$\mathbb{R}$ and ν over $\mathbb{R}_+$ independently. since the Jacobian of transformation $(\mu, \mu) \mapsto (\lambda, \nu)$ is unity, it maps $L^2(\mathbb{R}_+ \times \mathbb{R}_+, d\mu \, d\mu')$ onto $(\mathbb{R} \times \mathbb{R}_+, d\lambda \, d\nu)$ unitarily. Under this change of variables $\rho(\mu, \mu') \mapsto \rho)\lambda, \nu) \equiv \rho(\nu + 1/2|\lambda| + 1/2\lambda, \nu + 1/2|\lambda| - 1/2\lambda)$. Any operator $\rho \in \mathcal{B}_2(\mathcal{H})$ is thus represented by a kernel $\rho(\lambda, \nu) \in L^2(\mathbb{R} \times \mathbb{R}_+, d\lambda \, d\nu)$ and vice versa. e^{-iLt} is then represented by $e^{-i\lambda t}\rho(\lambda, \nu)$ and $L\rho$ by $\lambda\rho(\lambda, \nu)$.

Let us define F_s, s-real, by

$$(F_s\rho)(\lambda, \nu) = \frac{1}{2}\rho(\lambda, \nu) + \frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty} \frac{e^{is(x-\lambda)}}{x - \lambda}\rho(x, \nu) \, dx . \tag{13.21}$$

In particular

$$(F_0\rho)(\lambda, \nu) = \frac{1}{2}\rho(\lambda, \nu) + \frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty} \frac{\rho(x, \nu)}{x - \lambda} \, dx$$

or

$$F_0 = \frac{1}{2}\left(I - i\hat{H}\right)$$

on the variable λ for each fixed ν. Here $\hat{H}$ denotes the Hilbert transform (Titchmarch). We shall verify that the family $\{F_s\}$ has the properties of spectral projection and satisfies the imprimitivity condition with respect to $\{e^{-iLt}\}$.

Indeed, it is known that the operator $\hat{H}$ is unitary and $\hat{H}^2 = I$, so that $\hat{H}^* = -\hat{H}$. Thus

$$F_0^* = 1/2(I - i\hat{H})^* = 1/2(I + i\hat{H}^*) = 1/2(I - i\hat{H}) = F_0$$

$$F_0^2 = 1/2(I - i\hat{H})^2 = 1/4(I - 2i\hat{H} + \hat{H}^2) = 1/2(I - i\hat{H}) = F_0\,.$$

F_0 is thus a projection. That F_s is, for any real s, a projection follows from the imprimitivity condition

$$e^{-iLt}F_s e^{iLt} = F_{s+t}\,,$$

which is verified below:

$$\begin{aligned}
\left(e^{-iLt}F_s e^{iLt}\rho\right)(\lambda,\nu) &= e^{-i\lambda t}\left(F_s e^{iLt}\rho\right)(\lambda,\nu)\\
&= e^{-i\lambda t}\left[\frac{1}{2}\left(e^{iLt}\rho\right)(\lambda,\nu)\right.\\
&\qquad\left.+\frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty}\frac{e^{is(x-\lambda)}}{x-\lambda}\left(e^{iLt}\rho\right)(x,\nu)\,dx\right]\\
&= e^{-i\lambda t}\left[\frac{1}{2}e^{i\lambda t}\rho(\lambda,\nu)+\frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty}\frac{e^{i(s+t)x-i\lambda s}}{x-\lambda}\rho(x,\nu)dx\right]\\
&= \frac{1}{2}\rho(\lambda,\nu)+\frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty}\frac{e^{i(t+s)(x-\lambda)}}{x-\lambda}\rho(x,\nu)\,dx\\
&= F_{s+t}\,.
\end{aligned}$$

We need to verify that $F_s \leq F_t$, if $s \leq t$. For this it suffices to prove that $F_0 \leq F_t$, for $t \geq 0$ or $F_0 F_t = F_0$.

We have to show that if $(F_0\rho)(\lambda,\nu) = \rho(\lambda,\nu)$ then $(F_t\rho)(\lambda,\nu) = \rho(\lambda,\nu)$, for $t \geq 0$. For this we need the following result (Titchmarch):

> A necessary and sufficient condition for a function $\phi \in L^2$ to be the boundary function of an analytic function $\tilde{\phi}$ in the upper halfplane
>
> $$\phi(\lambda) = \lim_{\varepsilon\to 0^+}\tilde{\phi}(\lambda + i\varepsilon)\,,\quad \text{for almost all } \lambda \in \mathbb{R}\,,$$
>
> such that
>
> $$\sup_{\nu>0}\int_{-\infty}^{\infty}|\tilde{\phi}(\lambda + i\nu)|^2 < \infty$$
>
> is that ϕ is of the form
>
> $$\phi(\lambda) = (I - i\hat{H})f\,,\quad \text{for some } f \in L^2(\mathbb{R})\,.$$

The function ϕ that satisfies the above property is called a Hardy function, $\phi \in \mathcal{H}_+^2$.

Now if $(F_0\rho)(\lambda,\nu) = \rho(\lambda,\nu)$ then $\rho(\lambda,\nu) = 1/2\rho(\lambda,\nu) - i/2(\hat{H}\rho)(\lambda,\nu)$, for each fixed $\nu \in \mathbb{R}_+$ (from the definition of F_0). Thus $\rho(\cdot,\nu)$ satisfies the condition mentioned above. Hence $\rho(\lambda,\nu)$, as a function of λ, for each fixed ν is the limit as $\Im z \to 0$ of an analytic function $\tilde{\rho}(z,\nu)$ such that $\int_{-\infty}^{\infty} |\tilde{\rho}(\lambda + i\sigma,\nu)|^2 d\lambda < K < \infty$, with $\sigma > 0$ ($K = K(\nu)$).

If $t > 0$ then $e^{\lambda} t\rho(\lambda,\nu)$ is also the limit as $\Im z \to 0$ of the analytic function $e^{izt}\rho(z,\nu)$ in the upper halfplane and

$$\int_{-\infty}^{\infty} \left| e^{i(\lambda+i\sigma)t}\rho(\lambda+i\sigma,\nu)\right|^2 d\lambda = e^{-2\sigma t}\int_{-\infty}^{\infty} |\rho(\lambda+i\sigma,\nu)|^2\, d\lambda \le K\,.$$

Hence by the above characterization of Hardy functions $e^{i\lambda t}$ is of the form :

$$e^{i\lambda t}\rho(\lambda,\nu) = f_\nu(\lambda) - i\left[\hat{H}(f_\nu)\right](\lambda)$$

with $f_\lambda \in L^2(\mathbb{R})$, for each $\nu \in \mathbb{R}_+$. But this shows that $e^{i\lambda t}\rho(\lambda,\nu)$ belongs to the range of F_0. On the other hand $e^{i\lambda t}\rho(\lambda,\nu) = (e^{iLt}\rho)(\lambda,\nu)$. We have then

$$F_0 e^{iLt}\rho = e^{iLt}\rho$$

or

$$e^{-iLt}F_0 e^{iLt}\rho = F_t\rho = \rho\,,$$

for $t \ge 0$. This proves that if $F_0\rho = \rho$ then $F_t\rho = \rho$, for $t \ge 0$. This means

$$F_tF_0 = F_0\,,\quad \text{for } t \ge 0\,.$$

It remains to verify that

$$\lim_{s\to\infty} F_s \equiv F_{-\infty} = 0 \quad\text{and}\quad \lim_{s\to\infty} F_s \equiv F_{\infty} = I\,. \tag{13.22}$$

In order to do this let us rewrite equation (13.21) as

$$\begin{aligned}(F_s\rho)(\lambda,\mu) &= \frac{1}{2}\rho(\lambda,\mu) + \frac{\operatorname{sgn}(s)}{2\pi}\int_{-\infty}^{\infty}\frac{\sin(|s|x)}{x}\rho(x+\lambda,\nu)\,dx \\ &+ \frac{1}{2\pi i}\mathrm{P}\int_{-\infty}^{\infty}\frac{\cos(sx)}{x}\rho(x+\lambda,\nu)\,dx\,.\end{aligned} \tag{13.23}$$

If for each fixed ν the function $\rho(\cdot,\nu)$ is continuous and satisfies the conditions

(i) $\dfrac{\rho}{1+|\lambda|}$ is integrable

(ii) $\lambda \longmapsto \rho(\lambda,\nu)$ is of bounded variation

then by ([Titchmarch] Th. 12, or [Zygmund] Ch. XVI)

$$\lim_{s\to\pm\infty}\frac{\operatorname{sgn}(s)}{2\pi}\int_{-\infty}^{\infty}\frac{\sin(|s|x)}{x}\rho(x+\lambda,\nu)\,dx = \pm\frac{1}{2}\rho(\lambda,\nu)\,.$$

To evaluate the second integral in (13.23) assume that

(iii) the derivative $\frac{\partial}{\partial\lambda}\rho(\lambda,\nu)$ is continuous for each ν

and let $\delta > 0$. Then

$$\begin{aligned} \mathrm{P}\int_{-\infty}^{\infty}\frac{\cos(sx)}{x}\rho(x+\lambda,\nu)\,dx &= \left(\int_{-\infty}^{-\delta}+\int_{\delta}^{\infty}\right)\frac{\cos(sx)}{x}\rho(x+\lambda,\nu)\,dx \\ &+\lim_{\delta\to 0}\int_{-\delta}^{\delta}\frac{\cos(sx)}{x}\left[\rho(x+\lambda,\nu)-\rho(\lambda-x,\nu)\right]\,dx\,. \end{aligned} \tag{13.24}$$

The integrals over the intervals $(-\infty,-\delta)$ and (δ,∞) vanish by the Riemann-Lebesgue Lemma.

Since

$$\left|\frac{\rho(x+\lambda,\nu)-\rho(\lambda-x,\nu)}{x}\right| \le 2\left|\frac{\partial}{\partial\lambda}\rho(\lambda,x)\right| < \infty\,,$$

the third integral in (13.24) is bounded by $4\delta\left|\frac{\partial}{\partial\lambda}\rho(\lambda,x)\right|$, which can be made arbitrarily small. Thus the third integral vanishes. Therefore

$$\lim_{s\to\pm\infty}(F_s\rho)(\lambda,\mu) = \frac{1}{2}\rho(\lambda,x)\pm\frac{1}{2}\rho(\lambda,x) = \begin{cases}\rho(\lambda,x)\,, & s\to\infty\\ 0\,, & s\to-\infty,\end{cases}$$

for each ρ that satisfies conditions (i) – (iii).

The class of functions $\rho(\lambda,\nu)$ that satisfy conditions (i) – (iii) is dense in $L^2(\mathbb{R}\times\mathbb{R},d\lambda d\nu)$ (note that each function from the Schwartz space of rapidly decreasing functions, which is dense in $L^2(\mathbb{R}\times\mathbb{R},d\lambda d\nu)$, satisfies conditions (i) – (iii)).

It follows from the above considerations that (13.22) is satisfied on a dense subset of $L^2(\mathbb{R}\times\mathbb{R},d\lambda d\nu)$. Since the operators $F_{-\infty}$ and F_{∞} are bounded, (13.22) holds on the whole Hilbert space $L^2(\mathbb{R}\times\mathbb{R},d\lambda d\nu)$. This also concludes the proof that $\{F_s\}$ is a spectral family satisfying imprimitivity conditions with respect to e^{-iLt}. Thus

$$T = \int_{-\infty}^{\infty}\lambda\,dE_\lambda$$

is a time operator in any of the equivalent sense (a), (b) or (c) and hence also satisfies the canonical commutation relation with L.

If the Hamiltonian H is not simple (as is usually the case) we can always choose a set $A_1, A_2, \ldots$ of observables which together with H form a mutually independent complete set of commuting observables and consider simultaneous spectral representation of the complete set $\{H, A_1, A_2, \ldots\}$. Any $\rho\in\mathcal{H}$ will again be represented by a kernel $\rho(\mu,\sigma_1,\sigma_2,\ldots,\mu',\sigma_1',\sigma_2',\ldots)$, where σ_k belong to the spectrum of A_k. The previous argument can be carried through without any change except for the appearance of degenerating indices σ_k.

The most important problem from the point of view of application of time operator techniques is to construct explicitly time operator associated with perturbed Hamiltonian $H = H_0 + V$ when the spectral representation is known. One of the simplest physical Hamiltonians provides the Fridrichs model. Since spectral representation in the Fridrichs model (for a given V) can be explicitly constructed it is, in principle, possible to construct the spectral family $\mathcal{F}_s\}$ of T as indicated above. But this attempt meets technical difficulty and has not yet been done.

14

Intertwining dynamical systems with stochastic processes

14.1 MISRA-PRIGOGINE-COURBAGE THEORY OF IRREVERSIBILITY

The problem of irreversibility in statistical physics lies in understanding the relation between reversible dynamical laws and the observed irreversible evolutions that require probabilistic description.

The laws of classical mechanics are deterministic and symmetric with respect to inversion of time. Irreversibility of a physical process is, on the other hand, expressed by the second law of thermodynamics. The quantitative measure of irreversible evolutions in isolated systems is the *entropy* defined as a functional

$$\Omega(\rho) = -\int_{\mathcal{X}} \rho ln \rho \, d\mu \tag{14.1}$$

acting on probability densities. Irreversible time evolution $t \to \rho_t$ of densities in an isolated system is expressed by the increase of $\Omega(\rho_t)$ with time until it reaches its maximum at equilibrium. For classical dynamical systems the postulate of increasing entropy is, however, in contradiction with the laws of classical mechanics because the entropy $\Omega(\rho_t)$ is conserved as the result of the measure preserving character of dynamical evolution. Recall here that a reversible dynamical system is described by a quadruple $(\mathcal{X}, \Sigma, \mu; \{S_t\})$, where $\{S_t\}$ is a group of measure preserving transformations of the phase space $\mathcal{X}$. An equivalent description of time reversible dynamics is through the group of evolution operators $\{U_t\}$ acting on the phase space functions. On the Hilbert space $L^2_{\mathcal{X}}$ the group $\{U_t\}$ is unitary.

The prototypes of irreversible evolutions of phase space functions are Markov semigroups. Particularly these semigroups that “tend to equilibrium" as time tends to ∞. Such Markov semigroups may arise from Markov processes such as kinetic or diffusive process. They may also be dynamical semigroups of irreversible dynamical system with strong ergodic properties like exactness. The puzzling thing is that for some dynamical systems, which are reversible, the observed evolution can be more adequately described by Markov semigroups. This suggest the existence of a strong link between deterministic dynamical evolutions and probabilistic Markov processes.

Until the late 1970s it had been a general belief that stochastic processes can arise from deterministic dynamics only as a result of some form of “coarse-graining" or approximations. A new approach to the problem of irreversibility was proposed by Misra, Prigogine and Courbage [MPC,...]. According to this theory, which will be called *MPC theory of irreversibility*, there is an “equivalence" between highly unstable deterministic and stochastic evolution. The unitary group of evolution of a highly unstable system can be linked with a Markov semigroup through a non-unitary similarity transformation. In contrast with the coarse-graining approach there is no loss of information involved in such transition.

The similarity transformation linking two kind of dynamical semigroups, traditionally denoted by Λ, was establish for K-systems as a function of the time operator. In this section we shall present in detail the MPC-theory of irreversibility, i.e. the construction of Λ-transformation its properties, as well as, the properties of the resulting Markov semigroups.

Before addressing the Misra-Prigogine-Courbage theory of irreversibility let us first recall some basic facts. Consider an abstract dynamical flow given by the quadruple $(\mathcal{X}, \Sigma, \mu, \{S_t\})$, where $\{S_t\}$ is a group of one-to-one μ invariant transformations of $\mathcal{X}$ and either $t \in \mathbb{Z}$ or $t \in \mathbb{R}$. The invariance of the measure μ implies that the transformations U_t

$$U_t\,\rho(x) = \rho(S_{-t}x), \quad \rho \in L^2$$

are unitary operators on L^2. Generally, U_t is an isometry on the space L^p, $1 \leq p \leq \infty$.

The MPC-theory of irreversibility proposes to relate the group $\{U_t\}$, considered on the Hilbert space $L^2_{\mathcal{X}}$, with the irreversible semigroup

$$W_t \stackrel{\mathrm{df}}{=} \Lambda U_t \Lambda^{-1}\,, \ \ t \geq 0\,, \tag{14.2}$$

through a nonunitary operator Λ with the properties:

(a) $\Lambda\rho \geq 0$ if $\rho \geq 0$,

(b) $\int_{\mathcal{X}} \Lambda\rho\, d\mu = \int_{\mathcal{X}} \rho\, d\mu$, for $\rho \geq 0$,

(c) $\Lambda 1 = 1$,

(d) Λ has a densely defined inverse Λ^{-1}.

Moreover, it is also assumed that $\{W_t\}_{t\geq 0}$ satisfies conditions (a)–(c) and

(e) $\lim\limits_{t\to\infty} \|W_t\rho - 1\|_{L^2} = 0$,

for each square integrable density ρ.

Any operator that satisfies properties (a)–(c) is a positivity preserving contraction on L^p, $p \geq 1$ (see ???), thus $\{W_t\}$ is a Markov semigroup. Property (e) is interpreted as irreversible approach to equilibrium. Another interpretation of (14.2) is that group $\{U_t\}$ and semigroup $\{W_t\}$ are connected through a nonunitary, intertwining operator Λ:

$$W_t \Lambda = \Lambda U_t\,, \ \ t \geq 0\,.$$

A dynamical system for which such a construction is possible is called *intrinsically random* and the conversion of the reversible group $\{U_t\}$ into the irreversible semigroup $\{W_t\}$ through a nonunitary transformation Λ is called a *change of representation*.

So far all known constructions of the operator Λ have been done for dynamical systems which are K-flows. If the considered dynamical system is a K-flow then (see Section ???) there exists a generating sub-σ-algebra Σ_0 of Σ such that $\Sigma_t \overset{\mathrm{df}}{=} S_t(\Sigma_0)$, $t \geq 0$, have the properties (i)–(iii)listed in Section ???.

The construction of Λ is the following. With any K-flow we associate the family of conditional expectations $\{E_t\}$ with respect to the σ-algebras $\{\Sigma_t\}$, which projectors on the Hilbert space $L^2_{\mathcal{X}}$. These projectors form a resolution of identity on $L^2_{\mathcal{X}}$, thus determine the time operator T:

$$T = \int_{-\infty}^{+\infty} t dE_t\,, \tag{14.3}$$

which is defined on a dense subspace of $L^2_{\mathcal{X}}$ (see Section 3).

We shall show now that the intertwining operator Λ can be defined, up to constants, as a function of the operator T, i.e.

$$\Lambda = f(T) + E_{-\infty}\,, \tag{14.4}$$

where $E_{-\infty}$ is the expectation (the projection on constants).

We have shown in Section 3 that for any Borel measurable function f the operator function $f(T)$

$$f(T) = \int_{-\infty}^{\infty} f(t)\, dE_t$$

is correctly defined on a dense domain $D(f(T)) \subset L^2_{\mathcal{X}}$. If f is positive and nondecreasing then we can be more specific concerning the domain of $f(T)$. Namely, we have

Proposition 14.1 *If f is a nondecreasing and nonnegative function then the domain $D(f(T))$ contains all simple functions measurable with respect to the σ-algebra $\bigcup_t \Sigma_t$*

Proof. Recall that a necessary and sufficient condition for ρ to be in the domain $D(f(T))$ is (see Section 3):

$$\int_{-\infty}^{\infty} f^2(t)\, d\langle E_t\rho, \rho\rangle < \infty\,.$$

Let ρ be a simple function measurable with respect to $\bigcup_t \Sigma_t$. Since f is nondecreasing, there exists t_0 such that ρ is $\mathcal{A}_{t_0}$-measurable. Let n_0 be such a number that

$$(-\infty, t_0] \subset \{s \in \mathbb{R} \,:\, f(s) \le n_0\}\,.$$

By the orthogonality of increments of $\{E_t\}$ (or using the fact that $E_t\rho$ is a martingale with respect to $\{\Sigma_t\}$) we have

$$(E_t - E_{t_0})\rho = 0, \,, \quad \text{for each } t \ge t_0\,.$$

Hence the function $t \mapsto \langle E_t\rho, \rho\rangle$ is constant on the interval (t_0, ∞), and consequently

$$\begin{aligned}\int_{-\infty}^{\infty} f^2(s)\, d\langle E_t\rho, \rho\rangle \quad &= \int_{-\infty}^{t_0} f^2(s)\, d\langle E_t\rho, \rho\rangle \\ &\le n_0^2 \int_{-\infty}^{\infty} d\langle E_t\rho, \rho\rangle \\ &= n_0^2 \int_{\mathcal{X}} |\rho|^2\, d\mu < \infty\,.\end{aligned}$$

We are now ready to construct the operator Λ in MPC theory of irreversibility.

Theorem 14.1 *Assume, that f is a non increasing function on $\mathbb{R}$ such that $f(t) > 0$ for each t, and $f(-\infty) = 1$, $f(+\infty) = 0$. Then the operator*

$$\Lambda\rho = \int_{-\infty}^{\infty} f(s)\, dE_s\rho + E_{-\infty}\rho \tag{14.5}$$

has the following properties:

(i) Λ *is a bounded linear operator,*

(ii) Λ *is positively defined, i.e.* $\Lambda\rho \ge 0$ *if* $\rho \ge 0$,

(iii) $\Lambda 1 = 1$,

(iv) Λ *is one-to-one with a densely defined inverse* Λ^{-1}.

The function f is called the *scaling function*. The operator $E_{-\infty}$ is the projection on constants, i.e. $E_{-\infty}\rho = \int_{\mathcal{X}} \rho\, d\mu$.

In order to prove Theorem 14.1 it will be convenient to use another interpretation of the integral $\int f(s)\, dE_s$. Suppose that f is a function of bounded variation on $\mathbb{R}$. Then we can consider a mean-square integral $\int_{-\infty}^{\infty} E_s\rho\, df(s)$, i.e., a Stieltjes integral of the vector valued function $s \mapsto E_s\rho$ from $\mathbb{R}$ into $L^2_{\mathcal{X}}$. Since the family $\{E_t\}$ (martingale $\{E_t\}\rho$) is right mean-square continuous, this integral is well defined and we have the following “integration by parts formula”

Lemma 14.1 *If f is a function of bounded variation on $\mathbb{R}$ then*

$$\int_{-\infty}^{\infty} f(s)\, dE_s\rho = -\int_{-\infty}^{\infty} E_s\, df(s) + f(+\infty)\rho - f(-\infty)E_{-\infty}\rho\,,$$

for each $\rho \in L^2_{\mathcal{X}}$

Proof of Theorem 14.1. Property (i) is a consequence of Proposition 2 in Section 3, (ii) follows from Lemma, because we can write

$$\Lambda\rho = -\int_{-\infty}^{\infty} E_s\rho \, df(s) \,, \tag{14.6}$$

and E_s are positivity preserving as conditional expectations. Also (ii) is obvious, because

$$\Lambda 1 = \int_{-\infty}^{\infty} E_s 1 + E_{-\infty} = E_\infty 1 - E_{-\infty} 1 + E_{-\infty} 1 = 1 \,.$$

We shall prove now that operator Λ is injective. Indeed, assume that $\Lambda\rho = 0$. Then $E_t\rho\Lambda = 0$, for each t. Since

$$E_t\Lambda\rho = E_t \int_{-\infty}^{t} f(s)\, dE_s\rho + E_t \int_{t}^{\infty} f(s)\, dE_s\rho + E_{-\infty} = \int_{-\infty}^{t} f(s)\, dE_s\rho + E_{-\infty} \,,$$

then

$$0 = E_{t_2}\Lambda\rho - E_{t_1}\Lambda\rho = \int_{t_1}^{t_2} f(s)\, dE_s\rho \,,$$

for each $t_1 < t_2$. Hence

$$\int_{t_1}^{t_2} f^2(s)\, d\langle E_s\rho, \rho\rangle = \int_{\mathcal{X}} \left(\int_{t_1}^{t_2} f(s)\, dE_s\rho \right)^2 d\mu = 0 \,.$$

Since $f > 0$ the function $s \mapsto \langle E_s\rho, \rho\rangle$ must be constant on the interval $(t_1, t_2]$, as it is nondecreasing. This implies that $E_{t_2}\rho = E_{t_1}\rho$. If $t_1 \to \infty$, then the last equality implies that $E_t\rho = E_{-\infty}\rho$=const, for each $t \in \mathbb{R}$. Consequently, $\int_{-\infty}^{\infty} f(s)\, dE_s\rho = 0$ and (14.5) implies that $E_{-\infty} = 0$. Therefore $E_t\rho = 0$, for each t, so $\rho = E_\infty\rho = 0$.

Finally, put

$$\Lambda^{-1}\rho = \int_{-\infty}^{\infty} \frac{1}{f(s)}\, dE_s\rho + E_{-\infty}\rho \,.$$

From Proposition 1 follows that Λ^{-1} is well defined on the set of all simple functions that are $\sum_t \mathcal{A}_t$ measurable. Using Proposition ??? from Section 3 we obtain that $\Lambda\Lambda^{-1}\rho = \rho$ on a dense subspace of $L^2_{\mathcal{X}}$. We have also $\Lambda^{-1}\Lambda\rho = \rho$, because $\Lambda\rho$ is in the domain of Λ^{-1}, for each $\rho \in L^2_{\mathcal{X}}$.

Theorem 14.2 *Let f be as defined in Theorem 14.1, and assume additionally, that for each $t \geq 0$ $f(s)/f(s-t)$ is a bounded and non increasing function of s. Then the family of operators*

$$W_t = \Lambda U_t \Lambda^{-1} \,, \quad t \geq 0 \,,$$

is a Markov semigroup on $L^2_{\mathcal{X}}$ and $\|W_t\rho - 1\| \to 0$, as $t \to \infty$, for each probability density ρ.

Proof. First observe that the Markov operators W_t are of the form

$$W_t = \left(\int_{-\infty}^{\infty} \frac{f(s)}{f(s-t)} \, dE_s + E_{-\infty} \right) U_t \, .$$

Indeed, applying Proposition 3 and ??? from Section 3 we have

$$\begin{aligned} W_t &= \Lambda U_t \Lambda^{-1} \\ &= \left(\int_{-\infty}^{\infty} f(s) \, dE_s + E_{-\infty} \right) U_t \left(\int_{-\infty}^{\infty} \frac{1}{f(s)} \, dE_s + E_{-\infty} \right) \\ &= \left(\int_{-\infty}^{\infty} f(s) \, dE_s + E_{-\infty} \right) \left(\int_{-\infty}^{\infty} \frac{1}{f(s-t)} \, dE_s + E_{-\infty} \right) U_t \\ &= \left(\int_{-\infty}^{\infty} \frac{f(s)}{f(s-t)} \, dE_s + \int_{-\infty}^{\infty} f(s) \, dE_s E_{-\infty} \right. \\ &\qquad \left. + \int_{-\infty}^{\infty} \frac{1}{f(s)} \, dE_{-\infty} E_s + E_{-\infty} \right) U_t \, . \end{aligned}$$

But the second and third term are equal to 0 because the integrators $E_s E_{-\infty}$ and $E_{-\infty} E_s$ are constant.

It is now obvious that for each t W_t is positivity preserving. Moreover, applying Corollary 3.1 to the integral $\int f(s)/f(s-t) \, dE_s$ we obtain that W_t is bounded on $L^2_{\mathcal{X}}$.

No we show that W_t, $t > 0$, are contractions. If g is a bounded measurable function, say $|g| \le 1$, and $\rho \in L^2_{\mathcal{X}}$, then

$$\left\| \int_{-\infty}^{\infty} g(s) \, dE_s \rho \right\|^2_{L^2} = \int_{-\infty}^{\infty} g^2(s) \, d\langle E_s \rho, \rho \rangle - \left(\int_{-\infty}^{\infty} g(s) \, d\langle E_s \rho, \rho \rangle \right)^2 .$$

Hence

$$\begin{aligned} \|W_t \rho\|^2_{L^2} &= \left\| \int_{-\infty}^{\infty} \frac{f(s)}{f(s-t)} \, dE_s (U_t \rho) + E_{-\infty} (U_t \rho) \right\|^2_{L^2} \\ &= \left\| \int_{-\infty}^{\infty} \frac{f(s)}{f(s-t)} \, dE_s (U_t \rho) \right\|^2_{L^2} + \|E_{-\infty} U_t \rho\|^2_{L^2} \\ &\le \int_{\mathcal{X}} (U_t \rho)^2 \, d\mu - \left(\int_{\mathcal{X}} U_t \rho \, d\mu \right)^2 + \left(\int_{\mathcal{X}} U_t \rho \, d\mu \right)^2 \\ &= \|\rho\|^2_{\mathcal{X}} \, , \end{aligned}$$

for $t > 0$. In consequence $\{W_t\}_{t>0}$ is a contraction semigroup on $L^2_{\mathcal{X}}$.

Finally, because

$$\begin{aligned}\|W_t\rho - 1\|^2_{L^2} &= \left\| \int_{-\infty}^{\infty} \frac{f(s)}{f(s-t)}\, dE_s(U_t\rho) \right\|^2_{L^2} \\ &= \left\| U_t \int_{-\infty}^{\infty} \frac{f(s+t)}{f(s)}\, dE_s(U_t\rho) \right\|^2_{L^2} \\ &\le \|U_t\| \cdot \left\| \int_{-\infty}^{\infty} \frac{f(s+t)}{f(s)}\, dE_s(U_t\rho) \right\|^2_{L^2}\end{aligned}$$

and $f(s+t)$ tends monotonically to 0, as $t \to +\infty$, we get that $\|W_t\rho - 1\| \to 0$, as $t \to \infty$.

Remark. The assumptions of the above theorem are satisfied if the function f, which is called the *scaling function*, is assumed to be positive, non increasing, $f(-\infty) = 1$, $f(+\infty) = 0$ and such that $ln f$ is concave on $\mathbb{R}$. This means that f has the form

$$f(t) = e^{-\phi(t)}\,, \tag{14.7}$$

where $\phi(t)$ is a positive convex function which increases to $+\infty$ as $t \to +\infty$. Indeed, we have the following

Lemma 14.2 *If the function f is logarithmically concave then $\dfrac{f(s)}{f(s-t)}$ is a bounded decreasing function of s for every $t > 0$.*

Proof. We can assume that f is of the form (14.7. Therefore it is enough to show that the function $s \mapsto \phi(s-t) - \phi(s)$ is decreasing. Let $s_1 < s_2$, then $0 < t/(s_2+t-s_1) < 1$. Since the function ϕ is convex, we have

$$\begin{aligned}\phi(s_1-t)+\phi(s_2) &= \frac{t}{s_2+t-s_1}\phi(s_1-t) + \frac{s_2-s_1}{s_2+t-s_1}\phi(s_2) \\ &\quad + \frac{s_2-s_1}{s_2+t-s_1}\phi(s_1-t) + \frac{t}{s_2+t-s_1}\phi(s_2) \\ &\ge \phi\Big(\frac{t(s_1-t)+(s_2-s_1)s_2}{s_2+t-s_1}\Big) + \phi\Big(\frac{(s_2-s_1)(s_1-t)+ts_2}{s_2+t-s_1}\Big) \\ &= \phi(s_2-t)+\phi(s_1)\,,\end{aligned}$$

which implies the desired result

$$\phi(s_1-t) - \phi(s_1) \ge \phi(s_2-t) - \phi(s_2)\,.$$

The L^2-construction of the operator Λ and the Markov semigroup can be generalized in several directions. One possibility is that the family of projectors $\{E_t\}$ can be replaced by an operator valued martingale $\{M_t\}$. Namely consider the family $\{M_t\}$ of bounded linear operators on L^2 with properties:

(a) $\{M_t\}$ has orthogonal increments:

$$(M_{s_2} - M_{s_1})(M_{t_2} - M_{t_1}) = 0, \text{ for } s_1 < s_2 \leq t_1 < t_2$$

(b) $\{M_t\rho\}$ is a right continuous martingale for each $\rho \in L^2$. For such martingales it is possible to obtain a straightforward generalization of the MPC construction (see [SWRS,SW]).

Another direction in which the MPC-construction can be generalized is the enlargement of the space of admissible densities. The choice of L^2 as the space on which the unitary group of evolution, and consequently the MPC theory, is considered was purely technical. More natural, however, is to consider the evolution of any probability density ρ under the action of Frobenius-Perron operators defined on L^1. This is equivalent to consider the evolution of all probability measures on $(\mathcal{X}, \Sigma)$ that are absolutely continuous with respect to μ. We can go even further, considering the evolution of *any* probability measure under the action of the Frobenius-Perron operator. Indeed, if ν is an arbitrary measure then the time evolution $t \mapsto U_t\nu$ can be defined provided then exist a dense subspace $\mathcal{G}$ of $L^1_{\mathcal{X}}$, which is invariant with respect to the Koopman operator, i.e. $V_t\mathcal{G} \subset \mathcal{G}$, for each t and such that each $\rho \in \mathcal{G}$ is ν-integrable. Then we put

$$\langle U_t\nu, \rho\rangle = \int_{\mathcal{X}} V_t\rho \, d\nu \ , \text{ for each } \rho \in \mathcal{G} \, .$$

The question now is: Does the MPC theory of irreversibility remain valid on “larger" spaces? The answer is: Yes but not on the whole L^1-space. There is a class $\mathcal{D}$ of admissible densities (or, more generally, admissible measures), on which the Markov semigroup $\{W_t\}$, associated with $\{U_t\}$, converges to the equilibrium state. Class $\mathcal{D}$ contains all square integrable densities, but not all L^1-densities are in $\mathcal{D}$.

We shall focus our attention on L^1 extensions of MPC theory irreversibility. Such extension does not require an additional effort of defining the evolution operators but, on the other hand, reveals some interesting features concerning the choice of admissible densities.

In section 3 we have already developed the tools, which are necessary for L^1-extension. One of them is the stochastic integral with respect to L^1-martingales. This allows to define the integral of the form $\int_{-\infty}^{\infty} f(s)\, dE_s\rho$ where f is a Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ and $\rho \in L^1_{\mathcal{X}}$.

Denote by $\mathcal{D}_f$ the set of all $\rho \in L^1_{\mathcal{X}}$ for which $\int_{-\infty}^{\infty} f(t)\, dE_t\rho$ exists. In other words, $\rho \in \mathcal{D}_f$ if and only if there exists a sequence $\{f_n\}$ of simple Borel measurable functions on $\mathbb{R}$ such that

1° $f_n \to f$ a.s

2° $\int\limits_{-\infty}^{\infty} f\,(s)\ dE_s\rho$ converges in $L^1_{\mathcal{X}}$.

It follows from Proposition ???, in Section 3, that a sufficient condition for ρ being in $\mathcal{D}_f$, where f is bounded, is that

$$\|\sup_s |E_s\rho|\|_{L^1} < \infty\,. \tag{14.8}$$

What matters in the above condition is the behavior of the martingale in $+\infty$ (its uniform integrability). We know, by the backward martingale convergence theorem that the martingale $\{E_s\rho\}$ is convergent and uniformly integrable, as $s \to -\infty$. The space of all martingales $\{m_t\}$ with respect to the filtration $\{\Sigma_t\}$ and satisfying (14.8) is denoted by H^1 (sometimes called the *Hardy space of martingales* [DeMe]). Identifying each uniformly integrable martingale with its limit of infinity we have $H^1 \subset L^1$. The condition that ρ is an element of H^1 guarantees that the stochastic integral $\int f(s)\,dE_s\rho$ with respect to the martingale $\{E_s\rho\}$ can be defined for any bounded function f. As we shall see below this condition is in some circumstances also necessary. For this reason we may confine our further consideration concerning L^1-dynamics to H^1 spaces.

Now, we can reformulate Theorem 14.1 and 14.2 as follows.

Theorem 14.3 *For any decreasing and positive function f on $\mathbb{R}$ with the properties: $f(-\infty) = 1$, $f(+\infty) = 0$ the linear operator Λ on $L^1_{\mathcal{X}}$ defined by the formula*

$$\Lambda\rho = \int_{-\infty}^{\infty} f(s)\,dE_s\rho + E_{-\infty}\rho$$

satisfies:

(i) Λ is positively preserving, i.e. $\Lambda\rho \geq 0$ if $\rho \geq 0$
(ii) $\Lambda 1 = 1$
(iii) Λ restricted to H^1 is bounded with densely defined inverse Λ^{-1}.

Proof. Since f is bounded the integral $\int f(s)\,dE_s\rho$ exists for each $\rho \in H^1$. Moreover approximating f by step functions it is easy to check that Lemma 1 remains valid, i.e.

$$\Lambda\rho = -\int_{-\infty}^{\infty} E_s\rho\, df\,(s)\,.$$

This shows that Λ is positivity preserving. The rest of the proof also follows the same lines as the proof of Theorem 14.1. Only the boundeness of Λ is a result of stochastic integral inequalities that can be found in [Bi] (Th.7.2).

Now we can use the above defined operator Λ as a similarity transformation converting $\{U_t\}$ into a Markov semigroup. Putting

$$W_t = \Lambda U_t \Lambda^{-1}$$

we obtain a family of linear operators densely defined on $L^1_{\mathcal{X}}$.

Theorem 14.4 *Let f and W_t be as in Theorem 14.2. Then the semigroup $\{W_t\}_{t\geq 0}$ restricted to H^1 form a Markov semigroup such that*

(i) $W_t 1 = 1$, *for* $t \geq 0$
(ii) $\|W_t\rho - 1\|_{L^1} \to 0$, *as* $t \to \infty$, *for each probability density* $\rho \in H^1$.

Proof. The proof is the same as Theorem 14.2, because the commutation relation

$$U_t \left(\int_{-\infty}^{\infty} f(s)\, dE_s\rho \right) = \left(\int_{-\infty}^{\infty} f(s-t)\, dE_s \right) U_t$$

remains true for uniformly integrable martingales $\{E_t\rho\}$. The convergence of the semigroup $\{W_t\}$ is a direct consequence of the integrability of the bounded functions $s \mapsto \dfrac{f(s)}{f(s-t)}$ and their monotonical convergence as $t \to \infty$.

The next theorem shows that the evolution of densities under the action of the Markov semigroup $\{W_t\}$ has strictly increasing entropy.

Theorem 14.5 *If the measure* μ *is non-degenerated then for every probability density* ρ *with* $\Omega(\rho) < \infty$, *the Markov evolution* $\tilde{\rho}_t = W_t\rho$, $t \geq 0$, *has strictly increasing entropy.*

Proof. Since the function $\rho \ln \rho$ is integrable, the martingale $\{E_t\rho\}$ belongs to H^1 (cf. [DeMe], p. 261). Hence $\tilde{\rho}_t = W_t\rho \in L^1_{\mathcal{X}}$ by Theorem 14.2. Denote

$$\varphi(x) = x \ln x\,,\ x > 0\,.$$

First we show the inequality

$$\int_{\mathcal{X}} \varphi(W_t\rho)\, d\mu \leq \int_{\mathcal{X}} \varphi(\rho)\, d\mu\,. \tag{14.9}$$

Indeed, let ρ be a simple function, $\rho = \sum_1^n a_i \mathbb{1}_{A_i}$ such that $\bigcup_1^n A_i = \mathcal{X}$ and $A_i \in \bigcup_t \Sigma_t$. Then $W_t\rho \in L^1_{\mathcal{X}}$ and

$$\varphi(W_t\rho) = \varphi\left(\sum_1^n a_i W_t \mathbb{1}_{A_i}\right) \leq \sum_1^n \varphi(a_i)\, W_t \mathbb{1}_{A_i} = W_t\varphi(\rho)$$

because φ is convex. Hence

$$\int_{\mathcal{X}} \varphi(W_t\rho)\ d\mu \leq \int_{\mathcal{X}} W_t\varphi(\rho)\, d\mu = \int_{\mathcal{X}} \varphi(\rho)\, d\mu\,.$$

For a non-simple ρ it is enough to choose a sequence $\rho_n \to \rho$ in $L^1_{\mathcal{X}}$ and such that $\varphi(\rho_n) \to \varphi(\rho)$ in $L^1_{\mathcal{X}}$ and apply Theorem 14.2. Now putting $W_{t_2 - t_1}$ instead of W_t and $W_{t_1}\rho$ instead of ρ in (14.9) we have for $t_1 < t_2$

$$\Omega(\tilde{\rho}_{t_1}) \leq \Omega(\tilde{\rho}_{t_2})\ .$$

Since μ is non-degenerate and φ is strictly convex the above inequality is strict.

It is possible to obtain another generalization of the intertwining relation between the unitary dynamics $\{U_t\}$ and the Markov semigroup if we drop the assumption of reversibility of the transformation Λ. Namely, let $\{U_t\}$ be the group of operators on L^1 associated with a K-flow $(\mathcal{X}, \Sigma, \mu, \{S_t\})$ and let f be a non increasing function on the interval $[a, b]$ (can be $a = -\infty$ or $b = +\infty$) such that $f(a) = 1$ and $f(b) = 0$. Then the linear operator Λ on L^1 defined by the formula

$$\Lambda = \int_a^b f\,(s)\,dE_s + E_a \tag{14.10}$$

satisfies the properties (i)–(iii) of Theorem 14.1 (or 14.2).

Moreover, assume additionally that $0 < f(t) < 1$, for $a < t < b$, and that $\frac{f(s)}{f(s-t)}$) is a non increasing function of s in (a, b). Define:

$$W_t = \left(\int_a^b \frac{f\,(s)}{f\,(s-t)}\,dE_s + E_a \right) U_t\,. \tag{14.11}$$

Then the semigroup $\{W_t\}$, $t \geq 0$, is contractive on $L^1_{\mathcal{X}}$ and tends monotonically to the equilibrium state on the set $\mathcal{D}_0$ of all densities ρ which satisfy the property

$$\int_{\mathcal{X}} \sup_{a \leq t \leq b} |E_t \rho| d\mu < \infty\,. \tag{14.12}$$

Of course, assuming additionally that $a = -\infty$ and $b = +\infty$ we obtain the original intertwining relation.

The above extended meaning of the transformation Λ contains, in particular, the "coarse graining" procedure. Indeed, taking f equal to 1 on the interval $(-\infty, a]$ and 0 on (a, ∞) the condition (14.12) will be satisfied on the space $\mathcal{D}_0 = L^1_{\mathcal{X}}$, and we have $\Lambda = E_a$ and $W_t = U_t E_{a-t}$.

14.2 NONLOCALITY OF THE MISRA-PRIGOGINE-COURBAGE SEMIGROUP

In the Misra-Prigogine-Courbage theory of irreversibility the evolution operators that arise from point transformations of a phase space have been modified and leads to new evolution semigroups that need not to be related with the underlying point dynamics. The natural question is: Are these operators associated with other point transformations? In other words, we ask if, for example, modifications made on the level of evolution operators correspond to some modifications on the level of trajectories in the phase space. In a more general setting, we ask which linear operators on L^p spaces are implementable by point transformations? A related question is: Which time evolutions of states of physical systems that are described in terms of a semigroup $\{W_t\}$ of maps on an L^p-space can be induced by Hamiltonian flows? There are some partial answers to the above questions that will be presented below.

Consider the case of discrete time $t = 1, 2, \dots$ when the evolution semigroup $\{V_t\}$ is determined by a single transformation S

$$V_n = V^n \quad \text{and} \quad Vf(x) = f(Sx)\,.$$

The relation of the point dynamics with the Koopman operators is clarified by asking the question: What types of isometries on L^p spaces are implementable by point transformations? For L^p spaces with $p \neq 2$, all isometries induce underlying point transformations. Such theorems on the implementability of isometries on L^p spaces, $p \neq 2$, are known as Banach-Lamperti theorems [Ban,Lam]. The converse to Koopman's lemma in the case $p = 2$, which holds under the additional assumption that the isometry on L^2 is positivity preserving, can be found in [GGM]. The result is that an isometry V is implementable by a necessarily measure preserving transformation S

$$Vf(x) = f(Sx)\ , \quad x \in \mathcal{X}\,.$$

Consider now the group $\{U_t\}$ determined by a K-flow and the MPC semigroup $\{W_t\}$. Each operator U_t is the Frobenius-Perron operator associated with S_t and thus it is the adjoint $U_t = V_t^*$ of the Koopman operator V_t acting on $L^2_{\mathcal{X}}$. The operators W_t preserve the property of double stochasticity characteristic to Frobenius-Perron operators. Therefore the question is: are W_t Frobenius-Perron operators associated with some measure preserving transformations $\tilde{S}_t$ or, equivalently, is the adjoint W_t^* the Koopman operator

$$W_t^* f(x) = f(\tilde{S}_t x)\,.$$

We shall show below that the answer to this question is in general negative. Only the choice of Λ as a coarse graining projection gives implementability [AntGu].

Recall that the Λ-transformation is defined on the space $L^2_{\mathcal{X}}$ as

$$\Lambda\rho = \int_{-\infty}^{+\infty} f(s) dE_s \rho + E_{-\infty}\rho \quad \text{for K} - \text{flows}\,, \tag{14.13}$$

$$\Lambda\rho = \sum_{-\infty}^{+\infty} f(s)\big(E_s - E_{s-1}\big)\rho + E_{-\infty}\rho \quad \text{for K} - \text{systems}\ , \tag{14.14}$$

$f(s)$ is any positive function on the reals or integers with the following properties:

(i) Λ is decreasing with $f(\infty) = 0$ and $f(s) < f(-\infty) \leq 1$, for all $t \in \mathbf{R}$. Here we relax the original assumption of Misra, Prigogine and Courbage that Λ is strictly decreasing

(ii) Λ is a logarithmically concave function.

The transformation Λ leads to the irreversible Markov semigroup on the space $L^2_{\mathcal{X}}$

$$W_t = \Lambda U_t \Lambda^{-1} = \int_{-\infty}^{+\infty} \frac{f(s)}{f(s-t)} dE_s U_t + E_{-\infty} U_t \,, \tag{14.15}$$

for K-flows, $t \geq 0$,

$$W_t = \Lambda U_t \Lambda^{-1} = \sum_{-\infty}^{+\infty} \frac{f(s)}{f(s-t)} \big(E_s - E_{s-1}\big) U_t + E_{-\infty} U_t \,, \tag{14.16}$$

for K-systems, t positive integer.

In the sequel we shall need the following property of the semigroup $\{W_t\}$:

Lemma 14.3 *The operator*

$$W_t = \int_{-\infty}^{+\infty} E_s U_t dg(s) \,, \quad t \geq 0 \,, \tag{14.17}$$

with

$$g(s) \equiv -\frac{f(s)}{f(s-t)} \,, \tag{14.18}$$

vanishes or is positivity preserving on $L^2_{\mathcal{X}}$.

Proof. Using integration by parts, W_t can be written as the following Lebesgue-Stieltjes integral (see Lemma on p 75 [9]):

$$\begin{aligned} W_t \quad &= \int_{-\infty}^{\infty} E_s U_t d\Big(-\frac{f(s)}{f(s-t)}\Big) + \lim_{s\to+\infty} \frac{f(s)}{f(s-t)} U_t + \\ &- \lim_{s\to-\infty} \frac{f(s)}{f(s-t)} E_{-\infty} U_t + E_{-\infty} U_t \\ &= \int_{-\infty}^{\infty} E_s U_t d\Big(-\frac{f(s)}{f(s-t)}\Big) + \lim_{s\to+\infty} \frac{f(s)}{f(s-t)} U_t \,, \end{aligned}$$

where we have used

$$\lim_{s\to-\infty} \frac{f(s)}{f(s-t)} = 1 \,.$$

Since the function $\dfrac{f(s)}{f(s-t)}$ is positive and decreasing the limit

$$\lim_{s\to\infty} \frac{f(s)}{f(s-t)}$$

also exists and it is non-negative and (14.17) follows immediately.

Theorem 14.6 *The semigroup $W_t = \Lambda U_t \Lambda^{-1}$, $t \geq 0$ is not implementable, i.e., there does not exist a measurable point transformation $\tilde{S}_t$ of $\mathcal{X}$ such that $\tilde{S}_t$ preserves μ and for which*

$$W_t^* \rho(x) = \rho(\tilde{S}_t x) \quad \text{for all } t \geq 0 \,. \tag{14.19}$$

Since we deal with the adjoint operator W_t^*, it is convenient to use the following lemma.

Lemma 14.4 *Let W be a linear operator on $\mathcal{L}^2$ which is implementable (14.19) by a measure preserving point transformation $\tilde{S}$. Then for each measurable set Δ such that its image $\tilde{S}(\Delta)$ under $\tilde{S}$ is also measurable, the following holds:*

$$\int_{\mathcal{X}-\tilde{S}(\Delta)} W \mathbf{1}_\Delta \, d\mu = 0 \,. \tag{14.20}$$

Proof of Lemma 14.4. For any measurable set Δ for which $\tilde{S}\Delta$ is also measurable, we have

$$\begin{aligned} \int_{\mathcal{X}-\tilde{S}(\Delta)} W \mathbf{1}_\Delta \, d\mu &= \int_{\mathcal{X}} (W \mathbf{1}_\Delta) \cdot \mathbf{1}_{\mathcal{X}-\tilde{S}(\Delta)} \, d\mu = \Big(W \mathbf{1}_\Delta \Big| \mathbf{1}_{\mathcal{X}-\tilde{S}(\Delta)} \Big)_{L^2} \\ &= \Big(\mathbf{1}_\Delta \Big| W^* \mathbf{1}_{\mathcal{X}-\tilde{S}(\Delta)} \Big)_{L^2} = \int_\Delta \mathbf{1}_{\mathcal{X}-\tilde{S}(\Delta)} \circ \tilde{S} \, d\mu \\ &= \int_{\tilde{S}(\Delta)} \mathbf{1}_{\mathcal{X}-\tilde{S}(\Delta)} \, d\mu = 0 \,. \end{aligned}$$

Proof of the Theorem. Now, suppose W_t is implementable by the measure preserving transformation $\tilde{S}_t$ on $\mathcal{X}$ (14.19). Consider a measurable set Δ such that $\tilde{S}_t(\Delta)$ is also measurable and $0 < \mu(\Delta) < 1$. Thus by Lemma 14.4, we have

$$\int_{\mathcal{X}-\tilde{S}_t(\Delta)} W_t \mathbf{1}_\Delta \, d\mu = 0 \,. \tag{14.21}$$

However we shall show that

$$\int_B W_t \mathbf{1}_\Delta \, d\mu > 0 \quad \text{for each measurable } \; B \text{ with } \; \mu(B) > 0 \,, \tag{14.22}$$

which contradicts Lemma 14.4.

Indeed, using Lemma 14.3 and applying Fubini's theorem we have

$$\int_B W_t \mathbf{1}_\Delta \, d\mu \geq \int_B \left(\int_{-\infty}^{\infty} E_s U_t \mathbf{1}_\Delta dg(s) \right) \, d\mu = \int_{-\infty}^{\infty} \Big(\int_B E_s U_t \mathbf{1}_\Delta d\mu \Big) dg(s) \,, \tag{14.23}$$

but $U_t\mathbf{1}_\Delta = \mathbf{1}_{S_t(\Delta)}$ and $E_s\mathbf{1}_{S_t(\Delta)} \to E_{-\infty}\mathbf{1}_{S_t(\Delta)} = \mu(S_t(\Delta)) = \mu(\Delta)$ when $s \to -\infty$. Therefore

$$\int_B E_s U_t \mathbf{1}_\Delta \, d\mu \to \mu(\Delta)\mu(B) \quad \text{when} \quad s \to -\infty \, . \tag{14.24}$$

Consequently there is a number s_0 such that for $s < s_0$

$$\int_B E_s U_t \mathbf{1}_\Delta \, d\mu > \frac{1}{2}\mu(\Delta)\mu(B) \, .$$

Then using (14.23) and (14.24) we obtain

$$\begin{aligned} \int_B W_t \mathbf{1}_\Delta \, d\mu \quad &\geq \int_{-\infty}^{s_0} \Big(\int_B E_s U_t \mathbf{1}_\Delta \, d\mu \Big) dg(s) \\ &> \frac{1}{2}\mu(\Delta)\mu(B) \, \big(g(s_0) - g(-\infty) \big) > 0 \, , \end{aligned}$$

where the inequality $g(s_0) - g(-\infty) > 0$ follows from the fact that $f(s) < f(-\infty)$.

The proof for K-systems follows if we extend the function $f(s)$, $s = 0, \pm 1, \pm 2, \cdots$, to the whole real line in such a way that, for $s < x < s+1$, the value $f(x)$ is equal to the value on the segment joining the points $(s, f(s))$ and $(s+1, f(s+1))$.

15

Spectral and shift representations

15.1 GENERALIZED SPECTRAL DECOMPOSITIONS OF EVOLUTION OPERATORS

The idea behind the spectral analysis of the evolution semigroup $\{V_t\}$ on a Hilbert space $\mathcal{H}$ through the time operator T is to decompose T in terms of its eigenvectors $\varphi_{n,\alpha}$, $T\varphi_{n,\alpha} = n\varphi_{n,\alpha}$

$$T = \sum_n n \sum_\alpha |\varphi_{n,\alpha}\rangle\langle\varphi_{n,\alpha}|$$

in such a way that the system $\{\varphi_{n,\alpha}\}$ is complete in $\mathcal{H}$, i.e. $\sum_{n,\alpha} |\varphi_{n,\alpha}\rangle\langle\varphi_{n,\alpha}| = I$, and the Koopman operator V_t shifts the eigenvectors $\varphi_{n,\alpha}$

$$V_t\varphi_{n,\alpha} = \varphi_{n+t,\alpha} \, .$$

The index n labels the age and α the multiplicity of the spectrum of the time operator. As a result the eigenvectors $\varphi_{n,\alpha}$ of the time operator provide a shift representation of the evolution

$$f = \sum_{n,\alpha} a_{n,\alpha}\varphi_{n,\alpha} \implies V_t f = \sum_{n,\alpha} a_{n,\alpha}\varphi_{n+t,\alpha} = \sum_{n,\alpha} a_{n-t,\alpha}\varphi_{n,\alpha} \, .$$

The knowledge of the eigenvectors of T amounts therefore to a probabilistic solution of the prediction problem for the dynamical system described by the semigroup $\{V_t\}$. The spaces $\mathcal{H}_n$ spanned by the eigenvectors $\varphi_{n,\alpha}$ are called age eigenspaces or spaces of innovations at time n, as they correspond to the new information or detail brought at time n.

Another method of spectral analysis is based on generalized decompositions of evolution operators. Roughly speaking, the idea of this method is to find a generalized spectral decomposition of the evolution operator V in the form

$$V = \sum_k \lambda_k |\varphi_k\rangle\langle F_k| \, ,$$

where λ_k are eigenvalues of V (or U), called the *resonances*, $|\varphi_k\rangle$ are the eigenvectors of U, and $\langle F_k|$ the generalized eigenvectors of V.

The advantage of the general spectral decomposition is that the knowledge of the consecutive powers λ_k^n of resonances is sufficient to describe the time evolution $n \to V^n$ (or the evolution $n \to U^n$).

The notion of a generalized spectral decomposition of self-adjoint operators on a Hilbert space goes back to Dirac [1], who assumed that a given self-adjoint operator A must be of the form

$$A = \int_{\sigma(A)} d\lambda \; \lambda \; |\lambda\rangle\langle\lambda| \quad , \tag{15.1}$$

where $\sigma(A)$ is the spectrum of the operator A. This formula is a straightforward generalization of the familiar decomposition of a self-adjoint operator on a finite-dimensional Hilbert space

$$A = \sum_i \lambda_i \; |e_i\rangle\langle e_i| \, , \tag{15.2}$$

where λ_i and e_i are the eigenvalues and eigenvectors of A, respectively. In infinite dimensional Hilbert spaces, however, the situation is not so simple. The notion of an eigenvalue is replaced by the spectrum, but eigenvectors can be associated only with the discrete part of the spectrum. Nevertheless, a precise meaning can be given to the decomposition (15.1), if we replace eigenvectors by "generalized eigenvectors", which will in general lie outside the given Hilbert space. This is achieved by replacing the initial Hilbert space $\mathcal{H}$ by a dual pair $(\Phi, \Phi^\times)$, where Φ is a locally convex space, which is a dense subspace of $\mathcal{H}$ endowed with a topology, stronger than the Hilbert space topology. This procedure is referred to as rigging and the triple

$$\Phi \subset \mathcal{H} \subset \Phi^\times \tag{15.3}$$

is called a *rigged Hilbert space*.

Gelfand [3,4] was the first to give a precise meaning to the generalized eigenvectors, which was later elaborated by Maurin [5]. Although generalized eigenvectors have a very natural physical interpretation, generalized spectral decompositions have not been used in physics for a long time. Only a few papers had appeared by the end of the 60's (see, for example [6–8]), followed by a series of papers by Bohm and Gadella (see [2] and references therein). The latter publications are particularly significant, because they provide the basis for a rigorous and systematic approach to the problems of irreversibility and resonances in unstable quantum systems like the Friedrichs model [9]. The same ideas can be extended to chaotic dynamical systems, like Kolmogorov systems or exact systems. The observable phase functions of dynamical systems evolve according to the Koopman operator $Vf(x) = f(Sx)$, where S is an endomorphism or an automorphism of a measure space, and f is a square-integrable phase function.

The spectrum of the Koopman operator determines the time scales of the approach to equilibrium very much in analogy with quantum unstable systems, where the spectra of Hamiltonians determine the decay rates. More precisely, the eigenvalues of the Koopman operator or that of its adjoint – the Frobenius-Perron operator, are, according to the terminology introduced by Ruelle and others [12-15] the resonances of the power spectrum. Eigenvalues and eigenvectors of simple chaotic systems have been constructed just recently [16–24].

The question of the existence of a generalized spectral decomposition of extensions of the Koopman operator differs significantly from the original Gelfand-Maurin theory, which was constructed for operators which admit a spectral theorem, like normal operators, giving a generalized spectrum identical with the Hilbert space spectrum. The Koopman operator of unstable systems, however, either does not admit a spectral theorem, as in the case of exact systems, or the generalized spectrum is very different from the Hilbert space spectrum, as in the case of Kolmogorov systems. This requires an extension of the Gelfand-Maurin theory.

Summarizing for the reader's convenience, a dual pair $(\Phi, \Phi^{\times})$ of linear topological spaces constitutes a rigged Hilbert space for the linear endomorphism V of the Hilbert space $\mathcal{H}$ if the following conditions are satisfied:

1) Φ is a dense subspace of $\mathcal{H}$

2) Φ is complete and its topology is stronger than the one induced by $\mathcal{H}$

3) Φ is stable with respect to the adjoint $V^{\dagger}$ of V, i.e. $V^{\dagger}\Phi \subset \Phi$.

4) The adjoint $V^{\dagger}$ is continuous on Φ

The extension $V_{\rm ext}$ of V to the dual $\Phi^{\times}$ of Φ is then defined in the standard way as follows:

$$(\phi|V_{\rm ext}f) = (V^{\dagger}\phi|f)\,,$$

for every $\phi \in \Phi$.

In the sequel we shall not distinguish between V and $V_{\rm ext}$ if confusion is unlikely to arise.

The choice of the test function space Φ depends on the specific operator V and on the physically relevant questions to be asked about the system. For self-adjoint operators V, for example, the generalized spectral theorem can be justified for nuclear test function spaces.

Here, we shall discuss the problem of rigging for the generalized spectral decompositions of the Koopman operators for a few specific but typical models of chaotic systems beginning with the Renyi map.

In the case of the Renyi map various riggings exist so our task will be also to choose a tight rigging. We call a rigging 'tight' if the test function space is the (set-theoretically) largest possible within a chosen family of test function spaces, such that the physically relevant spectral decomposition is meaningful.

The β-adic Renyi map S on the interval [0,1) is the multiplication, modulo 1, by the integer $\beta \geq 2$

$$S : [0,1) \to [0,1) \ : \qquad x \mapsto Sx = \beta x \pmod 1 \,.$$

The probability densities $\rho(x)$ evolve according to the Frobenius–Perron operator U

$$U\rho(x) \equiv \sum_{y, S(y)=x} \frac{1}{|S'(y)|}\rho(y) = \frac{1}{\beta}\sum_{r=0}^{\beta-1}\rho\big(\frac{x+r}{\beta}\big) \,.$$

The Frobenius-Perron operator is a partial isometry on the Hilbert space L^2 of all square integrable functions over the unit interval; it is, moreover, the dual of the isometric Koopman operator V

$$V\rho(x) = U^{\dagger}\rho(x) = \rho(Sx)\,.$$

Using a general algorithmic approach developed in [AT] the following spectral decomposition of the Koopman operator of the Renyi map can be constructed:

$$V = \sum_{n=0}^{\infty} \frac{1}{\beta^n} \, |\tilde{B}_n\rangle\langle B_n| \ \ , \tag{15.4}$$

where $B_n(x)$ is the n-degree Bernoulli polynomial defined by the generating function [29, §9]

$$\frac{ze^{zx}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n$$

and

$$|\widetilde{B}_n\rangle \ = \begin{cases} |1\rangle \,, & n = 0 \\ |\dfrac{(-1)^{(n-1)}}{n!}\{\delta^{(n-1)}(x-1) - \delta^{(n-1)}(x)\}\rangle & n = 1, 2, \ldots \end{cases}$$

The bras $\langle\cdot|$ and kets $|\cdot\rangle$ denote linear and antilinear functionals, respectively. Formula (15.4) defines a spectral decomposition for the Koopman and Frobenius-Perron operators in the following sense

$$(\rho|Vf) = (U\rho|f) = \sum_{n=0}^{\infty} \frac{1}{\beta^n}(\rho|\tilde{B}_n)\,(B_n|f),$$

for any density function ρ and observable f in the appropriate pair $(\Phi, \Phi^{\times})$. Consequently, the Frobenius-Perron operator acts on density functions as

$$U\rho(x) = \int_0^1 dx'\rho(x') + \sum_{n=1}^{\infty} \frac{\rho^{(n-1)}(1) - \rho^{(n-1)}(0)}{n!\beta^n} B_n(x) \,.$$

The orthonormality of the system $|\tilde{B}_n\rangle$ and $\langle B_n|$ follows immediately, while the completeness relation is just the Euler–MacLaurin summation formula for the Bernoulli polynomials [29, §9]

$$\rho(x) = \int_0^1 dx' \rho(x') + \sum_{n=1}^{\infty} \frac{\rho^{(n-1)}(1) - \rho^{(n-1)}(0)}{n!} B_n(x) \; . \tag{15.5}$$

The Bernoulli polynomials are the only polynomial eigenfunctions as any polynomial can be uniquely expressed as a linear combination of the Bernoulli polynomials.

The spectral decomposition (15.4) has no meaning in the Hilbert space L^2, as the derivatives $\delta^{(n)}(x)$ of Dirac's delta function appear as right eigenvectors of V. A natural way to give meaning to formal eigenvectors of operators which do not admit eigenvectors in Hilbert space is to extend the operator to a suitable rigged Hilbert space. A suitable test function space is the space $\mathcal{P}$ of polynomials. The space $\mathcal{P}$ fulfills the following conditions:

i) $\mathcal{P}$ is dense in L^2 (see [30, ch.15]),

ii) $\mathcal{P}$ is a nuclear LF-space [30, ch.51] and thus, complete and barreled,

iii) $\mathcal{P}$ is stable with respect to the Frobenius-Perron operator U, and

iv) U is continuous with respect to the topology of $\mathcal{P}$, because U preserves the degree of polynomials.

It is, therefore, an appropriate rigged Hilbert space, which gives meaning to the spectral decomposition of V.

We shall, however, look for a tight rigging. The test functions should at least provide a domain for the Euler-MacLaurin summation formula (15.5). The requirement of absolute convergence of the series (15.5) means that

$$\sum_{n=1}^{\infty} \Big| \frac{\phi^{(n-1)}(y)}{n!} B_n(x) \Big| < \infty \qquad (y = 0, 1)$$

This implies [18] that the appropriate test functions are restrictions on [0,1) of entire functions of exponential type c with $0 < c < 2\pi$. For simplicity we identify the test functions space with the space $\mathcal{E}_c$ of entire functions $\phi(z)$ of exponential type $c > 0$ such that

$$|\phi(z)| \leq K e^{c|z|} \; , \quad \forall z \in \mathbb{C}, \text{ for some } K > 0 \, .$$

Each member of the whole family $\mathcal{E}_c$, $0 < c < 2\pi$ is a suitable test function space, since properties 1–4 are fulfilled. Indeed, each space $\mathcal{E}_c$ is a Banach space with norm [30, ch.22]:

$$\|\phi\|_c \equiv \sup_{z \in \mathbb{C}} \big|\phi(z)\big| e^{-c|z|} \; ,$$

which is dense in the Hilbert space L^2, as $\mathcal{E}_c$ includes the polynomial space $\mathcal{P}$. Each $\mathcal{E}_c$ is stable under the Frobenius-Perron operator U, and it is easily verified that U is continuous on $\mathcal{E}_c$. Now, observe that the spaces are ordered

$$\mathcal{E}_c \subset \mathcal{E}_{c'} \; , \qquad c < c',$$

and consider the space

$$\tilde{\mathcal{E}}_{2\pi} \equiv \bigcup_{c<2\pi} \mathcal{E}_c \, .$$

The space $\tilde{\mathcal{E}}_{2\pi}$, also preserved by U, is the (set-theoretically) largest test function space in our case. Since $\tilde{\mathcal{E}}_{2\pi}$ is a natural generalization of the space $\mathcal{P}$ of polynomials, we want to equip it with a topology which is a generalization of the topology of $\mathcal{P}$.

Recall that $\mathcal{P}$ was given the strict inductive limit topology of the spaces $\mathcal{P}^n$ of all polynomials of degree $\leq n$. A very important property of this topology is that the strict inductive limit of complete spaces is complete. Moreover, it is exceptionally simple to describe convergence in this topology. For example, a sequence $\{w_n\}$ of polynomials converges in $\mathcal{P}$ if and only if the degrees of all w_n are uniformly bounded by some n_0 and $\{w_n\}$ converges in $\mathcal{P}^{n_0}$.

We cannot, however, define the strict inductive topology on $\tilde{\mathcal{E}}_{2\pi}$, because for $c < c'$ the topology on $\mathcal{E}_c$ induced by $\mathcal{E}_{c'}$ is essentially stronger than the initial one. Nevertheless, as we shall see in the theorem below, it is possible to define a topology on $\tilde{\mathcal{E}}_{2\pi}$, which is a natural extension of the topology on $\mathcal{P}$ in the following sense

Theorem 15.1 *There is a locally convex topology $\mathcal{T}$ on $\tilde{\mathcal{E}}_{2\pi}$ for which it is a nuclear, complete Montel space. Moreover, a sequence $\{f_n\} \subset \tilde{\mathcal{E}}_{2\pi}$ is convergent in the $\mathcal{T}$ topology if and only if there is $c_0 \in (0, 2\pi)$ such that*

1^o $f_n,\ n = 1, 2, \ldots$, *are of exponential type c_0*

2^o $\{f_n\}$ *converges in $\|\cdot\|_{c_o}$ - norm.*

Proof. Denote by $\hat{f}$ the Fourier transform of a function f and by $\check{f}$ its converse. By Schwartz's extension of the Paley-Wiener Theorem [31, vol.II, p.106] a function f belongs to $\mathcal{E}_c$ if and only if $\hat{f}$ is a distribution with compact support contained in the interval $[-c, c]$.

Note that, if the function $f \in \tilde{\mathcal{E}}_{2\pi}$ is integrable or square integrable then $\hat{f}$ is a function. However, for an arbitrary function its Fourier transform is correctly defined only as a distribution with compact support, i.e. as a continuous linear functional on the space $C^\infty(\Omega)$ of all infinitely differentiable functions on the interval $\Omega = (-2\pi, 2\pi)$, endowed with the topology of uniform convergence on compact subsets of Ω of functions together with all their derivatives.

The Fourier transform, therefore, establishes an isomorphism between $\tilde{\mathcal{E}}_{2\pi}$ and the topological dual $C^\infty(\Omega)^\times$ of the space $C^\infty(\Omega)$. Consequently, the strong dual topology of $C^\infty(\Omega)^\times$ can be transported through the inverse Fourier transform to the space $\tilde{\mathcal{E}}_{2\pi}$. The strong dual topology is the topology of uniform convergence on bounded subsets of $C^\infty(\Omega)$. Then $C^\infty(\Omega)^\times$ is nuclear [30, p.530], complete [31, vol.I, p.89] and a Montel space [30, prop. 34.4 and 36.10]. In this way we obtain on $\tilde{\mathcal{E}}_{2\pi}$ a topology with the same properties.

We shall now prove the second part of the theorem. Let $\{f_n\}$ be convergent to zero in $\tilde{\mathcal{E}}_{2\pi}$. This means that $\{\check{f}_n\}$ converges in $C^\infty(\Omega)^\times$. Therefore $\{\check{f}_n\}$ is a bounded subset of $C^\infty(\Omega)^\times$, which implies [30, th. 34.4, p.359] that the supports of all $\check{f}_n$ are contained in a compact set $K \subset \Omega$.

Take c with $c < 2\pi$ and $K \subset (-c, c)$. Therefore (see [31, vol.I, th.XXVI] and the remark afterwards which remain true if we replace $\mathbb{R}^1$ by the open set $\Omega = (-2\pi, 2\pi)$), there is a number $p \geq 0$ and a family of continuous functions $g_{j,n}$ such that the supports of $g_{j,n}$ are contained in the interval $(-c, c)$,

$$\check{f}_n = \sum_{j \leq p} D^j g_{j,n}$$

(D^j denotes the j-th derivative, classical or in the sense of distributions) and $g_{j,n}(x)$ converges uniformly to zero as $n \to \infty$.

Using the above representation of $\check{f}_n$ we obtain that $\check{f}_n$ converges to zero uniformly on each set U_A

$$U_A \equiv \{f \in C^\infty(\Omega) : \sup_{x \in [-c,c]} |\frac{d^j}{dx_j} f(x)| \leq A, \ \ j = 0, 1, \ldots, p\} ,$$

where $A > 0$. Indeed, for each j

$$\begin{aligned} |\langle D^j g_{j,n}, f\rangle| &= |(-1)^j \textstyle\int_{-c}^{c} g_{j,n}(x) \frac{d^j}{dx_j} f(x)\, dx| \\ &\leq A \textstyle\int_{-c}^{c} |g_{j,n}(x)|\, dx \to 0 \ , \end{aligned}$$

as $n \to \infty$.

Let us take any $c_0 \in (c, 2\pi)$. Then for each $z \in \mathbb{C}$ the function

$$x \longmapsto e^{izx} e^{-c_0|z|}, \ \ |x| \leq c \ , \tag{15.6}$$

belongs to U_A. Indeed

$$\begin{aligned} \left|\tfrac{d^j}{dx_j}(e^{izx} e^{-c_0|z|})\right| &\leq |z|^j e^{|z|(|x| - c_0)} = |z|^j e^{|z|(|x|-c)} e^{-(c_o - c)|z|} \\ &\leq |z|^j e^{-(c_o - c)|z|} \ . \end{aligned}$$

The right hand side is bounded, for each $j = 0, 1, ..., p$, by some constant A_j. Thus taking $A = \max_{0 \leq j \leq p} A_j$ we see that the functions (15.6) belong to U_A, for each $z \in \mathbb{C}$.

From

$$f_n(z) = (\check{f}_n)\hat{}\,(z)$$

and uniform convergence of $\check{f}_n$ on U_A we have

$$\sup_{z \in \mathbb{C}} |f_n(z)| e^{-c_0|z|} = \sup_{z \in \mathbb{C}} |\langle \check{f}_n, e^{iz\cdot} e^{-c_0|z|}\rangle| \longrightarrow 0 \ ,$$

as $n \to \infty$, which means that $\|f_n\|_{c_0} \longrightarrow 0$. This proves 2^o. Condition 1^o is also satisfied because we have chosen $c_0 > c$. Thus, the supports of the $\check{f}_n$s are also contained in $(-c_0, c_0)$ and by the Paley-Wiener-Schwartz theorem the f_ns are of exponential type c_0.

The converse of the second part of the theorem is now trivial. If $\{f_n\}$ satisfies $1^{\rm o}$ and $2^{\rm o}$ then by applying the Paley-Wiener-Schwartz theorem again we obtain convergence of $\{\hat{f}_n\}$ in $C^\infty(\Omega)^\times$.

Remarks

1. Using the above method one can show an analogous criterion of convergence for bounded nets in $\tilde{\mathcal{E}}_{2\pi}$ but not for an arbitrary net.
2. Note that it is not always possible to obtain convergence of the type given in the above theorem. Actually, to prove the second part of the theorem we needed the following property:

Let F be a Frechet space and let $\{x'_n\}$ be a sequence in its dual F' which converges to zero in the strong dual topology. Then there exists an open subset U of F such that

$$|\langle x'_n, x\rangle| \longrightarrow 0\ , \quad \text{uniformly for } x \in U\ . \tag{15.7}$$

As mentioned in [32], some concrete F-spaces have this property although it is not true in general. It was stated there as an open problem to describe those F-spaces for which (15.7) is true. This situation motivated us to include the full proof.

Let us review briefly another generalized spectral decompositions

The tent maps

The family of tent maps is defined by:

$$T_m : [0,1) \longrightarrow [0,1)$$

$$T_m = \begin{cases} m\left(x - \frac{2n}{m}\right), & \text{for} \quad x \in \left[\frac{2n}{m}, \frac{2n+1}{m}\right) \\[2ex] m\left(\frac{2n+2}{m} - x\right), & \text{for} \ \ x \in \left[\frac{2n+1}{m}, \frac{2n+2}{m}\right) \end{cases}$$

where $m = 2, 3, \cdots$, $n = 0, 1, \cdots, \left[\frac{m-1}{2}\right]$ and $[y]$ denotes the integer part of real number y. The case $m = 2$ corresponds to the well known tent map. The absolutely continuous invariant measure is the Lebesgue measure dx for all maps T_m. The Frobenius-Perron operator for the tent maps has the form

$$U_T\rho\,(x) = \frac{1}{m}\left\{\sum_{n=0}^{\left[\frac{m-1}{2}\right]} \rho\left(\frac{2n+x}{m}\right) + \sum_{n=0}^{\left[\frac{m-2}{2}\right]} \rho\left(\frac{2n+2-x}{m}\right)\right\}$$

The spectrum consists of the eigenvalues [21]:

$$z_i = \frac{1}{m^{i+1}}\left\{\left[\frac{m-1}{2}\right] + 1 + (-1)^i\left(\left[\frac{m-2}{2}\right] + 1\right)\right\}$$

which means that for the even tent maps, $m = 2, 4, \cdots$, the eigenvalues are

$$z_i = \begin{cases} \frac{1}{m^i}, & i \text{ even} \\ 0, & i \text{ odd} \end{cases} \tag{2.4}$$

and for the odd tent maps, $m = 3, 5, \cdots$, the eigenvalues are

$$z_i = \begin{cases} \frac{1}{m^i}, & i \text{ even} \\ \frac{1}{m^{i+1}}, & i \text{ odd} \end{cases} \tag{2.5}$$

The eigenvectors of the tent maps can be expressed in terms of the Bernoulli and Euler polynomials.
For even tent maps $m = 2, 4, \cdots$

$$f_i(x) = \begin{cases} B_i\left(\frac{x}{2}\right), & i = 0, 2, 4, \cdots \\ \frac{i+1}{2m^{i+1}} E_i(x), & i = 1, 3, 5, \cdots \end{cases}$$

For odd tent maps $m = 3, 5, \cdots$:

$$f_i(x) = \begin{cases} 1, & i = 0 \\ E_i(x), & i = 1, 3, 5, \cdots \\ B_i(x) + E_{i-1}(x), & i = 2, 4, \cdots \end{cases}$$

The Bernoulli polynomials are defined by the generating function (2.3). The Euler polynomials are defined by the generating function

$$\frac{2e^{xt}}{e^t + 1} = \sum_{n=0}^{\infty} E_n(x) \frac{t^n}{n!}, \qquad |t| < \pi.$$

The left eigenvectors are given by the following formulas
For even tent maps $m = 2, 4, \cdots$

$$F_i(x) = \begin{cases} m^i \widetilde{B}_i(x), & i = 0, 2, 4, \cdots \\ m^{i+1}\left(\frac{2}{i+1}\widetilde{E}_i(x) + \widetilde{B}_{i+1}(x)\right), & i = 1, 3, 5, \cdots \end{cases}$$

For odd tent maps $m = 3, 5, \cdots$:

$$F_i(x) = \begin{cases} 1, & i = 0 \\ \widetilde{E}_i(x), & i = 1, 3, 5, \cdots \\ \widetilde{B}_i(x) + \widetilde{E}_{i-1}(x), & i = 2, 4, \cdots \end{cases}$$

The expressions $\widetilde{B}_i(x)$, $\widetilde{E}_i(x)$ are given by the following formulas as in the case of the Renyi maps

$$\widetilde{B}_i(x) = \begin{cases} 1, & i = 0 \\ \frac{(-1)^{i-1}}{i!}\left\{\delta^{(i-1)}(x-1) - \delta^{(i-1)}(x)\right\}, & i = 1, 2, \cdots \end{cases}$$

$$\widetilde{E}_i(x) = (-1)^i 2(i!)\left\{\delta^{(i)}(x-1) + \delta^{(i)}(x)\right\}, \quad i = 0, 1, \cdots.$$

From Eqs. (2.4) and (2.5) we observe that the spectrum of the symmetric tent maps $m = 2, 4, \ldots$, does not contain the odd powers of $\frac{1}{m}$. This is a general property of symmetric maps and is due to the fact that the antisymmetric eigenfunctions are in the null space of the Frobenius-Perron operator. Summarizing, the spectra of the tent maps T_m depend upon m but the eigenvectors depend only on the evenness of m.

The logistic map

The logistic map in the case of fully developed chaos is defined by

$$S(x) = 4x(1-x), \quad \text{for } x \in [0,1] \ .$$

The logistic map is a typical example of exact system. The invariant measure for the logistic map is [2] $d\nu(x) = \frac{1}{\pi\sqrt{x(1-x)}}dx$.

The spectral decomposition of the logistic map can be obtained from the spectral decomposition of the dyadic tent map through the well known topological equivalence of these transformations [22]. The transformation $g : [0,1) \to [0,1)$ defined by:

$$g(x) = \frac{1}{\pi}\arccos(1-2x)$$

defines a topological equivalence between the logistic map and the dyadic tent map $T : [0,1) \to [0,1)$

$$T(x) = \begin{cases} 2x, & \text{for } x \in \left[0, \frac{1}{2}\right) \\ 2(1-x), & \text{for } x \in \left[\frac{1}{2}, 1\right) \end{cases}$$

expressed through the formula

$$S = g^{-1} \circ T \circ g.$$

The transformation g transforms the Lebesgue measure which is the invariant measure of the tent map to the invariant measure of the logistic map.

The transformation G intertwines the Koopman operator V of the logistic map with the Koopman operator V_T of the tent map

$$V = GV_TG^{-1}.$$

The intertwining transformation G and G^{-1} suitably extended, map the eigenvectors of V_T onto the eigenvectors of V. Therefore

$$V = \sum_{n=0}^{+\infty} z_n G\,|\Phi_n)\,(\varphi_n|\,G^{-1}$$

$$V = \sum_{n=0}^{+\infty} z_n\,|F_n)\,(f_n| \tag{2.6}$$

with

$$z_n = \frac{1}{2^{2n}}\,,$$

$$F_n\left(x\right) = 2^{2n}\widetilde{B}_{2n}\left(\frac{1}{\pi}\arccos\left(1-2x\right)\right)\,,$$

$$f_n\left(x\right) = B_{2n}\left(\frac{1}{2\pi}\arccos\left(1-2x\right)\right)\,.$$

In formula (2.6) the bras and kets correspond to the invariant measure of the logistic map. The meaning of the spectral decomposition (2.10) is inherited from the meaning of the spectral decomposition of the tent map in terms of the dual pair of polynomials, i.e. it can be understood in the sense

$$(\rho|Vf) = \sum_{n=0}^{+\infty}\frac{1}{2^{2n}}\left(\rho\left|2^{2n}\widetilde{B}_{2n}\left(\frac{1}{\pi}\arccos\left(1-2x\right)\right)\right)\left(B_{2n}\left(\frac{1}{2\pi}\arccos\left(1-2x\right)\right)\right|f\right)\,,$$

for any state ρ in the space $\mathcal{P}_{\left(\frac{1}{\pi}\arccos(1-2x)\right)}$, and any observable f in the anti-dual space ${}^{\times}\mathcal{P}_{\left(\frac{1}{\pi}\arccos(1-2x)\right)}$.

The baker's transformations

The β-adic, $\beta = 2, 3, \ldots$, baker's transformation B on the unit square $[0,1)\times$[0,1) is a two-step operation: 1) squeeze the 1×1 square to a $\beta \times 1/\beta$ rectangle and 2) cut the rectangle into β $(1 \times 1/\beta)$-rectangles and pile them up to form another 1×1 square

$$(x,y) \longmapsto B(x,y) = (\beta x - r, \frac{y+r}{\beta}) \quad (\text{for } \frac{r}{\beta} \leq x < \frac{r+1}{\beta},\ r = 0, \cdots \beta - 1)\,. \tag{2.7}$$

The invariant measure of the β-adic baker transformation is the Lebesgue measure on the unit square. The Frobenius-Perron and Koopman operators are unitary on the Hilbert space $L^2 = L^2_x \otimes L^2_y$ of square integrable densities over the unit square and has countably degenerate Lebesgue spectrum on the unit circle plus the simple eigenvalue 1 associated with the equilibrium (as is the case for all Kolmogorov automorphisms).

The Koopman operator V has a spectral decomposition involving Jordan blocks, which was obtained [7,23] using a generalized iterative operator method based on the subdynamics decomposition

$$V = |F_{00}\rangle\langle f_{00}| + \sum_{\nu=1}^{\infty}\left\{\sum_{r=0}^{\nu}\frac{1}{\beta^{\nu}}\ |F_{\nu,r}\rangle\langle f_{\nu,r}| + \sum_{r=0}^{\nu-1}|F_{\nu,r+1}\rangle\langle f_{\nu,r}|\right\}, \tag{2.8}$$

The vectors $|F_{\nu,r}\rangle$ and $\langle f_{\nu,r}|$ form a Jordan basis

$$\langle f_{\nu,r}|\, V = \begin{cases} \dfrac{1}{\beta^\nu}\{\langle f_{\nu,r}| + \langle f_{\nu,r+1}|\} & (r = 0,\ldots,\nu-1) \\ \dfrac{1}{\beta^\nu}\langle f_{\nu,r}| & (r=\nu) \end{cases}$$

$$V\,|F_{\nu,r}\rangle = \begin{cases} \dfrac{1}{\beta^\nu}\{\,|F_{\nu,r}\rangle + |F_{\nu,r-1}\rangle\} & (r = 1,\ldots,\nu) \\ \dfrac{1}{\beta^\nu}\,|F_{\nu,r}\rangle & (r=0) \end{cases}$$

$$\langle f_{\nu,r}|\, F_{\nu',r'}\rangle = \delta_{\nu\nu'}\delta_{rr'}\ , \quad \sum_{\nu=0}^{\infty}\sum_{r=0}^{\nu} |F_{\nu,r}\rangle\langle f_{\nu,r}| = I\,.$$

While the Koopman operator V is unitary in the Hilbert space L^2 and thus has spectrum on the unit circle $|z| = 1$ in the complex plane, the spectral decomposition (2.8) includes the numbers $1/\beta^\nu < 1$ which are not in the Hilbert space spectrum. The spectral decomposition (2.8) also shows that the Frobenius–Perron operator has Jordan-block parts despite the fact that it is diagonalizable in the Hilbert space. As the left and right principal vectors contain generalized functions, the spectral decomposition (2.8) has no meaning in the Hilbert space L^2.

That the principal vectors $f_{\nu,j}$ and $F_{\nu,j}$ are linear functionals over the spaces $L^2_x \otimes \mathcal{P}_y$ and $\mathcal{P}_x \otimes L^2_y$, respectively [7].

15.2 RELATION BETWEEN SPECTRAL AND SHIFT REPRESENTATIONS

First, let us remind some spectral properties of unilateral shift operators. It is well known that the spectrum of any unilateral shift V and the spectrum of the adjoint shift V^* coincide with their essential spectrum which is the closed unit disk $\overline{D} = \{z \in \mathbb{C}\,|\,|z| \le 1\}$. The unit circle ∂D is the continuous spectrum of both V and $V^\dagger$. The open unit disk $D = \{z \in \mathbb{C}\,|\,|z| < 1\}$ is the point spectrum of $V^\dagger$ with multiplicity equal to the multiplicity of the shift V. The disk D coincides also with the residual spectrum of V.

The following theorem gives an explicit representation of eigenvectors of $V^\dagger$ in terms of the generating basis.

Theorem 15.2 *Let V be a unilateral shift of multiplicity m on a Hilbert space $\mathcal{H}$ and $\{g_a^n\,|\,n = 0,1,2,\ldots,1 \le a < m+1\}$ be a generating basis for V. Then for any $z \in \mathbb{C}$ with $|z| < 1$ the corresponding eigenspace of the adjoint shift $V^\dagger$ is:*

$$\{f \in \mathcal{H}\,|\,V^\dagger f = zf\} = \left\{ \sum_{1\le a<m+1}\sum_{n=0}^{\infty} c_a z^n g_a^n \;\middle|\; \sum_{1\le a<m+1} |c_a|^2 < \infty \right\}.$$

Proof. First, let $z \in D$ and

$$f = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n g_a^n \,,$$

where $\sum_{1 \le a < m+1} |c_a|^2 < \infty$. The series $\sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n g_a^n$ converges unconditionally in $\mathcal{H}$ because g_a^n is an orthonormal basis and $\sum_{1 \le a < m+1} \sum_{n=0}^{\infty} |c_a z^n|^2 < \infty$. Therefore f is a well-defined element of $\mathcal{H}$ and

$$V^\dagger f = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n V^\dagger g_a^n =$$

$$= \sum_{1 \le a < m+1} \sum_{n=1}^{\infty} c_a z^n V^\dagger g_a^{n-1} = z \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n g_a^n = zf.$$

Thus the inclusion

$$\left\{ \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n g_a^n \,\Big|\, \sum_{1 \le a < m+1} |c_a|^2 < \infty \right\} \subseteq \{ f \mid V^\dagger f = zf \}$$

is proved. Now let $f \in \mathcal{H}$, $z \in D$, $V^\dagger f = zf$, and $(f, g_a^n) = f_a^n$. Then $f = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} f_a^n g_a^n$. Therefore

$$V^\dagger f = \sum_{1 \le a < m+1} \sum_{n=1}^{\infty} f_a^n g_a^{n-1} = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} f_a^{n+1} g_a^n.$$

By the equality $V^\dagger f = zf$ we have $f_a^{n+1} = z f_a^n$ for all a. Hence, $f_a^n = z^n f_a^0$. Denoting f_a^0 by c_a we obtain $f_a^n = z^n c_a$. Prom Parseval's equality $\|f\|^2 = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} |f_a^n|^2$ we get that $\sum_{1 \le a < m+1} |c_a|^2 < \infty$. So, $f = \sum_{1 \le a < m+1} \sum_{n=0}^{\infty} c_a z^n g_a^n$, where $\sum_{1 \le a < m+1} |c_a|^2 < \infty$. The opposite inclusion is also proved.

Remark When the multiplicity of V is finite the condition $\sum_{1 \le a < m+1} |c_a|^2 < \infty$ is always satisfied.

Theorem 15.3 *Let V be a unilateral shift on a Hilbert space $\mathcal{H}$ with multiplicity m. If the vectors $\{f_a : 1 \le a < m+1\} \subset \mathcal{H}$ are eigenvectors of $V^\dagger$, associated to z_a, where*

$z_a \in \mathbb{C}$ *and* $|z_a| < 1$*:* $V^\dagger f_a = z_a f_a$*. Then* $h_a = f_a - z_a V f_a \in \mathcal{N}_0 = \operatorname{Null} V^\dagger$*. Moreover if* h_a *are linearly independent and their linear hull is dense in* $\operatorname{Null} V^\dagger$ *then the set of vectors*

$$g_a = \sum_{l=1}^{a} c_a^l h_l,$$

obtained by applying the Gramm–Schmidt orthonormalization algorithm to h_a *is orthonormal basis in* $\operatorname{Null} V^\dagger$ *and*

$$\{g_a^n = V^n g_a : n = 0, 1, 2, \ldots, \, 1 \le a < m+1\}$$

is a generating basis for V*.*

Proof Since $V^\dagger f_a = z_a f_a$ we have that $V^\dagger h_a = V^\dagger(f_a - z_a V f_a) = V^\dagger f_a - z_a f_a$. Therefore $h_a \in \operatorname{Null} V^\dagger = \mathcal{N}_0$. If h_a are linearly independent then the Gramm–Schmidt orthonormalization algorithm can be applied to h_a. Let

$$g_a = \sum_{l=1}^{a} c_a^l h_l$$

be the result of this algorithm. If the linear hull of h_a is dense in $\operatorname{Null} V^\dagger$ then the linear hull of the orthonormal system g_a is dense in $\mathcal{H}$. Thus $\{g_a \,|\, 1 \le a < m+1\}$ is the orthonormal basis in $\mathcal{N}_0$. It remains to apply Proposition 6.1.

In Section 7 we have constructed a time operator of the Renyi map using the Walsh basis as a generating basis for the semigroup $\{V^n : n = 0, 1, 2, \ldots\}$, where V is the Koopman operator. We shall obtain now another generating basis starting from Bernoulli polynomials and using Theorem 15.3. The *Bernoulli polynomials* $B_n(x)$, $n = 0, 1, \ldots$, are defined by the generating expansion

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!} \,, \quad |t| < 2\pi \,,$$

for all values of x. It lows from the *multiplication theorem* (Olver 1974):

$$B_n(mx) = m^{n-1} \sum_{k=0}^{m-1} B_n\left(x + \frac{k}{m}\right) \,, \quad m = 0, 1, \ldots \,.$$

that B_n are eigenfunctions of the Frobenius-Perron operator of the Renyi map:

$$UB_n = 2^{-n} B_n \,.$$

It is easy to see that the function $h_n = B_n - 2^{-n} V B_n$ for $n = 1, 2, \ldots$ has the form

$$h_n = \begin{cases} P_{n-1}(x), \text{ for } x \in (0, 1/2), \\ -P_{n-1}(x - 1/2), \text{ for } x \in (1/2, 1), \end{cases}$$

where P_n is a polynomial of degree n. Then applying Gramm–Schmidt algorithm to h_n we arrive at the functions

$$g_a = \begin{cases} L_{n-1}(2x), \text{ for } x \in (0, 1/2), \\ -L_{n-1}(2x-1), \text{ for } x \in (1/2, 1), \end{cases}$$

where L_n, $n = 1, 2, \ldots$ is a sequence of polynomials ($\deg L_n = n$) orthonormal in $L^2_{[0,1]}$. Therefore L_n are the normalized Legendre polynomials:

$$L_n(x) = \sqrt{\frac{2n+1}{n}} \frac{1}{2^n n!} \frac{d^k}{dx^k} \left((x^2-1)^k\right) .$$

By Theorem 15.3, $\{g_n\}$ is an orthonormal basis in $\mathcal{N} \equiv \mathrm{Null}\, V^*$. It follows from Theorem 6.1 that the set

$$\{g_n^k \,:\, g_n^k = V^k g_n \,,\ n \in \mathbb{N} \,,\ k = 0, 1, 2, \ldots\} \tag{15.8}$$

is a generating basis for V on $L^2_{[0,1]} \ominus [1]$.

The generalized Fourier transform, relating the spectral representation furnished by the Bernoulli polynomials B_a with the shift representation furnished by the age eigenfunctions g_a^n (15.8) is obtained by Theorem 15.2:

$$B_a = \sum_{k=0}^{\infty} \sum_{m=1}^{a} c_a^m 2^{-ak} V^k g_a \,, \tag{15.9}$$

where $c_a^m = 2 \int_0^{1/2} (B_a(x) - 2^{-a} B_a(2x)) L_{a-1}(2x)\, dx$.

NOTES

1. Problems related to generalized spectral decompositions:

(i) *A general approach to the problem of rigging* that would allow to derive a generalized spectral decomposition of an arbitrary Frobenius-Perron (Gelfand-Maurin approach can not be applied here because the operators are non-self-adjoint). This problem has been partially resolved in [AT].

(ii) *Appropriate choice of the space* Φ. The dual $\Phi^\times$ should not be too big since otherwise the extended operator V can have eigenvalues not belonging to the Hilbert space spectrum of the operator. In particular, the existence of the tightest rigging. However, the choice of Φ is strictly model-dependent. The raised above problem of tightest rigging has been resolved in the case of the Renyi map.

(ii) *Uniqueness of generalized spectral decompositions.* Spectral decompositions are not unique. For example, in the case of the Renyi map the spectral decomposition with eigenvalues 2^{-n} is associated with analytic observables. The choice of Cantor type observables allows to obtain the spectral decomposition with eigenvalues 3^{-n}.

2. The functions g_a^n (15.8) are another age eigenstates for the uniform time operator of the Renyi map, which do not correspond to the natural projection of the Baker's age eigenstates found in [Pr].

3. Formula (15.9) allows to obtain spectral representation from shift representation. Both strategies are available for the probabilistic prediction of chaotic dynamical systems.

4. (Cusp map). Although the spectral representation is formally equivalent to the shift representation, in the case of the cusp map, the spectral analysis seems to be quite involved and not possible for analytic functions [AnShYa]. However the shift representation (8.8) is quite simple and explicit, allowing for probabilistic predictions.

Appendix A
Probability

A.1 PRELIMINARIES - PROBABILITY

A probability space $(\Omega, \mathcal{F}, P)$ consists of a space Ω with points ω (*elementary events*), σ-algebra $\mathcal{F}$ of sets in Ω (*events*) and a probability measure P on $\mathcal{F}$ (*probability*). Let $(\mathcal{X}, \Sigma)$ be a measurable space, where $\mathcal{X}$ is a set and Σ a σ-algebra of its subsets. An $\mathcal{X}$*-valued random variable* $X = X(\omega)$ is a $(\mathcal{F}, \Sigma)$ measurable function $X : \Omega \to \mathcal{X}$, i.e. such function for which the counter-image $X^{-1}(A)$ of every set $A \in \Sigma$ belongs to $\mathcal{F}$. If $\mathcal{X}$ is the set of real numbers and Σ its Borel σ-algebra, i.e. the smallest σ-algebra containing all intervals, then the $\mathcal{X}$-valued random variable X is called simply the *random variable*. If $\mathcal{X} = \mathbb{R}^n$ then X is called the *random vector*. If $\mathcal{X} = \mathbb{C}$ - the set of complex numbers then X is called the *complex random variable*.

If X is an $\mathcal{X}$ valued random variable then the family of all sets: $X^{-1}(A),\ A \in \Sigma$, forms also a σ-algebra, which will be denoted by $\sigma(X)$. Thus $\sigma(X) \subset \mathcal{F}$ and $\sigma(X)$ is the smallest σ-algebra of subsets of $\mathcal{X}$ with respect to which the function X is measurable. If $\{X_i\}_{i \in I}$ is an arbitrary family of $\mathcal{X}$-valued random variables then $\sigma\{X_i : i \in I\}$ will denote the smallest σ-algebra containing all the $\sigma(X_i),\ i \in I$.

If X is a random variable, then the function

$$F(x) = P\{\omega : X(\omega) \leq x\}$$

is non-decreasing and continuous to the right, which tends to 0 as $x \to -\infty$ and to 1 as $x \to \infty$. The function F is called the distribution function of the random variable X. The knowledge of $F(x)$ for all x determines the probability $P(X \in A)$, for each Borel set $A \subset \mathbb{R}$, which in turn determines the Borel measure $\mu = \mu_x$ on $\mathbb{R}$:

$$\mu(A) = P(X \in A)$$

called the *distribution* of X. The measure μ can be, as every Borel measure, decomposed into three parts

$$\mu = \mu_{\mathrm{d}} + \mu_{\mathrm{sc}} + \mu_{\mathrm{ac}}$$

that consists of a discrete, singular continuous and absolutely continuous measure. The random variable for which the distribution consists of only one component will be called discrete, singular continuous and absolutely continuous respectively. If the random variable X is discrete then its distribution function is a step function with a finite or denumerable number of steps. If X is absolutely continuous then there exists a nonnegative function $f(x)$ called the *probability density* of X such that

$$F(x) = \int_{-\infty}^{x} f(y)\, dy\,.$$

The *expectation* or *mean value* of the random variable X is defined as the Lebesgue, or the Lebesgue-Stieltjes, integrals

$$EX = \int_{\Omega} X(\omega) P(d\omega) = \int_{-\infty}^{\infty} x\, dF(x)\,.$$

provided the integral exists.

If h is a Borel measurable function on $\mathbb{R}$ then the expectation of the function $h(X)$ of X is defined as

$$Eh(X) = \int_{-\infty}^{\infty} h(x) dF(x)\,.$$

the mean value $EX^n,\ n \in \mathbb{N}$, is called the n*-th moment* of X. The mean value of $E(X - EX)^2$ is called the *variance* of X and is denoted by $\mathrm{Var} X$.

If $X_1, \ldots, X_n$ are random variables defined on the space Ω, they jointly induce the distribution function

$$F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$$

in n-dimensional Euclidean space $\mathbb{R}^n$. Similarly, as in one-dimensional case, the distribution function F determines the probability $P((X_1, \ldots, X_n) \in A)$, for each Borel set $A \subset \mathbb{R}^n$. Consequently, F determines the Borel measure μ on $\mathbb{R}^n$ that can be interpreted as the distribution of the *random vector* $(X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$.

Let $X_1, \ldots, X_n$ be random variables with the distributions $F_1(x_1), \ldots, F_n(x_n)$ and let $F(x_1, \ldots, x_n)$ be the joint distribution of the random vector $(X_1, \ldots, X_n)$. If

$$F(x_1, \ldots, x_n) = F(x_1) \ldots F(x_n)\ , \quad \text{for every } x_1, \ldots, x_n \in \mathbb{R}\,,$$

then the random variables $X_1, \ldots, X_n$ are called to be *independent*.

If h is a Borel measurable function on $\mathbb{R}^n$ then the function $h(X_1, \ldots, X_n)$ is a random variable and its expectation is defined as

$$Eh(X_1, \ldots, X_n) = \int_{\mathbb{R}^n} h(X_1, \ldots, X_n)\, dF(X_1, \ldots, X_n)\,.$$

In particular if X and Y are random variables and $h(x, y) = (x - EX)(y - EY)$ then the expectation of $h(X, Y)$

$$Eh(X, Y) = E(X - EX)(Y - EY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - EX)(y - EY)\, dF(x, y)$$

is called the *covariance* of X and Y and denoted by $\mathrm{Cov}(X, Y)$.

The value $\dfrac{\mathrm{Cov}(x, y)}{\sqrt{\mathrm{Var}X\mathrm{Var}Y}}$ is called the *correlation coefficient*.

If X and Y are random variables then $X + iY$ defines the *complex random variable* on $\mathbb{R}$. The distribution of $X + iY$ is identified with the joint (X, Y)-distribution.

The characteristic function $\varphi = \varphi_x$ of a random variable X is defined as the mean value of the function e^{itX}:

$$\varphi(t) = Ee^{itX} = \int_{-\infty}^{\infty} e^{itX}\, dF(x)\,.$$

The characteristic function is uniformly continuous on the real line and such that $\varphi(0) = 1$ and $|\varphi(t)| \leq 1$ for every t. The distribution function F is uniquely determined by the corresponding characteristic function. In particular, we have

$$F(x_2) - F(x_1) = \frac{1}{2\pi i}\lim_{\tau\to\infty}\int_{-\tau}^{\tau}\frac{e^{-itx_1} - e^{-itx_2}}{t}\varphi(t)\, dt\,,$$

for all the points $x_1, x_2 \in \mathbb{R}$ in which F is continuous.

Among the most important properties of characteristic functions are those that concerns sums of independent random variables, moments and convergence of distributions.

Let $X_1, \ldots, X_n$ be independent random variables with distribution functions F_1, ..., F_n and the characteristic functions $\varphi_1, \ldots, \varphi_n$. Then the characteristic function φ of the sum $X_1 + \ldots + X_n$ is:

$$\varphi(t) = \varphi_1(t)\ldots\varphi_n(t)\,.$$

Note that the distribution function F of the sum $X_1 + \ldots + X_n$ is the convolution of $F_1, \ldots, F_n$.

$$F(x) = F_1 * \ldots * F_n(x)$$

Recall that the convolution of two distribution functions F and G is $F * G(x) = \int_{-\infty}^{\infty} F(x - y)\, dG(y)$.

If the random variable X has finite n-th moment, $n \in \mathbb{N}$, then its characteristic function φ is of the form

$$\varphi(t) = \sum_{k=1}^{n} m_k \frac{(it)^k}{k!} + 0(t^n)\,,$$

where $m_k = EX^k$ and $\frac{0(t^n)}{t^n} \to 0$ as $t \to 0$.

Let $X_1, X_2, \ldots$ be random variables with the distribution functions $F_1, F_2, \ldots$. The sequence $\{X_n\}$ of random variables (or equivalently the sequence $\{F_n\}$ of distribution) is said to be convergent *in distributions* if there is a distribution function F such that

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for every point of continuity x of F. A necessary and sufficient condition that $\{F_n\}$ is convergent in distributions is that the corresponding sequence $\{\varphi_n\}$ of characteristic functions is pointwise convergent to a function φ which is continuous at the point 0. If this holds then φ is the characteristic function of the limiting distribution function F.

The concept of characteristic function has a straightforward generalization on random vectors, which is called the *characteristic functional*. The characteristic functional φ of the random vector $(X_1, \ldots, X_n)$ is defined as the expected value

$$\varphi(t_1, \ldots, t_n) = \int_{\mathbb{R}^n} e^{i(t_1x_1 + \ldots + t_nx_n)}\, dF(x_1, \ldots, x_n)\,,$$

is a function of n-variables $(t_1, \ldots, t_n) \in \mathbb{R}^n$. The properties of characteristic functionals are similar as characteristic functions.

One of the most important distribution in probability theory, which will be also encountered in the sequel, is the normal (or Gauss) distribution. The *normal distribution* with parameters (m, σ^2) on the real line is an absolutely continuous distribution with the density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-m)^2}{\sigma^2}}\,.$$

All the moments of the normal distribution are finite. Its mean value is m and the variance is σ^2. The probability density of the normal distribution with the parameters $m = 0$ and $\sigma^2 = 1$ will be denoted by ϕ, i.e.

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

and the corresponding distribution function by $\Phi(x)$, $\Phi(x) = \int_{-\infty}^{\infty} \phi(y)\, dy$. If $X_1, \ldots, X_n$ are independent random variables having normal distributions with parameters $(m_1, \sigma_1^2), \ldots, (m_n, \sigma_2^2)$ respectively, then the sum $X_1 + \ldots + X_n$ has normal distribution with the parameters $(m_1 + \ldots + m_n, \sigma_1^2 + \ldots + \sigma_n^2)$. The importance of the normal distribution stems from the fact that it is a limit distribution of sums of independent random variables.

For example, if $X_i,\ i = 1, \ldots, n$ are arbitrary but independent random variables with mean values m_i and variances $\sigma_i^2,\ i = 1, \ldots, n$, then the sequence $\{Y_n\}$ of random variables

$$Y_n = \frac{X_1 + \ldots + X_n - (m_1 + \ldots + m_n)}{\sqrt{\sigma_1^2 + \ldots + \sigma_n^2}}$$

converges in distribution to the normal distribution with parameters (0,1). This result is known as the *Lindenberg-Lévy theorem*.

The characteristic function φ on the normal distribution with the parameters (m, σ^2) is

$$\varphi(t) = e^{mit - \frac{\sigma^2 t^2}{2}}\ ,\ t \in \mathbb{R}\,.$$

As it was already noted characteristic functions determine uniquely the probability distributions. Using this property it is convenient to define the normal distribution of a random vector. Namely, assume that $\mathbf{A} = [a_{ij}],\ a_{ij} \in \mathbb{R}$, is a $m \times n$-matrices, which is symmetric and positive. The function

$$\varphi(t_1, \ldots, t_n) = \exp\left(-i\sum_{k=1}^{n} m_k t_k - \frac{1}{2}\sum_{k,l=1}^{n} a_{kl} t_k t_l\right)\ ,\ (t_1, \ldots, t_n) \in \mathbb{R}^n\,,$$

is a joint distribution of some random variables $X_1, \ldots, X_n$ and is called the *characteristic function of n-dimensional normal (Gaussian) distribution*. Then the numbers m_n are expectations and a_{kl} are the covariances of X_k:

$$m_k = EX_k\ ,\ a_{kl} = E(X_k - m_n)(X_l - m_l)\ ,\ k, l = 1, \ldots, n\,.$$

In the particular case when the matrix **A** has the rank n the probability density of the joint Gaussian distribution has the form

$$f(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{A}|}} \exp\left(-\frac{1}{2|\mathbf{A}|}\sum_{k,l=1}^{n} a_{kl}(x_k - m_k)(x_l - m_l)\right)\,.$$

Stable distributions

We introduce now another important class of distributions, which, like the Gaussian distribution, are limit distributions of sums of independent random variables.

Let X be a nondegenerate random variable, i.e. X is not a constant almost everywhere. Let the random variables X_1 and X_2 be independent copies of X. We say that X (or the distribution of X) is *stable* if for each pair a and b of positive numbers there is a positive number c and a number d such that the distribution of the random variable $aX_1 + bX_2$ coincides with the distribution of $cX + d$. An equivalent definition of stable distribution is that for each number $X_1, \ldots, X_n$ of independent copies of X

there exists numbers $c_n > 0$ and d_n such that the distribution of the sum $X_1, \ldots, X_n$ coincides with the distribution of $c_n X + d_n$. If $d_n = 0$ then X is called strictly stable. It can be proved that (see [Feller], vol.II) that only the norming constants $c_n = n^{\frac{1}{\alpha}}$, where $0 < \alpha \leq 2$, are possible. The number α is called the *characteristic exponent* of the stable distribution and the term α*-stable distribution* is often used.

One of the versions of the central limit theorem says that if $X_1, X_2, \ldots$ are independent identically distributed random variables with mean 0 and variance 1, then the distribution of the sum $\frac{X_1+\ldots+X_n}{\sqrt{n}}$ converges to the normal distribution with mean 0 and variance 1. If the distribution of X_i does not have the second moment then it still may happen that there exist constants $a_n > 0$ and b_n such that the distribution of

$$\frac{X_1 + \ldots + X_n}{a_n} + b_n$$

converge to some distribution R. If this holds then it is said that the distribution of X_i belongs to the *domain of attraction* of R. It can be proved that a distribution R possesses a nonempty domain of attraction if and only if R is stable. In particular each stable distribution belongs to its own domain of attraction. Stable distributions are absolutely continuous but the explicit form of the density function is known only in a few particular cases. It is known however the characteristic function φ of each stable distribution. Namely

$$\varphi(t) = \exp\left\{ i\mu t - \sigma^{\alpha}|t|^{\alpha} \left(1 - i\beta(\mathrm{sgn}t) \tan \frac{\pi\alpha}{2}\right\} \right. ,$$

if $\alpha \neq 1$, $0 < \alpha \leq 2$, and

$$\varphi(t) = \exp\left\{ i\mu t - \sigma|t|(1 + i\beta\frac{2}{\pi}(\mathrm{sgn}t)\ln|t|)\right\}$$

if $\alpha = 1$, where $-1 \leq \beta \leq 1$, $\sigma > 0$ and $\mu \in \mathbb{R}$.

The stable distribution is thus characterized by four parameters α, β, σ and μ, where α is the *characteristic exponent*, β is the *skeweness* parameter, σ is the *scale* parameter and μ is the *location* parameter.

Putting $\alpha = 2$ we obtain the characteristic function of the Gaussian distribution. If $\alpha = 1$ and $\beta = 0$ then $\varphi(t)$ is the characteristic function of the Cauchy distribution, which has the density

$$f(x) = \frac{1}{\pi}\frac{a}{a^2 + (x-b)^2} ,$$

for some numbers a, b.

Particularly important is the class of α-stable distributions with the parameters $\beta = \mu = 0$. In this case the stable random variable is symmetric about 0 and the characteristic function takes the form

$$\varphi(t) = e^{-\sigma^{\alpha}|t|^{\alpha}}$$

for every $\alpha \in [0, 2]$. A characteristic feature of an α-stable random variable with $0 < \alpha < 2$ is that it does not have finite second moment. It can be only proved that

$$E|X|^p < \infty$$

for $0 < p < \alpha$. In particular, if $\alpha \leq 1$ then the expectation of X does not exists. In fact the α-stable distribution of X has for $\alpha < 2$ “heavy tails", i.e. the density of X behaves asymptotically, when $x \to \infty$, as the function $\frac{1}{x^{\alpha+1}}$.

Conditional expectation

One of the basic tools used in this book in the study of both stochastic and deterministic system is the conditional expectation. Let X be random variable on the probability space $(\Omega, \mathcal{F}, P)$ such that $E|X| < \infty$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. The random variable Y is called the *conditional expectation* of X with respect to the σ-algebra $\mathcal{G}$ if it satisfies the following conditions:

1) Y is $\mathcal{G}$-measurable, i.e. Y is random variable on the probability space $(\Omega, \mathcal{G}, P)$

2) $\int_A X\,dP = \int_A Y\,dP$, for every $A \in \mathcal{G}$.

The conditional expectation of X with respect to $\mathcal{G}$ will be denoted by $E(X|\mathcal{G})$. The existence of the conditional expectation follows from the Radon-Nikodym theorem. The function $E(X|\mathcal{G})$ is defined up to sets with probability 1. Conditional expectation has the following properties that are valid P-a.e. on Ω.

1) $E(\cdot|\mathcal{G})$ is a linear operator, i.e. if X are random variables with finite expectation and a, b are real numbers then $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$.

2) If $X_1(\omega) \leq X_2(\omega)$ for almost all $\omega \in \Omega$ then $E(X_1|\mathcal{G}) \leq E(X_2|\mathcal{G})$.

3) $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$.

4) If h is a convex function on $\mathbb{R}$ such that $Eh(X)$ exists then $h(E(X|\mathcal{G})) \leq E(h(X)|\mathcal{G})$.

5) If X is $\mathcal{G}$ measurable function and Y an arbitrary random variable such that $E|XY| < \infty$ then $E(XY|\mathcal{G}) = XE(Y|\mathcal{G})$.

6) In particular, if X is $\mathcal{G}$ measurable then $E(X|\mathcal{G}) = X$.

7) If X is independent with respect to the σ-algebra $\mathcal{G}$, i.e. if $P(X^{-1}(B) \cap A) = P(X^{-1}(B))P(A)$, for each $B \in \mathcal{B}_{\mathbb{R}}$ and $A \in \mathcal{G}$, then

$$E(X|\mathcal{G}) = EX$$

8) If $\{X_n\}$ is a sequence of random variables such that $|X_n| \leq Y$, $n = 1, 2, \ldots$, $EY < \infty$, and $X_n(\omega) \to X(\omega)$ for almost all $\omega \in \mathbb{R}$ then

$$E(X_n|\mathcal{G}) \to E(X|\mathcal{G})$$

P-almost everywhere.

The last property is an analog of the Lebesgue majorized convergence theorem.

A.2 STOCHASTIC PROCESSES

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(\mathcal{X}, \Sigma)$ a measurable space and I an index set. An $\mathcal{X}$-valued stochastic process on $(\Omega, \mathcal{F}, P)$ is a family $\{X_t\}_{t\in I}$ of $(\mathcal{F}, \Sigma)$ measurable functions

$$X_t : \Omega \to \mathcal{X}, \ \ t \in I\,.$$

The index set is assumed to be a subset of real numbers and the set $\mathcal{X}$ an Euclidean space. If $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$, then $\{X_t\}$ is simply called a *stochastic process*. In the case $\mathcal{X} = \mathbb{R}^n$ or $\mathcal{X} = \mathbb{C}^n$ we speak about a *vector valued stochastic process*. For a fixed $\omega \in \Omega$ the function:

$$I \ni t \longmapsto X_t(\omega) \in \mathcal{X}$$

is called a *realisation* or a *trajectory* of the process $\{X_t\}$.

A *finite dimensional distribution* of the $\mathcal{X}$-valued stochastic process $\{X_t\}_{t\in I}$ is defined for a given set $\{t_1, \ldots, t_n\} \in I,\ t_1 < \ldots < t_n$, as the probability measure $\mu_{t_1,\ldots,t_n}$, on the product space measure space $(\mathcal{X}^n, \Sigma^n)$

$$\mu_{t_1,\ldots,t_n}(B) = P\{\omega : (X_{t_1}(\omega), \ldots, X_{t_n}\,, (\omega)) \in B\} \tag{A.1}$$

for all $B \in \Sigma^n$.

Each stochastic process $\{X_t\}_{t\in I}$ determines the family of finite dimensional distributions of the form (A.1). Conversely, suppose that it is given a family of probability measures $\{\mu_{t_1,\ldots,t_n}\}_{t_1,\ldots,t_n\in I}$, $t_1 < \ldots < t_n$, $n \in \mathbb{N}$, defined on the product spaces $(\mathcal{X}^n, \Sigma^n)$, which satisfies the following consistency conditions. For each two time ordered sets of indices $\{s_1, \ldots, s_k\}$ and $\{t_1, \ldots, t_n\}$ such that

$$\{s_1, \ldots, s_k\} \subset \{t_1, \ldots, t_n\}$$

the measure $\mu_{s_1,\ldots,s_n}$ is the projection of the measure $\mu_{t_1,\ldots,t_n}$. The last statement means that $\mu_{s_1,\ldots,s_k}$ coincides with $\mu_{t_1,\ldots,t_n}$ on the sets of the form $B = B_{t_1} \times \ldots \times B_{t_n}$ provided we replace the corresponding B_{t_k} by $\mathcal{X}$. Such family of measures determines a probability measure P on the measure space $(\Omega, \mathcal{F})$, where $\Omega = \mathcal{X}^I (= \prod_{t\in I} \mathcal{X}_t,\ \mathcal{X}_t = \mathcal{X})$, and $\mathcal{F}$ is the σ-algebra generated by the cylinder subsets of $\mathcal{X}$, such that each projection of P onto the product space $\mathcal{X}_{t_1} \times \mathcal{X}_{t_n}$ coincides with $\mu_{t_1,\ldots,t_n}$. This allows, in turn, to determine an $\mathcal{X}$-valued stochastic process $\{X_t\}_{t\in I}$ on $(\Omega, \mathcal{F}, P)$ such that the measures $\{\mu_{t_1,\ldots,t_n}\}$ are its finite dimensional distributions. Indeed, since each element $\omega \in \Omega$ is an $\mathcal{X}$ valued function on I the process $\{X_t\}_{t\in I}$ can be defined by putting

$$X_t(\omega) \stackrel{\mathrm{df}}{=} \omega(t)\,.$$

This result, known as the *Kolmogorov extension theorem*, can be found in textbooks on measure theory or stochastic processes, (e.g. Neveu 1965 or Loevé 1977-78).

Each stochastic process determines its *natural filtration* $\{\mathcal{F}_t\}_{t\in I}$ that can be obtained by putting

$$\mathcal{F}_t = \sigma(X_s)_{s\le t}\,. \tag{A.2}$$

Thus $\mathcal{F}_t$ is the smallest σ-algebra generated by all random variables X_s, for $s \le t$. In general, let $\{\mathcal{F}_t\}_{t\in I}$ be a family of sub-σ-algebras of $\mathcal{F}$. We say that the process $\{X_t\}$ is $\{\mathcal{F}_t\}$ *adapted* if every random variable X_t, $t \in I$, is $\mathcal{F}_t$ measurable, i.e. every X_t is measurable as a function from $(\Omega, \mathcal{F}_t)$ into (X, Σ). Family $\{\mathcal{F}_t\}_{t\in I}$ is called a *filtration* of process $\{X_t\}_{t\in I}$ if:

(1) $\mathcal{F}_s \subset \mathcal{F}_t$, for each $s < t$

(2) $\{X_t\}$ is $\{\mathcal{F}_t\}$ *adapted*.

A filtration $\{\mathcal{F}_t\}_{t\in I}$ is called *right continuous* if for each $t \in I$ we have

$$\mathcal{F}_t = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} . \tag{A.3}$$

It is a common assumption that the filtration $\{\mathcal{F}_t\}$ associated with a stochastic process $\{X_t\}$ is right continuous. Such an assumption is justified because it is always possible to augment the σ-algebras $\mathcal{F}_t$ in such a way that condition (A.3) is satisfied and, in addition, all $\mathcal{F}_t$ are complete (Dellacherie and Meyer 1978, p.115).

A stochastic process $\{X_t\}_{t\in I}$ is called the L^p*-stochastic process*, $p > 0$, if all random variables X_t, $t \in I$ are p-integrable. An L^p-process $\{X_t\}_{t\in I}$ can be interpreted as a trajectory

$$t \longmapsto X_t$$

in the Fréchet space L^p of all $\mathcal{X}$-valued p-integrable random variables endowed with the norm

$$\|X\|_{L^p} = (E|X|^p)^{1\wedge\frac{1}{p}} < \infty$$

(L^p is a Banach space if $p \geq 1$ and a Hilbert space if $p = 2$).

If $p = 1$ (or, more generally, if $p \geq 1$) then an L^p-process is called an *integrable processes*. If a process $\{X_t\}$ is integrable and $\{\mathcal{F}_t\}$ is its filtration, then we can associate with $\{X_t\}$ the family of conditional expectation $\{E_t\}$

$$E_t \overset{\text{df}}{=} E(\,\cdot\,|\mathcal{F}_t) \ , \quad t \in I$$

If $p = 2$ then E_t are orthogonal projectors on the Hilbert space L^2.

In order to deal with continuous time processes it is often necessary to impose some additional assumptions. We introduce below a few of the most important of them.

A stochastic process $\{X_t\}_{t\in I}$ is called *separable* if there exist a countable dense set $I_0 \subset I$ and a set $\Omega_0 \in \mathcal{F}$, $P(\Omega_0) = 0$, such that if $\omega \in \Omega \setminus \Omega_0$ and $t \in I$, then there is a sequence $\{t_n\} \subset I_0$, $t_n \to t$, such that $X_{t_n}(\omega) \to X_t(\omega)$.

Proposition A.1 *For each stochastic process $\{X_t\}_{t\in I}$ there exists a separable process $\{\tilde{X}_t\}_{t\in I}$ defined on the same probability space such that*

$$P\{\tilde{X}_t = X_t\} = 1 \, ,$$

for each $t \in I$.

The assumption of separability is usually accompanied with the assumption that the probability space $(\Omega, \mathcal{F}, P)$ is complete, i.e. $\mathcal{F}$ contains all subsets of the sets of P-measure 0. These two assumptions allow to deal with expressions concerning an uncountable number of random variables X_t. For example, the expressions like $\sup_t X_t$ or $\inf_t X_t$ (X_t are real valued) represents then random variables. Both assumptions are not too restrictive and are usually taken as granted. Indeed, any probability space $(\Omega, \mathcal{F}, P)$ can be always completed and, by the above theorem, any stochastic process has a separable version.

Another important assumption on a stochastic process, which is needed when dealing with integrals of the sample functions, is the measurability. A stochastic process $\{X_t\}_{t\in I}$ is called *measurable* if the function

$$I \times \Omega \ni (t, \omega) \longmapsto X_t(\omega) \in \mathcal{X}$$

is $(\mathcal{B} \otimes \mathcal{F}, \mathcal{B}_{\mathcal{X}})$-measurable, where $\mathcal{B}_{\mathcal{X}}$ is the Borel σ-algebra of subsets of $\mathcal{X}$ (recall that I is an interval in $\mathbb{R}$ and $\mathcal{X}$ is either $\mathbb{R}$ or $\mathbb{C}$).

Theorem A.1 *Assume that the process $\{X_t\}_{t\in I}$ is continuous in probability, i.e. for each $t_0 \in I$ and $\varepsilon > 0$ holds*

$$\lim_{t\to t_0} P\{|X_t - X_{t_0}| \geq \varepsilon\} = 0\,.$$

Then there exists a separable and measurable process $\{\tilde{X}_t\}_{t\in I}$ defined on the same probability space such that

$$P\{X_t = \tilde{X}_t\} = 1\,, \text{ for each } t \in I\,,$$

and such that each countable subset of I is the set of separability for $\{\tilde{X}_t\}_{t\in I}$.

The process $\{\tilde{X}_t\}_{t\in I}$ from the above theorem is called a *separable and measurable modification* of $\{X_t\}_{t\in I}$. The condition of continuity can be easily verified for some classes of stochastic processes. For example, the Brownian motion is continuous in probability.

A.3 MARTINGALES

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{\mathcal{F}_t\}_{t\in I}$ a filtration. An integrable $\{\mathcal{F}_t\}$ adapted process $\{X_t\}_{t\in I}$ is called a *martingale* if

$$X_s = E(X_t|\mathcal{F}_s)\,, \text{ for all } s < t\,,\ s, t \in I\,. \tag{A.4}$$

If (A.4) holds with "$=$" replaced by $\leq$ (resp. by $\geq$) then the process $\{X_t\}$ is called *submartingale* (resp. *supermartingale*). If time is discrete, say $I = \mathbb{N}$, then a process $\{X_n\}_{n\in\mathbb{N}}$ is a martingale with respect to $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ if

$$X_n = E(N_{n+1}|\mathcal{F}_n)\,, \text{ for each } n\,. \tag{A.5}$$

A similar condition to (A.5) but with the equality replaced by an inequality is sufficient for $\{X_n\}$ to by a submartingale or a supermartingale. Using the definition of conditional expectation condition (A.5) can be written more explicitly:

$$\int_A X_n\, dP = \int_A X_{n+1}\, dP\,, \text{ for each } A \in \mathcal{F}_n\,,\ n \in \mathbb{N}\,.$$

Examples of martingales:

1) Let X be an integrable random variable and $\{\mathcal{F}_t\}_{t\in I}$ a filtration. Define

$$X_t \overset{\text{df}}{=} E(X|\mathcal{F}_t)\,, \quad \text{for each } t \in I\,.$$

Then $\{X_t\}_{t\in I}$ is a martingale with respect to $\{\mathcal{F}_t\}_{t\in I}$. 2) Let X_n, $n \in \mathbb{N}$, be a sequence of independent integrable random variables on probability space $(\Omega, \mathcal{F}, P)$ and put

$$\mathcal{F}_n \overset{\text{df}}{=} \sigma(X_1, \ldots, X_n)\,, \quad S_n \overset{\text{df}}{=} X_1 + \ldots + X_n\,.$$

If $EX_n = 0$, for each $n \in \mathbb{N}$, then $\{S_n\}$ is a martingale.
If $EX_n \geq 0$, for each $n \in \mathbb{N}$, then $\{S_n\}$ is a submartingale.
If $EX_n \leq 0$, for each $n \in \mathbb{N}$, then $\{S_n\}$ is a supermartingale.

3) Let $\{X_t\}_{t>0}$ be a Brownian motion starting from 0. Recall that process $\{X_t\}_{t\geq 0}$ has the following properties: $P(X_0 = 0) = 1$, the trajectories $t \mapsto X_t(\omega)$ are continuous, $EX_t = 0$ and $E(X_s X_t) = s \wedge t$. Moreover this process has independent increments random variable $X_t - X_s$, $0 \leq s < t$, has normal distribution with mean 0 and variance $t - s$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the natural filtration of $\{X_t\}_{t\geq 0}$. There is a several martingales connected with the Brownian motion $\{X_t\}$:

(1) $\{X_t\}_{t\geq 0}$

(2) $\{X_t^2 - t\}_{t\geq 0}$

(3) $\{\exp(i\theta X_t + \frac{1}{2}\theta^2 t)\}_{t\geq 0}$.

The fact that each of the above three processes is a martingale with respect to filtration $\{\mathcal{F}_t\}_{t\geq 0}$ follows easily from the properties of conditional expectation and Brownian motion. In fact, since $X_t - X_s$ is independent of $\mathcal{F}_s$,

(1) follows from: $E(X_t - X_s|\mathcal{F}_s) = 0$

(2) follows from: $t - s = E\left[(X_t - X_s)^2 |\mathcal{F}_s\right] = E\left(X_t^2|\mathcal{F}_s\right) - X_s^2$

(3) follows from: $E\left[e^{(i\theta(X_t - X_s)}|\mathcal{F}_s\right] = Ee^{i\theta(X_t - X_s)} = e^{-\frac{1}{2}\theta^2(t-s)}$.

It can be also shown that conditions (1) and (2) characterize Brownian motions. In fact, a continuous process $\{X_t\}_{t\geq 0}$, $X(0) = 0$, is a Brownian motion if and only if $\{X_t\}_{t\geq 0}$ and $\{X_t^2 - t\}_{t\geq 0}$ are martingales.

Exponential martingale (3) can be used to prove the law of the iterated logarithm:

$$P\left(\limsup_{t\to\infty} \frac{X_t}{\sqrt{2t\log\log t}} = 1\right) = 1\,.$$

Let $\{X_t\}_{t\in I}$ be a martingale with respect to $\{\mathcal{F}_t\}_{t\in I}$ and f a convex function on $\mathbb{R}$, then $\{f(X_t)\}_{t\in I}$ is a submartingale with respect to $\{\mathcal{F}_t\}_{t\in I}$. Similarly, if f is concave on $\mathbb{R}$, then $\{f(X_t)\}_{t\in I}$ is a supermartingale. This property is an immediate consequence of a generalized version of Jensen's inequality (Neveu 1965). It follows from this property that if $\{f(X_t)\}$ is a martingale, then $\{|X_t|)\}$, $\{X_t^2\}$, or $\{|X_t|^p\}$, $p > 1$, is a submartingale.

Particularly important are *uniformly integrable martingales (submartingales)*, i.e. such martingales (submartingales) for which $\{X_t\}$ is a family of uniformly integrable random variables. Recall here that *uniform integrability* of $\{X_t\}$ means that

$$\lim_{c\to\infty} \sup_{t\in I} \int_{\{|X_t|>c\}} |X_t|\, dP = 0\,, \quad \text{for each } c > 0\,.$$

It can be shown easily that the martingale $\{X_t\}_{t\in I}$ from Example 1 is uniformly integrable. This particular martingale is generated from a single random variable. As we shall see below each uniformly integrable martingale is in fact of such form. First, however, let us introduce a very useful criterion of uniform integrability.

If $\{X_t\}$ is a uniformly integrable family of functions then it is bounded in L^1, i.e. $\sup_t E|X_t| < \infty$. The converse is not true we have however the following:

Proposition A.2 *A family $\{X_t\}$ of random variables is uniformly integrable if and only if there exists a nonnegative, nondecreasing and convex function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ which satisfies:*

(i) $\displaystyle\lim_{x\to\infty} \frac{\phi(x)}{x} = \infty$

(ii) $\displaystyle\sup_{t\in I} E\phi(|X_t|) < \infty$.

Observe that the function $\phi(x) = x^p$ which is positive, increasing and convex on $\mathbb{R}_+$ satisfies, for $p > 1$, condition (i) of the above proposition. Therefore it follows from Proposition A.2 that each family $\{X_t\}_{t\in I}$ which satisfies the condition

$$\sup_{t\in I} E|X_t|^p < \infty$$

is uniformly integrable.

Proposition A.3 *Let $\{X_t\}_{t\in I}$ be a martingale with respect to $\{\mathcal{F}_t\}_{t\in I}$.*

(a) If $\{X_t\}$ is uniformly integrable (therefore bounded in L^1) and $t \nearrow t_0$, $t_0 \leq \infty$, then $\lim_{t\to t_0} X_t = X_{t_0}$ exists a.e., and in L^1, and

$$X_t = E(X_{t_0}|\mathcal{F}_t)\,, \quad \text{for each } t < t_0\,.$$

(b) If $\{X_t\}$ is bounded in L^p, $p > 1$, then $X_t \longrightarrow X_{t_0}$ in L^p.

An analog of proposition A.3 (a) remains true for submartingales provided $\{X_t^+\}_{t\in I}$ is uniformly integrable. Then we have

$$X_t \leq E(X_{t_0}|\mathcal{F}_t)\,, \quad \text{for each } t < t_0\,.$$

We also have the following

Proposition A.4 (Backward Martingale Convergence Theorem) *Let $\{X_t\}_{t\in I}$ be a martingale with respect to $\{\mathcal{F}_t\}_{t\in I}$ and let $\mathcal{F}_{s_0} = \bigcap_{t>s_0}$, $s_0 \geq -\infty$. If $\searrow s_0$ then the limit $\lim_{t\to s_0} = X_{s_0}$ exists a.e., and in L^1.*

If $\{X_n\}_{n\leq 0}$ is discrete time martingale and $\mathcal{F}_{-\infty} = \bigcap_{n\leq 0}\mathcal{F}_n$, then the family $\{X_n\}_{n\leq 0}$ is also uniformly integrable, and the limit random variable $X_{-\infty}$ from Proposition A.4 can be obtained by putting

$$X_{-\infty} = E(X_k|\mathcal{F}_{-\infty})\,,$$

where k is an arbitrary parameter. If $\mathcal{F}_{-\infty}$ is a trivial σ-algebra then

$$\lim_{n\to\infty} X_n = EX_k\,.$$

A.4 STOCHASTIC MEASURES INTEGRALS

Stochastic integral

$$\int_I f(t)\,dX_t\,,$$

where f is a deterministic function and $\{X_t\}$ is a stochastic process plays a crucial role in the construction of time operators. If the random event ω is fixed, then the trajectory $t \longmapsto X_t(\omega)$ is a real or complex valued function and the stochastic integral resembles the Riemann-Stieltjes (or Lebesgue-Stieltjes) integral. However, in spite of this similarity, there is a fundamental difference between these two kinds of integrals. Trajectories of a stochastic process are usually functions of unbounded variation, which means that the expression $\int_I f(t)\,dX_t(\omega)$ becomes meaningless in the classical sense.

We shall present below the idea behind the construction of the stochastic integral and show its properties. For the clarity of presentation we confine ourself to the simplest possible case, which nevertheless covers the main needs of this book. Namely, to the integral with respect to L^2-processes with orthogonal increments. For the purpose of the construction of stochastic integral we also introduce the concept of stochastic measure. Although the use of stochastic measures is not necessary, it helps to construct the stochastic integral in a very transparent way and also indicates possible generalizations.

Let $(\Omega,\mathcal{F},P)$ be a probability space, $L^2 = L^2(\Omega,\mathcal{F},P)$, and let $\{X_t\}_{t\in I}$ be an L^2-process with orthogonal increments. We initially assume that the index set is a finite interval, $I = [a,b]$. Let $\mathcal{B}_{[a,b]}$ denotes the Borel σ-algebra of subsets of $[a,b]$. Given $\{X_t\}_{t\in[a,b]}$ we define

$$M((s,t]) = X_t - X_s\,,\quad \text{for } a \leq s < t \leq b\,.$$

M as a function of intervals has the following properties:

1) $M(\emptyset) = 0$

2) $M\left(\bigcup_{n=1}^{\infty}\Delta_n\right)=\sum_{n=1}^{\infty}M(\Delta_n)$, if $\Delta_i\cap\Delta_j=\emptyset$ for $i\neq j$, where the series on the right hand side is convergent in L^2

3) $EM(\Delta_1)M(\Delta_2)=0$, for $\Delta_1\cap\Delta_2=\emptyset$.

It can be shown that the set function M extends on the whole Borel σ-algebra $\mathcal{B}_{[a,b]}$ preserving properties 1), 2) and 3). Condition 2) is called the *countable additivity* and condition 3), which follows from the assumption that $\{X_t\}$ has orthogonal increments, is called the *orthogonal scattering*. The measure M is called the *stochastic measure*.

Given a stochastic measure M one can also define an associated set function m on $\mathcal{B}_{[a,b]}$ by putting

$$m(\Delta)=E|M(\Delta)|^2\,.$$

It follows from properties 1)–3) m that m is a countable additive nonnegative measure on $\mathcal{B}_{[a,b]}$. It is called the *control measure* of M (or of the stochastic process $\{X_t\}$).

One of the basic properties of a control measure is that $m(\Delta)=0$ if and only if $M(\Delta)=0$. A control measure of a stochastic process plays an important role in the description of the class of stochastically integrable functions.

Example. If $\{X_t\}$ is a Brownian motion then $E|X_t-X_s|^2=t-s$, which implies that on intervals we have

$$m((s,t])=t-s\,.$$

Therefore the control measure determined by the Brownian motion is the Lebesgue measure.

Denote be $\mathcal{L}^2$ the Hilbert space $L^2([a,b],\mathcal{B}_{[a,b]},m)$ and by $\mathcal{L}_0^2$ the space of all simple functions in $\mathcal{L}^2$. For $f\in\mathcal{L}_0^2$,

$$f(t)=\sum_{k=1}^{n}a_k\mathbb{1}_{\Delta_k}(t)\,,\quad \Delta_k\in\mathcal{B}_{[a,b]}\,,$$

we define the stochastic integral with respect to M by putting

$$\int_a^b f(t)\,M(dt)=\sum_{k=1}^{n}a_kM(\Delta_k)\,.$$

We shall write shortly $\int f\,dM$. The integral just defined has the following properties:

1) $\int(af+bg)\,dM=a\int f\,dM+b\int g\,dM$

2) $E\left(\int f\,dM\overline{\int g\,dM}\right)=\int f(t)\overline{g(t)}\,m(dt)$.

Property 1) means that the transformation

$$\mathcal{L}_0^2\ni f\longmapsto\int f\,dM\in L^2 \tag{A.6}$$

is a linear operator. By Property 2) this transformation is also an isometry

$$\left\| \int f \, dM \right\|_{L^2} = \|f\|_{\mathcal{L}^2}$$

from a dense subspace of the Hilbert space $\mathcal{L}^2$ into L^2. Therefore transformation (A.6) extends to an isometry on the whole $\mathcal{L}^2$. Thus we can add one more property:

3) $f_n \longmapsto f$ in $\mathcal{L}^2$ $\Leftrightarrow$ $\int f_n \, dM \longmapsto \int f_n \, dM$ in L^2 .

A.5 PREDICTION, FILTERING AND SMOOTHING

Let us begin from the following problem: Find the best estimation $\widehat{X}$ of the values of the random variable X observing some stochastic process $\{Z_t\}_{t\in I}$. This means, we have to find a function $f(\{Z_t\}_{t\in I})$ of the set of random variables $\{Z_t\}$ for which

$$X \approx \widehat{X} = f(\{Z_t\}_{t\in I})$$

with the smallest possible error.

We distinguish here the following problems:

I) *Prediction* – estimate the value of the stochastic process in moment t^*, knowing its values for times $t < t^*$.

II) *Filtering* – in the moments $t \in T' \subset T$ we observe the process

$$Z_t = X_t + \eta_t$$

("signal" + "noise"). The task is to separate signal from the noise, i.e. for given $t^* \in T$ find the best approximation

$$X(t^*) \approx \widehat{X} = f(\{Z_t\}_{t\in T'}) \, .$$

III) *Smoothing* – the information about X_t need *not* become available at t and measurements derived later than t can be used for estimation of X_t.

The term filtering means also the recovery at time t of some information about the signal X_t using measurements until time t, i.e. Z_s , $s \leq t$ *not* later.

An illustrative example of smoothing is the way how the human brain tackles the problem of reading hastily written handwriting - words are tackled sequentially; reaching a word difficult to interpret we go several words after (and before) to deduce it.

The meaning of the "best approximation" is the following: We shall look for a function f for which the *error*

$$\delta \overset{\text{df}}{=} (E \mid X - f(\{Z_t\}_{t\in I}) \mid^2)$$

is minimal.

The meaning of $f(\{Z_t\}_{t\in I})$ is the following: If I is a finite set $f(\{Z_t\}_{t\in I})$ is a Borel measurable function of the arguments $Z_t\,,\ t\in I$. If I is infinite then this symbol denotes a random variable measurable with respect to the σ- algebra.

$$\mathcal{F}\overset{\mathrm{df}}{=}\sigma\{Z_t\,,\ t\in I\}\,.$$

We shall assume in the sequel that f has finite second moment.

Proposition A.5 *The best approximation of a random variable X with the finite second order by using $\mathcal{F}=\sigma(Z_t\,, t\in I)$ is given by*

$$\widehat{X}=E(X\mid\mathcal{F})$$

Proof. Put $Y=E(X\mid\mathcal{F})$. Then

$$\delta^2=E(X-X)^2=E(X-Y)^2+2E(X-Y)(Y-\widehat{X})+E(Y-\widehat{X})^2\,.$$

Since $Y-\widehat{X}$ is $\mathcal{F}$-measurable,

$$\begin{aligned}E(X-Y)(Y-\widehat{X}) &= E(E\{(X-Y)(Y-\widehat{X})\}\mid\mathcal{F})\\ &= E((Y-\widehat{X})E\{(X-Y)\mid\mathcal{F}\})\\ &= E((Y-\widehat{X})E\{X_E(X\mid\mathcal{F})\})\,.\end{aligned}$$

But $E(X-E(X\mid\mathcal{F}))=E(X\mid\mathcal{F})-E(X\mid\mathcal{F})=0$. Thus $E(X-Y)(Y-\widehat{X})=0$. Therefore $\delta^2=E(X-Y)^2+E(Y-\widehat{X})^2$ and δ^2 is minimal when $\widehat{X}=Y=E(X\mid\mathcal{F})$.

Remark The above proposition is not very useful from the practical point of view. It is sometimes easier to solve the problem of minimum if we restrict ourself to a smaller class of admissible solutions than all measurable functions. One of such possibilities is the following:

Let $L^2\{Z_t\}_{t\in I}$ be the closure in L^2 of the linear space spanned by $Z_t\,,\ t\in I$, and constant functions. The best linear approximation $\tilde{X}$ of X is this element of $L^2\{Z_t\}$ for which

$$\delta^2=E\mid\tilde{X}-X\mid^2\leq E\mid X'-X\mid^2,\ \text{for all } X'\in L^2\{Z_t\}.$$

Such $\tilde{X}$ always exists. $\tilde{X}$ is the orthogonal projection of X onto $L^2\{Z_t\}$. It is also unique. A practical way to find $\tilde{X}$ is to solve the system of equations

$$(X-\tilde{X},X')=0, X'\in L^2\{Z_t\}\,.$$

Linear estimations are *not* always acceptable, however for gaussian random variable we have:

Proposition A.6 *If the family $\{X,Z_t\,,\ t\in I\}$ is gaussian with $EX=EZ_t=0$ then the best approximation $\widehat{X}$ of X with respect to $\sigma\{Z_t\}$ coincides with the best linear approximation $\widetilde{X}$.*

Proof
$$(X - \widetilde{X}, Z_t) = 0 \Longleftrightarrow E(X - \widetilde{X})\bar{Z}_t = 0 \Rightarrow X - \widetilde{X}$$
and Z_t are uncorrelated and since both are gaussian, they are independent.

Note that $\widetilde{X} \in L^2\{Z_t\}$ which implies that $\widetilde{X}$ is gaussian and, consequently, $X - \widetilde{X}$ is also gaussian. Therefore $X - \widehat{X}$ and $\sigma\{Z_t\}$ are independent and

$$\begin{aligned}\widetilde{X} = E(X \mid \mathcal{F}) &= E(X - \widetilde{X} + \mid \mathcal{F}) \\ &= E(X - \widetilde{X} \mid \mathcal{F}) + E(\widetilde{X} \mid \mathcal{F}) \\ &= E(X - \widetilde{X}) + \widetilde{X} = \widetilde{X}\,.\end{aligned}$$

An easily solvable example concern the prediction on an interval $[a, b]$ of mean square continuous stochastic processes (in particular stationary stochastic processes). Such processes can be represented as $Z(t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\varphi_n(t)Z_n$, where $\lambda_k, \varphi_k(t)$ are respectively eigenvalues and eigenvectors of the correlation function

$$R(s,t) = E(Z_s - m)(Z_t - m) \;\; (\text{where } m = EX_t),$$

i.e.
$$\lambda_k\varphi_k(t) = \int_a^b R(s,t)\varphi_k(s)ds\,.$$
In such case $\widetilde{X} = \sum_{n=1}^{\infty} c_n Z_n$, where c_n can be expressed through the correlation function R.

A straight forward generalization of above examples leads to the problem of prediction of stochastic processes given in the form of stochastic integral

$$\int f(u,t)\, dZ_u\,.$$

A.6 KARHUNAN-LOEVE EXPANSION

Let $\{X_t\}_{t\in[a,b]}$ be a mean-square continuous stochastic process with the covariance function $R(s,t) = E(X_t - EX_t)\overline{(X_s - EX_s)}$. Then X_t can be expanded into the series
$$X(\omega, t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\varphi_n(t)Y_n(\omega)$$
which is convergent in L^2_Ω for each $t \in [a, b]$. Here Y_n is an orthonormal (in L^2_Ω) sequence of random variables, λ_n are eigenvalues and $\varphi_n(t)$ are eigenfunction of the covariance function, i.e.

$$\int_a^b R(s,t)\varphi_n(s)\, ds = \lambda_n\varphi_n(t)\;,\; a \le t \le b\,.$$

Remark The above theorem is a consequence of the following fact:

1° $\{X_t\}_{t\in[a,b]}$ is mean-square continuous $\Longrightarrow$ $R(s,t)$ is a continuous function on the square $[a,b]^2$.
2° $R(s,t)$ is a kernel of the following integral operator

$$L^2_{[a,b]} \ni \varphi(t) \longmapsto \int_a^b R(s,t)\varphi(s)\,ds$$

3° According to the Mercer theorem $R(s,t)$ has the expansion

$$R(s,t) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(s)\overline{\varphi_n(t)}\,,$$

where the series converges uniformly on $[a,b]\times[a,b]$, and $\lambda_n, \varphi_n(t)$ are the eigenfunctions of above kernel operator
4°

$$Y_n \overset{\text{df}}{=} \int_a^b X(t)\varphi_n(t)\,dt$$

this is a Riemann-kind integral of an L^2-valued continuous function defined on $[a,b]$.

Example (*of the Karhunan-Loeve expansion*)
Let $X(t)$ be a Brownian motion on $[0,1]$ (here $r(s,t)=s\wedge t$) then

$$X(t) = tY_0 + \sqrt{2}\sum_{n=1}^{\infty} Y_n \frac{\sin \pi n t}{n\pi}\,,$$

where $Y_0, Y_1, \ldots$ i.i.d. random variables $N(0,1)$. This series converges also with probability 1.

Return to the problem of estimation (prediction) of a mean-square continuous stochastic process. We have

$$Z(t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n}\varphi_n(t) Z_n$$

Obviously $\{Z_n\}_{n=1,2,\ldots}$ form an ON basis in $L^2(\{Z_t\}_{t\in[a,b]})$. Therefore the best linear approximation must be of the form

$$\widetilde{X} = \sum_{n=1}^{\infty} c_n Z_n\,.$$

We want to compute c_n. Recall first that $\widetilde{X}$ realized the minimum

$$\| X - \widetilde{X} \| = \min \| X - X' \|$$

when $\widetilde{X}$ is the projection of X onto $L^2\{Z(t)\}$. This is equivalent to

$$\forall t : (X - \widetilde{X}, Z(t)) = 0$$

which is in turn equivalent to

$$\forall t : (X, Z(t)) = (\widetilde{X}, Z(t))\,.$$

We have

$$c_n = \left(\sum_{n=1}^{\infty} c_n Z_n, Z_n\right) = (\widetilde{X}, Z_n)\,.$$

On the other hand, the orthonormality of the family $\{\varphi_n(t)\}$ implies

$$\sqrt{\lambda_n} Z_n = \int_a^b \left(\sum_{n=1}^{\infty} \lambda_n \varphi_n(t) Z_n\right) \overline{\varphi_n(t)}\, dt = \int_a^b Z(t)\overline{\varphi_n(t)}\, dt\,.$$

Therefore

$$\begin{aligned} c_n = (\widetilde{X}, Z_n) = \frac{1}{\sqrt{\lambda_n}} \left(\widetilde{X}, \int_a^b Z(t)\overline{\varphi_n(t)}\, dt\right) \quad &= \frac{1}{\sqrt{\lambda_n}} E\widetilde{X} \int_a^b \overline{Z(t)\overline{\varphi_(t)}\, dt} \\ &= \frac{1}{\sqrt{\lambda_n}} \int_a^b (E\widetilde{X}\overline{Z(t)})\varphi_n(t)\, dt \\ &= \frac{1}{\sqrt{\lambda_n}} \int_a^b R_{XZ}(t)\varphi_n(t)\, dt = c_n\,, \end{aligned}$$

where $R_{XZ}(t) = (X, Z(t))$.

Appendix B
Operators on Hilbert and Banach spaces.

Let $\mathcal{H}$ be a vector space over the field of real or complex numbers. A *scalar product* on $\mathcal{H}$ is a mapping from the Cartesian product $\mathcal{H} \times \mathcal{H}$ into the field of scalars such that it assigns to each pair $(\psi, \varphi) \in \mathcal{H} \times \mathcal{H}$ a number $\langle \psi | \varphi \rangle$ with the following properties:

(i) If $\psi, \varphi, \phi \in \mathcal{H}$, and $\alpha, \beta \in \mathbb{C}$ then:

$$\langle \psi | \alpha\, \varphi + \beta\, \phi \rangle = \alpha\, \langle \psi | \varphi \rangle + \beta\, \langle \psi | \phi \rangle$$

(ii) For any $\psi, \varphi \in \mathcal{H}$,

$$\langle \psi | \varphi \rangle = \overline{\langle \varphi | \psi \rangle}\,,$$

where the bar denotes complex conjugation.

(iv) For any $\psi \in \mathcal{H}$,

$$\langle \psi | \psi \rangle \geq 0$$

and

$$\langle \psi | \psi \rangle = 0 \Rightarrow \psi = 0\,.$$

Let us introduce some notions associated with the concept of scalar product.

We say that two vectors $\psi, \varphi \in \mathcal{H}$ are *orthogonal* if

$$\langle \psi | \varphi \rangle = 0 \,.$$

Let $\psi \in \mathcal{H}$, where $\mathcal{H}$ is an arbitrary vector space. A *norm* on $\mathcal{H}$ is a mapping from $\mathcal{H}$ into the set of non negative real numbers assigning to every $\psi \in \mathcal{H}$ a positive number $\|\psi\|$ fulfilling the following conditions:

1. $\|\psi\| = 0$ if and only if $\psi = 0$.
2. For any $\psi \in \mathcal{H}$ and $\lambda \in \mathbb{C}$, $\|\lambda\psi\| = |\lambda| \, \|\psi\|$.
3. For any $\psi, \varphi \in \mathcal{H}$ the *triangle inequality* is satisfied, i.e.,

$$\|\psi + \varphi\| \leq \|\psi\| + \|\varphi\|$$

A vector space $\mathcal{H}$ with a norm is called the *normed space*.

A norm determines on $\mathcal{H}$ a distance on $\mathcal{H}$ defined as

$$d(\psi, \varphi) = \sqrt{\|\psi - \varphi\|}$$

Let $\mathcal{H}$ be a normed space. We say that the sequence $\{\psi_n\}$ of vectors in $\mathcal{H}$ *converges* to $\psi \in \mathcal{H}$ if for any positive number $\varepsilon > 0$, there exists a natural number N such that if $n > N$ then

$$\|\psi_n - \psi\| < \varepsilon$$

A *Cauchy sequence* in $\mathcal{H}$ is a sequence of vectors in $\mathcal{H}$, $\{\psi_n\}$, such that for any positive number $\varepsilon > 0$, there exists a natural number N with the property that if $n, m > N$ then

$$\|\psi_n - \psi_m\| < \varepsilon$$

Any convergent sequence is a Cauchy sequence, but the converse is not true in general. If a normed space $\mathcal{H}$ has the property that each of its Cauchy sequences has a limit on $\mathcal{H}$, the space is called *complete*. A complete normed space is called the *Banach space*.

A vector space with a scalar product is a normed space. Its norm is defined as:

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle} \tag{B.1}$$

A property of a particular interest is the *Cauchy-Schwarz inequality*. Let $\psi, \varphi \in \mathcal{H}$, were $\mathcal{H}$ is a vector space with scalar product then

$$|\langle \psi | \varphi \rangle| \leq \|\psi\| \, \|\varphi\|$$

A *Hilbert space* is a vector space with a scalar product that is complete under the norm (**??**).

Examples of Hilbert spaces:

1. Let $\mathcal{H} = \mathbb{C}^n$, the n dimensional vector space of sequences of n complex numbers, with the addition and multiplication by complex numbers defined as:

$$\begin{aligned}(\alpha_1, \alpha_2, \ldots, \alpha_n) + (\beta_1, \beta_2, \ldots, \beta_n) &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n) \\ \lambda\,(\alpha_1, \alpha_2, \ldots, \alpha_n) &= (\lambda\,\alpha_1, \lambda\,\alpha_2, \ldots, \lambda\,\alpha_n) \qquad \text{(B.2)}\end{aligned}$$

where $\lambda, \alpha_i, \beta_i$ ($i = 1, 2, \ldots, n$) are complex numbers. For $\psi \equiv (\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $\varphi \equiv (\beta_1, \beta_2, \ldots, \beta_n)$ we define the scalar product of these two vectors as:

$$\langle\psi|\varphi\rangle = \sum_{i=1}^{n} \overline{\alpha_i}\,\beta_i$$

This scalar product produces the following norm:

$$\|\psi\| = \sqrt{\sum_{i=1}^{n} |\alpha_i|^2}$$

Then, $\mathbb{C}^n$ is finite dimensional normed space. Since all finite dimensional normed spaces are complete, we conclude that $\mathbb{C}^n$ is a Hilbert space.

2. Let us consider the set of all Lebesgue measurable functions from $\mathbb{R}^n$ to $\mathbb{C}$ in which we identify those functions that differ on a set of zero *Lebesgue measure* only. Consider the class $L^2_{\mathbb{R}^n}$ of all functions that are square integrable in the Lebesgue sense

$$\int_{\mathbb{R}^n} |f(\mathbf{x})|^2\, d^n\mathbf{x} < \infty$$

in the *Lebesgue sense*. If f and g are square integrable then it follows from the Schwartz inequality

$$\left\{\int_{\mathbb{R}^n} |f(\mathbf{x}) + g(\mathbf{x})|^2\, d^n\mathbf{x}\right\}^{1/2} \leq \left\{\int_{\mathbb{R}^n} |f(\mathbf{x})|^2\, d^n\mathbf{x}\right\}^{1/2} + \left\{\int_{\mathbb{R}^n}^{\infty} |g(\mathbf{x})|^2\, d^n\mathbf{x}\right\}^{1/2}$$

that the set of square integrable functions form a vector space.

For $f, g \in L^2_{\mathbb{R}^n}$ let us define the scalar product:

$$\langle f|g\rangle = \int_{\mathbb{R}^n} \overline{f(\mathbf{x})}\, g(\mathbf{x})\, d^n\mathbf{x}\,,$$

which is well defined by the Hölder inequality:

$$\left|\int_{\mathbb{R}^n} \overline{f(\mathbf{x})}\, g(\mathbf{x})\, d^n\mathbf{x}\right| \leq \left\{\int_{\mathbb{R}^n} |f(\mathbf{x})|^2\, d^n\mathbf{x}\right\}^{1/2} \cdot \left\{\int_{\mathbb{R}^n} |g(\mathbf{x})|^2\, d^n\mathbf{x}\right\}^{1/2}.$$

The norm determined by this scalar product is:

$$\|f\| = \sqrt{\int_{\mathbb{R}^n} |f(\mathbf{x})|^2\, d^n\mathbf{x}}$$

With this scalar product and norm $L^2_{\mathbb{R}^n}$ is a Hilbert space. If $n = 1$, we have $L^2_{\mathbb{R}}$ as a particular case

Analogously, we can consider the set of square integrable functions from a Borel subset I of $\mathbb{R}^n$ into $\mathbb{C}$, also identifying these functions which differ on a subset of I of zero Lebesgue measure. The definition of the space L^2_I, its scalar product and its norm goes in complete analogy with $L^2_{\mathbb{R}^n}$. L^2_I is a Hilbert space.

3. A straightforward generalization of $L^2_{\mathbb{R}^n}$ is the space of square integrable functions over an arbitrary measure space $(\mathcal{X}, \Sigma, \mu)$, where $\mathcal{X}$ is a nonempty set, Σ a σ-algebra of subsets of $\mathcal{X}$ and μ a measure on Σ. We identify those functions which differ on a set of zero μ-measure only. This space is denoted by $L^2(\mathcal{X}, \Sigma, d\mu)$. As in the previous case $L^2(\mathcal{X}, \Sigma, d\mu)$ is a Hilbert space.

4. Let us consider the set of all countably infinite sequences of complex numbers $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ (in short $\{\alpha_n\}$) such that

$$\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$$

This set is denoted by l^2. Let $\{\alpha_n\}$ and $\{\beta_n\}$ be two elements of l^2. With the following definition for the addition and multiplication by complex numbers

$$\{\alpha_n\} + \{\beta_n\} = \{\alpha_n + \beta_n\} \qquad ; \qquad \lambda\{\alpha_n\} = \{\lambda a_n\}$$

l^2 is a vector space, in which a scalar product is defined as:

$$\langle \{\alpha_n\} | \{\beta_n\} \rangle = \sum_{n=1}^{\infty} \overline{\alpha_n}\, \beta_n$$

Because of (**??**) this sum converges, so that the scalar product is well defined. It produces the following norm:

$$\|\{\alpha_n\}\| = \sum_{n=1}^{\infty} \sqrt{|\alpha_n|^2} \tag{B.3}$$

The space l^2 is complete under this norm and therefore, it is a Hilbert space.

Let $\mathcal{S}$ be a subset of $\mathcal{H}$. We say that $\mathcal{S}$ is an *orthonormal system* if

1. $\|\psi\| = 1, \forall \psi \in \mathcal{S}$.
2. $\langle \psi | \phi \rangle = 0, \forall \psi, \phi \in \mathcal{S}$.

A Hilbert space $\mathcal{H}$ is *separable* if there exists a countable subset $\mathcal{S} \subset \mathcal{H}$ such that for any $\varepsilon > 0$ and any $\psi \in \mathcal{H}$, it exists a $\varphi \in \mathcal{S}$ such that $d(\psi, \varphi) = \|\psi - \varphi\| < \varepsilon$.

Proposition B.1 *(Bessel inequalities). Let $\mathcal{S}$ be an orthonormal system in $\mathcal{H}$, $\{\psi_1, \psi_2, \ldots\}$ a sequence (finite or infinite) of vectors in $\mathcal{S}$ and φ, ϕ two arbitrary vectors of $\mathcal{H}$. Then*

$$\sum_{n=0}^{N} |\langle\varphi|\psi_n\rangle|^2 \leq \|\varphi\|^2 \tag{B.4}$$

$$\sum_{n=0}^{N} |\langle\varphi|\psi_n\rangle\langle\psi_n|\phi\rangle| \leq \|\varphi\|\,\|\phi\| \tag{B.5}$$

where N stands for the number of vectors in the sequence. If the sequence is infinite $N = \infty$.

Let $\mathcal{S}$ be an orthonormal system on $\mathcal{H}$. $\mathcal{S}$ is called the *complete orthonormal system* if there is no another orthonormal system $B \subset \mathcal{H}$ such that $A \subset B$ and $A \neq B$ Equivalently, $\mathcal{S}$ is a complete orthonormal system if and only any vector ψ orthogonal to all the vectors of $\mathcal{S}$ is necessarily the zero vector.

The concept complete orthonormal system in Hilbert spaces generalizes the idea of basis for finite dimensional spaces, in the sense that every vector of the Hilbert space can be written in terms of vectors on a complete orthonormal system. Namely, if $\mathcal{H}$ is a Hilbert space, then:

1. There exist a complete orthonormal system in$\mathcal{H}$

2. Any orthonormal system $\mathcal{S}$ can be completed adding additional vectors to $\mathcal{S}$ in such a way that the extended system $\tilde{\S}$ is a complete and orthonormal.

3. A Hilbert space $\mathcal{H}$ is separable if and only if every complete orthonormal system in $\mathcal{H}$ is countable.

In practically all applications of Hilbert spaces in Physics, only separable Hilbert spaces are used. Therefore, unless stated otherwise, the word Hilbert space will mean separable Hilbert space

Proposition B.2 *Let $\psi_1, \psi_2, \ldots, \psi_n \ldots$ be an orthonormal system, either complete or not, in the Hilbert space $\mathcal{H}$. Let $\alpha_1, \alpha_n, \ldots, \alpha_n, \ldots$ be a sequence of scalars. Then, the series*

$$\sum_{n=0}^{N} \alpha_n\, \psi_n$$

converges if and only if

$$\sum_{n=0}^{\infty} |\alpha_n|^2 < \infty.$$

This proposition has the following consequence:

Corollary B.1 *Let $\mathcal{H}$ be a Hilbert space either finite or infinite dimensional and let $\psi_1, \psi_2, \ldots, \psi_n \ldots$ be a complete orthonormal system. If $\psi \in \mathcal{H}$ then*

$$\psi = \sum_{n=0}^{N} \alpha_n\, \psi_n$$

where $\alpha_n = \langle \psi_n | \psi \rangle$. *In addition,*

$$\|\psi\|^2 = \sum_{n=0}^{N} |\alpha_n|^2 = \sum_{n=0}^{N} |\langle \psi_n | \psi \rangle|^2$$

where N *is either finite or* ∞.

The last identity in the above proposition is called the *Parseval identity*. Note that the expansion for a vector in $\mathcal{H}$ in terms of a complete orthonormal system is unique and depends only on the vector and the complete orthonormal system.

Examples of complete orthonormal systems.

i) Let $\mathcal{H} = \mathbb{C}^n$, the space of sequences of n complex numbers. Since $\mathbb{C}^n$ is a finite dimensional Hilbert space, any orthonormal basis in $\mathbb{C}^n$ is also a complete orthonormal system. The simplest one is the canonical basis.

$$(1, 0, 0, \dots, 0); \quad (0, 1, 0, \dots, 0); \; \dots \; (0, 0, 0, \dots, 1)$$

ii) Take $\mathcal{H} \equiv l^2$ and define the following sequence of vectors in l^2:

$$\begin{aligned} \psi_1 &= (1, 0, 0, \dots, 0, \dots) \\ \psi_2 &= (0, 1, 0, \dots, 0, \dots) \\ &\quad \dots\dots\dots\dots \\ \psi_n &= (0, 0, 0, \dots, 1, 0, \dots) \qquad \text{(B.6)} \\ &\quad \dots\dots\dots\dots \end{aligned}$$

If $\xi = (\xi_1, \xi_2, \dots, \xi_k, \dots) \in l^2$, then we can write:

$$\xi = \sum_{k=1}^{\infty} \xi_k \, \psi_k$$

iii) Take $L^2[-\pi, \pi]$, the Hilbert space of square integrable complex functions on the real interval $[-\pi, \pi]$. A complete orthonormal system in $L^2[-\pi, \pi]$ is given by:

$$\psi_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \quad n = \dots, -1, 0, 1, \dots, \quad t \in [-\pi, \pi].$$

iv) Take now $L^2[0, 2\pi]$. In this case

$$\psi_0(t) = \frac{1}{\sqrt{2\pi}}, \qquad \psi_1(t) = \frac{\cos t}{\sqrt{\pi}}, \qquad \psi_2(t) = \frac{\sin t}{\sqrt{\pi}}, \qquad \psi_4(t) = \frac{\cos 2t}{\sqrt{\pi}}, \quad \dots$$

is a complete orthonormal system.

v) The Hermite functions are the following

$$\varphi_n(x) := A_n\, H_n(x)\, e^{-x^2/2}$$

where $H_n(x)$ are the polynomials of n order defined as:

$$H_n(x) := e^{x^2/2} \left[\frac{1}{2^n\, n!\, \pi^{1/2}}\right]^{1/2} \left\{x - \frac{d}{dx}\right\}^n e^{-x^2/2}$$

and A_n are normalization constants, so that

$$\int_{-\infty}^{\infty} |\varphi_n(x)|^2\, dx = 1$$

The Hermite functions form a complete orthonormal system in $L^2_{\rm I\!R}$.

Two normed spaces X and Y are called *isometric* if there is a linear map F from X to Y such that

$$\|\psi\|_X = \|F(\psi)\|_Y \;, \qquad \forall \psi \in \; X \;,$$

where $\|\cdot\|_X$ and $\|\cdot\|_Y$ denote the norm of X and Y respectively. X and Y are called *isometrically isomorphic* if the map F is in addition *onto* and *one-to-one*.

If two normed spaces are isometrically isomorphic then they have the same properties from both algebraic and topological point of view. In particular, if $\mathcal{H}$ and $\mathcal{G}$ are two finite dimensional Hilbert spaces of the same dimension N then $\mathcal{H}$ and $\mathcal{G}$ are isometrically isomorphic. We also have

Proposition B.3 *Any infinite dimensional separable Hilbert space is isometrically isomorphic to l^2.*

Let $\mathcal{H}$ be a Hilbert space and F a continuous linear mapping from $\mathcal{H}$ into the field of scalars. We call F the *linear functional.* This means that if $\{\psi_n\}$ is a convergent sequence in $\mathcal{H}$ and $\psi \in \mathcal{H}$ is its limit then

$$F(\psi_n) \longmapsto F(\psi)\,.$$

Equivalently, a linear linear functional is a mapping F from the Hilbert space $\mathcal{H}$ into the field of scalars, which is bounded, i.e. if there exists a positive number $K > 0$ such that

$$|F(\psi)| \leq K\, \|\psi\|\,,$$

for all $\psi \in \mathcal{H}$.

An important example of a linear functional on $\mathcal{H}$ is the following. Let ψ be a fixed vector in $\mathcal{H}$ and define

$$F_\psi(\varphi) := \langle\psi|\varphi\rangle \;, \qquad \forall \varphi \in \mathcal{H}\,.$$

Obviously, F_ψ is a linear mapping from $\mathcal{H}$ to $\mathbb{C}$. To see that it is also continuous we use the Schwarz inequality:

$$|F_\psi(\varphi)| \leq |\langle\psi|\varphi\rangle| \leq \|\psi\| \, \|\varphi\| \, , \qquad \forall\varphi \in \mathcal{H} \, .$$

Let us denote by $\mathcal{H}^\times$ the set of functionals on $\mathcal{H}$. We endow $\mathcal{H}^\times$ with a vector space structure defining the sum and multiplication with scalars as follows. If $F, G \in \mathcal{H}^\times$ and $\alpha, \beta \in \mathbb{C}$,

$$(\alpha\, F + \beta\, G)\,(\psi) := \alpha\, F(\psi) + \beta\, G(\psi) \ , \qquad \forall\psi \in \mathcal{H} \, .$$

Furthermore, we can define on $\mathcal{H}^\times$ a norm as follows. If $F \in \mathcal{H}^\times$,

$$\|F\| := \inf\{K \ \ : \ \ |F(\psi)| \leq K\,\|\psi\| \ , \ \ \forall\psi \in \mathcal{H}\}$$

The norm on can be equivalently expressed as follows

$$\|F\| = \sup_{\|\psi\|\leq 1} |F(\psi)|$$

We have endowed $\mathcal{H}^\times$ with the structure of normed space. This is a general property of normed spaces, i.e., the dual $X^\times$ of a normed space X is also a normed space. Furthermore, $X^\times$ is always complete. But the dual of a Hilbert space is also a Hilbert space. This is the result of the following:

Theorem B.1 (Riesz). *Let $\mathcal{H}$ be a Hilbert space and F an arbitrary functional on $\mathcal{H}$. Then, there exists a unique vector $\psi \in \mathcal{H}$ such that*

$$F(\varphi) = \langle\psi|\varphi\rangle \, , \qquad \forall\varphi \in \mathcal{H} \, .$$

The Riesz theorem has the following consequence: The dual $\mathcal{H}^\times$ of a Hilbert space $\mathcal{H}$ is a Hilbert space which is isometrically isomorphic to $\mathcal{H}$. In particular, if $\mathcal{H}$ has finite dimension n, so has $\mathcal{H}^\times$ and $\mathcal{H}$ is infinite dimensional, so is $\mathcal{H}^\times$.

Operators are linear mappings on vector spaces. In this section we restrict ourself to Hilbert spaces. Two kind of operators will be considered – bounded and unbounded operators. The class of bounded operators contains, in particular, projections and evolution operators. Among examples of unbounded operators are Hamiltonias or the generators of groups of evolutions. Many of the properties of bounded operators are is straightforward generalization a generalization of the properties of matrices on the Euclidean space. Unbounded operators present new and important features that distinguish them from bounded operators. We present the relevant properties of both classes below begin with the study of bounded operators.

Let $\mathcal{H}$ be a Hilbert space. A *bounded operator* on $\mathcal{H}$ is a linear mapping A from $\mathcal{H}$ into $\mathcal{H}$ such that for all vectors $\psi \in \mathcal{H}$, there exists a positive constant $K > 0$ with

$$\|A\psi\| \leq K\,\|\psi\| \, .$$

Similarly to linear functionals bounded operators are continuous and conversely. Recall that an operator A on $\mathcal{H}$ is continuous if and only if is is continuous at the origin. This means that if $\psi_n \longmapsto 0$, then, $A\psi_n \longmapsto 0$.

Assume now that A is a continuous linear mapping from a dense subspace $\mathcal{D}$ of the Hilbert space $\mathcal{H}$ to $\mathcal{H}$. Then, there exists a unique bounded operator on $\mathcal{H}$ such that it coincides with A on $\mathcal{D}$. This operator, also denoted by A, is called the extension of A to $\mathcal{H}$.

Denote by $L(\mathcal{H})$ the class of all bounded operators on $\mathcal{H}$. The sum of operators and the multiplication by scalars is defined in the usual way:

$$(\alpha\, A + \beta\, B)\,(\psi) := \alpha\, A\psi + \beta\, B\psi\ , \quad \psi \in \mathcal{H}\,, \quad \alpha, \beta \in \mathbb{C}$$

Thus $L(\mathcal{H})$ is a linear space: if $A, B \in L(\mathcal{H})$, $\alpha\, A + \beta\, B \in L(\mathcal{H})$. Moreover defining the product of two operators A and B as their composition AB we see that $L(\mathcal{H})$ is an algebra.

The *norm* of a bounded operator $A \in L(X)$ is defined as

$$\|A\| = \inf\{K \ : \ \|A\psi\| \le K\,\|\psi\|,\ \forall\psi \in \mathcal{H}\}\,. \tag{B.7}$$

An equivalent definition of the operator norm is

$$\|A\| = \sup_{\|\psi\|\le 1} \|A(\psi)\|\,. \tag{B.8}$$

As an immediate consequence of (B.8) we obtain

$$\|A\psi\| \le \|A\|\,\|\psi\|\ , \qquad \forall\psi \in \mathcal{H}$$

so that $\|A\|$ is one of the positive numbers K that fulfill (B.7).

Proposition B.4 *The space $L(\mathcal{H})$ with the norm (B.7) is a complete normed space, i.e. a Banach space (or a Banach algebra, if the multiplication is taken into account).*

Since $L(\mathcal{H})$ is a Banach space, we have a notion of convergence in $L(\mathcal{H})$. For many applications this sense of convergence is not sufficient and we need also a weaker sense of convergence of bounded operators. Below, we list the three most important notions of convergence on $L(\mathcal{H})$:

A sequence $\{A_n\} \subset L(\mathcal{H})$ converges *uniformly* to $A \in L(\mathcal{H})$ if $\{A_n\}$ converges to A in the operator norm of $L(\mathcal{H})$, i.e. if $\lim_{n\to\infty} \|A_n - A\| < \varepsilon$.

A sequence $\{A_n\} \subset L(\mathcal{H})$ converges in the *strong sense* to $A \in L(\mathcal{H})$ if for each $\psi \in \mathcal{H}$, $A_n\psi \longmapsto A\psi$ in $\mathcal{H}$. Strong convergence is a convergence in the norm of the Hilbert space $\mathcal{H}$.

A sequence $\{A_n\} \subset L(\mathcal{H})$ converges in the *weak sense* to $A \in L(\mathcal{H})$ if for each pair $\psi, \varphi \in \mathcal{H}$, $\langle\varphi|A_n\psi\rangle \longmapsto \langle\varphi|A\psi\rangle$.

The three notions of convergence of sequences of operators have the following relations:

$$\text{uniform convergence} \Rightarrow \text{strong convergence} \Rightarrow \text{weak convergence}$$

but the converses is false. The weak convergence does not imply either strong or uniform convergence and strong convergence does not imply uniform convergence.

Since the space $L(\mathcal{H})$ is a Banach space, when we consider it as a normed space, any Cauchy sequence (with respect to the norm in $L(\mathcal{H})$) of operators converges to an operator in $L(\mathcal{H})$. Such Cauchy sequence is called the *uniform Cauchy sequence*. A sequence $\{A_n\}$ of bounded operators on $\mathcal{H}$ is called the *strong Cauchy sequence* if for each $\psi \in \mathcal{H}$, the sequence $\{A_n\psi\}$ is a Cauchy sequence in $\mathcal{H}$. An important result, which is known as the Banach-Steinhaus theorem, says that any strong Cauchy sequence of bounded in the operator norm.

Let $\mathcal{H}$ be a Hilbert space, and A a bounded operator on $\mathcal{H}$, i.e., $A \in L(\mathcal{H})$. For any fixed $\psi \in \mathcal{H}$, let us define the following functional on $\mathcal{H}$:

$$f(\varphi) = \langle\psi|A\varphi\rangle\ , \qquad \forall\varphi \in \mathcal{H}$$

By the properties of scalar product, f is a linear mapping from $\mathcal{H}$ into $\mathbb{C}$. Then, the Cauchy-Schwartz inequality implies that f is bounded:

$$|f(\varphi)| = |\langle\psi|A\varphi\rangle| = K\,\|\varphi\|\ , \qquad \forall\varphi \in \mathcal{H}\,,$$

with the bound $K = \|A\|\,\|\psi\|$. The Riesz theorem shows that there exists a unique vector $\varphi \in \mathcal{H}$ such that

$$\langle\phi|\varphi\rangle = \langle\psi|A\varphi\rangle\ , \qquad \forall\varphi \in \mathcal{H}.$$

and this holds for all $\psi \in \mathcal{H}$. Thus the mapping:

$$A^*\psi \longmapsto \phi\,,$$

is well defined and linear on $\mathcal{H}$, so it is an operator on $\mathcal{H}$. It has the property that, for any pair of vectors $\psi, \varphi \in \mathcal{H}$

$$\langle\psi|A\varphi\rangle = \langle A^*\psi|\varphi\rangle \tag{B.9}$$

Furthermore, it follows easily from (**??**) and the Cauchy-Schwartz inequality that A^* is bounded and its norm satisfies

$$\|A^*\| \leq \|A\| \tag{B.10}$$

The operator A^* is called the *adjoint* of A.

The mapping $A \longmapsto A^*$ is from $L(\mathcal{H})$ into $L(\mathcal{H})$. Moreover, since it satisfies the properties

$$\begin{aligned} (A+B)^* &= A^* + B^* \\ (\lambda A)^* &= \bar{\lambda} A^* \\ (AB)^* &= B^* A^* \end{aligned} \qquad \text{(B.11)}$$

where $A, B \in L(\mathcal{H})$, $\lambda \in \mathbb{C}$, it is a bounded operator on the Banach algebra $L(\mathcal{H})$. In fact † is an isometry on $L(\mathcal{H})$ because we also have

$$\|A\| = \|A^*\| \, .$$

An operator $A \in L(\mathcal{H})$ is *self-adjoint* if $A = A^*$, i.e. if it is equal to its adjoint. A is *isometric* if $\|A\psi\| = \|\psi\|$, $\forall \psi \in \mathcal{H}$. If A is isometric and onto $\mathcal{H}$ it is called *unitary*.

Proposition B.5 *Let $A \in L(\mathcal{H})$ then A is unitary if and only if there exists its inverse A^{-1} and $A^{-1} = A^*$ or, equivalently, A is unitary if and only if $AA^* = A^*A = I$, where I is the identity operator on $\mathcal{H}$.*

A is self-adjoint if and only if

$$\langle \psi | A\psi \rangle = \langle A\psi | \psi \rangle \, , \qquad \forall \psi \in \mathcal{H}$$

Bounded operators of great importance are the orthogonal projections also called projectors, since they are the key objects in the theory of *spectral decompositions* of self-adjoint operators. Let us recall their main properties.

An *orthogonal projection* or a *projector* on a Hilbert space $\mathcal{H}$ is a bounded operator on $\mathcal{H}$ with the following properties:

i) Idempotency: $P^2 = P$.

ii) Self adjointness: $P = P^*$.

There are two trivial examples of projectors: the identity operator I, which transforms any vector in $\mathcal{H}$ into itself and the zero operator O that transforms every vector in $\mathcal{H}$ into the zero vector. Simple projectors transform (project) all $\mathcal{H}$ onto the subspace spanned by an arbitrary nonzero vector ψ.

Let $\psi \in \mathcal{H}$, $\|\psi\| = 1$, be an arbitrary nonzero vector in the Hilbert space $\mathcal{H}$. Define

$$P\varphi = \langle \psi | \varphi \rangle \, \psi \, , \qquad \forall \varphi \in \mathcal{H}.$$

It is easy to see that P is a projector. The image of $\mathcal{H}$ by P is the one dimensional subspace spanned by ψ, which is closed because it is finite dimensional. Similarly taking a countable family $\{\psi_n\}$ of orthonormal vectors and putting

$$P\varphi = \sum_n \langle \psi_n | \varphi \rangle \, \psi_n \, , \qquad \forall \varphi \in \mathcal{H}, ,$$

we define the orthogonal projector from $\mathcal{H}$ onto the (closed) subspace spanned by $\{\psi_n\}$.

In general we have one-to-one correspondence between closed subspaces of a Hilbert space and the projector. Namely, if P is a projector on $\mathcal{H}$ then the space $\mathcal{M} = P(\mathcal{H})$ is a closed. Conversely, let $\mathcal{M}$ be a closed subspace of $\mathcal{H}$. Then we have the following direct sum decomposition

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp} ,$$

where $\mathcal{M}^{\perp}$ is the *orthogonal complement* of $\mathcal{M}$

$$\mathcal{M}^{\perp} := \{\psi \in \mathcal{H} \quad : \quad \langle \psi | \varphi \rangle = 0 \quad , \quad \forall \varphi \in \mathcal{M}\} .$$

This means that each $\varphi \in \mathcal{H}$ has a unique decomposition

$$\varphi = \psi_1 + \psi_2 ,$$

where $\psi_1 \in \mathcal{M}$, $\psi_2 \in \mathcal{M}^{\perp}$ and $\langle \psi_1 | \psi_2 \rangle = 0$. Thus the mapping P:

$$P\varphi = \psi_1$$

defines the orthogonal projection from $\mathcal{H}$ onto $\mathcal{M}$.

In particular we have

$$P\psi = \psi \, , \ \text{ for each } \psi \in \mathcal{M} \text{ and } P\psi = 0 \text{ if } \varphi \in \mathcal{M}^{\perp} .$$

Observe also that if $P \in L(\mathcal{H})$ is a projector different from the zero operator then $\|P\| = 1$.

Let P_1 and P_2 be two distinct projections different from the zero operator on the Hilbert space $\mathcal{H}$. They are *mutually orthogonal* if and only if $P_1 P_2 = O$, where O is the zero operator. If the projectors P_1 and P_2 are mutually orthogonal, then they commute, i.e. $P_1 P_2 = P_2 P_1$.

Proposition B.6 *Let P_1 and P_2 be two mutually orthogonal projectors and let $\mathcal{M}_1$ and $\mathcal{M}_2$ be their respective associate subspaces. Then, $\mathcal{M}_1 \cap \mathcal{M}_2 = \{0\}$. Furthermore, $\mathcal{M}_1 \subset \mathcal{M}_2^{\perp}$ and $\mathcal{M}_2 \subset M_1^{\perp}$. Conversely, if $\mathcal{M}_1 \subset \mathcal{M}_2^{\perp}$, then, $\mathcal{M}_2 \subset \mathcal{M}_1^{\perp}$, $\mathcal{M}_1 \cap \mathcal{M}_2 = \{0\}$ and $P_1 P_2 = O$.*

Unbounded operators.

Assume till the end of this section that $\mathcal{H}$ is an infinite dimensional Hilbert space. An unbounded operator A acts on the vectors of a dense subspace of $\mathcal{H}$ called the *domain* of A. This domain is denoted as $\mathcal{D}(A)$. Thus, A is a linear mapping from $\mathcal{D}(A)$ onto its *range* or the image space $\mathcal{R}(A)$. The range of A is always a subspace of $\mathcal{H}$ because A is a linear mapping. In general $\mathcal{R}(A)$ is not a subspace of $\mathcal{D}(A)$ and we have to take this in account if we want to define the powers of A. Furthermore, the intersection of two dense subspaces of $\mathcal{H}$ is not, in general, dense and can even be as small as to contain the zero vector only. This should be considered if we want to define the sum or the product of two unbounded operators.

Let A be an operator on $\mathcal{H}$ with the domain $\mathcal{D}(A)$. A is *closed* if and only if for any sequence $\{\psi_n\}$ of vectors in $\mathcal{D}(A)$ such that

$$\psi_n \longmapsto \psi \quad \text{and} \quad A\psi_n \longmapsto \varphi,$$

i.e., both sequences converge, we have:

$$\psi \in \mathcal{D}(A) \quad \text{and} \quad A\psi = \varphi\,.$$

Observe that, if A is continuous, it is closed. The converse is not true, because for an unbounded operator, there exists convergent sequences $\psi_n \longmapsto \psi$ such that the sequence of their images, $\{A\psi_n\}$, does not converge (otherwise A would be continuous).

Let A be an operator on $\mathcal{H}$ with the domain $\mathcal{D}(A)$. The *graph*, $\mathcal{G}(A)$, of A is the subset of $\mathcal{H}\times\mathcal{H}$ of elements of the form $(\psi, A\psi)$, $\forall\psi\in\mathcal{D}(A)$.

Proposition B.7 *Let A be an operator with the domain $\mathcal{D}(A)$ and the graph $\mathcal{G}(A)$. A is closed if and only if $\mathcal{G}(A)$ is closed.*

Note that $\mathcal{G}(A)$ is closed if for any sequence $\{\psi_n\}$ such that

$$\psi_n \longmapsto \psi \quad \text{and} \quad A\psi_n \longmapsto \varphi\,,$$

where $\psi, \varphi \in \mathcal{D}(A)$ we have $\varphi = A\psi$. Hence, $(\psi_n,\, A\psi_n) \longmapsto (\psi, A\psi)$ and $\mathcal{G}(A)$.

Theorem B.2 (Closed Graph Theorem) *Let A be a closed operator with the domain $\mathcal{D}(A) \equiv \mathcal{H}$. Then, A is bounded.*

The above theorem shows that an unbounded operator can not act on the whole Hilbert space. Its domain has to be always taken into account. In consequence some concepts concerning bounded operators do not have a straightforward generalization on unbounded operators. Below we define the adjoint of an unbounded operator and introduce the important notion of self-adjointness.

Let A be an unbounded operator with the domain $\mathcal{D}(A)$. Consider the following subspace of $\mathcal{H}$

$$\mathcal{D}(A^*) := \{\psi \in \mathcal{H} \;\; : \;\; \exists\,\varphi \in \mathcal{H} \;\; ; \;\; \langle\psi|A\phi\rangle = \langle\varphi|\phi\rangle \;, \;\; \forall\phi \in \mathcal{H}\} \tag{B.12}$$

and the transformation:

$$A^*\psi = \varphi\,, \qquad \forall\psi \in \mathcal{D}(A^*) \tag{B.13}$$

A^* is called the *adjoint* of A and $\mathcal{D}(A^*)$ is the *domain* of A^*.

It is easy to see that $\mathcal{D}(A^*)$ is a subspace of $\mathcal{H}$. Moreover, if $\mathcal{D}(A)$ (the domain of A) is dense in $\mathcal{H}$ then A^* is well defined linear operator on $\mathcal{D}(A^*)$. We want to stress the fact that the denseness of $\mathcal{D}(A)$ in $\mathcal{H}$ plays the crucial role in the definition of the adjoint. Without the denseness of $\mathcal{D}(A)$ we cannot guarantee the uniqueness of the image of ψ by A^*.

We see that the adjoint A^* of A is a linear mapping from $\mathcal{D}(A^*)$ into $\mathcal{H}$ and hence an operator on $\mathcal{H}$. However, its domain $\mathcal{D}(A^*)$ is not necessarily dense in $\mathcal{H}$. In such case, the adjoint A^{**} of A^*, does not exists.

Let us point out some simple properties of the adjoints of unbounded operators.

1. If A is bounded, then $\mathcal{D}(A) = \mathcal{H}$. Consequently, $\mathcal{D}(A^*) = \mathcal{H}$.

2.- A^* is always a closed operator. This property is even independent on whether $\mathcal{D}(A^*)$ is dense in $\mathcal{H}$ or not. At this point, it is necessary to remark that for the most general definition of closed and closable operator the denseness of its domain is not needed.

3. If $\alpha \in \mathbb{C}$, then $(\alpha\, A)^* = \bar{\alpha}\, A^*$.

4. Let A and B be two operators on $\mathcal{H}$ with the respective domains $\mathcal{D}(A)$ and $\mathcal{D}(B)$. We say that B *extends* A if

$$\mathcal{D}(A) \subset \mathcal{D}(B) \quad \text{and} \quad A\psi = B\psi\,, \quad \forall \psi \in \mathcal{D}(A)\,.$$

Thus B extends A if and only if the domain of A is contained in the domain of B and A and B coincide on their common domain which is $\mathcal{D}(A)$. We denote this as:

$$A \prec B\,.$$

If B extends A then the adjoint of A extends the adjoint of B:

$$A \prec B \;\Rightarrow\; B^* \prec A^*\,.$$

5. We defined previously the sum and the product of bounded operators simply as we do it in elementary algebra. This could be done because bounded operators are defined on the entire Hilbert space. Since this is not the case for unbounded operators, we have to be more careful in this case. For the sum of two operators A and B, it seems natural to define $\mathcal{D}(A+B) = \mathcal{D}(A) \cap \mathcal{D}(B)$. We have mentioned that the intersection of the domains of two bounded operators may not be dense, so that $A+B$ will not have an adjoint in general. On the other hand, if $\mathcal{D}(A+B)$ is dense in $\mathcal{H}$ then $(A+B)^*$ exists and has the following property:

$$A^* + B^* \prec (A+B)^*\,.$$

6. Similarly, in order to define the product AB, we define first the domain of AB:

$$\mathcal{D}(AB) = \{\psi \in \mathcal{H} \;:\; \psi \in \mathcal{D}(B) \;\text{ and }\; B\psi \in \mathcal{D}(A)\}$$

If $\mathcal{D}(AB)$ is dense the adjoint of AB exists and

$$B^*A^* \prec (AB)^*$$

7. If $\alpha \in \mathbb{C}$, then, $(A + \alpha\, I)^* = A^* + \bar{\alpha}\, I$.
8. If $\mathcal{D}(A^*)$ is dense in $\mathcal{H}$, then A^{**} exists and $A \prec A^{**}$.

We have seen that, in the case of bounded operators, A is self-adjoint if and only if $\langle A\psi|\varphi\rangle = \langle\psi|A\varphi\rangle$ for any pair of vectors $\psi, \varphi \in \mathcal{H}$. This condition is however not sufficient for an unbounded operator A when only replacing the condition $\psi, \varphi \in \mathcal{H}$ by $\psi, \varphi \in \mathcal{D}(A)$. It follows from the definition of the domain of A^* that the identity $\langle A\psi|\varphi\rangle = \langle\psi|A\varphi\rangle$, $\forall\psi, \varphi \in \mathcal{D}(A)$ implies that $\mathcal{D}(A) \subset \mathcal{D}(A^*)$ and nothing else. To have $A = A^*$, we must have $\mathcal{D}(A) = \mathcal{D}(A^*)$. An operator A with the domain $\mathcal{D}(A)$ dense in $\mathcal{H}$ is called *symmetric* if for any $\psi, \varphi \in \mathcal{H}$, we have:

$$\langle A\psi|\varphi\rangle = \langle\psi|A\varphi\rangle$$

In other words, A is symmetric if and only if $\mathcal{D}(A) \subset \mathcal{D}(A^*)$.

An operator A on the Hilbert space $\mathcal{H}$ is *self-adjoint* if and only if $A = A^*$. Equivalently A is self-adjoint if and only if $\mathcal{D}(A) = \mathcal{D}(A^*)$ and $A\psi = A^*\psi$, for any $\psi \in \mathcal{D}(A^*)$.

We would like to end this section giving two important examples of unbounded self-adjoint operators. Consider the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$ and the subspace

$$\mathcal{D}(Q) := \{\varphi(x) \in L^2(\mathbb{R}) \,/\, x\varphi(x) \in L^2(\mathbb{R})\}$$

This subspace is dense in $L^2(\mathbb{R})$ because it contains the Hermite functions which form a complete orthonormal system on $L^2(\mathbb{R})$. Let us define the operator $Q : \mathcal{D}(Q) \longmapsto L^2(\mathbb{R})$ as

$$Q\,\varphi(x) := x\,\varphi(x)\,, \qquad \forall\varphi(x) \in \mathcal{D}(Q)$$

This operator is self-adjoint on $L^2(\mathbb{R})$ and is called the one dimensional *position operator*.

Now, let $\mathcal{D}(P)$ be the vector space of all square integrable functions that admit a derivative almost elsewhere with respect to the Lebesgue measure on the real line and such that this derivative is also square integrable. Define

$$P\,\varphi(x) = -i\varphi'(x)\,, \qquad \forall\varphi(x) \in \mathcal{D}(P)\,.$$

P is also self-adjoint operator on $L^2(\mathbb{R})$. It is called the one dimensional *momentum operator*.

The position and momentum operators can be easily generalized to n dimensions. In this case

$$\mathcal{D}(Q) := \{\varphi \in L^2(\mathbb{R}^n) \;:\; x_i\,\varphi(\mathbf{x}) \in L^2(\mathbb{R}^3)\; i = 1, \ldots, n\}\,,$$

where $\mathbf{x} = (x_1, \ldots, x_n)$. Then, we define the i-th component of the position operator as

$$Q_i\,\varphi(\mathbf{x}) := x_i\,\varphi(\mathbf{x})\,, \qquad \forall\varphi \in L^2(\mathbb{R}^n)$$

The operators $Q_1, \ldots, Q_n$ are self-adjoint on their respective domains $\mathcal{D}(Q_i)$.

Similarly the i-th component of the momentum operator on $L^2(\mathbb{R}^n)$ is given by $P_i = -i\partial/\partial x_i$.

Let us also note that neither position nor momentum operators are self-adjoint when defined on $L^2_{[a,b]}$, where at least one of the ends of the interval is finite.

A self-adjoint operator, either bounded or unbounded does not have residual spectrum. The spectrum, which is a closed subspace of the real line and eventually may coincide with the whole real line, is divided only into two categories: eigenvalues and continuous spectrum.

Let $\mathcal{H}$ be a Hilbert space and $\mathcal{P}$ the class of projectors on $\mathcal{H}$. A *spectral measure*, $E(\cdot)$, on $\mathbb{R}$ is a mapping from the set, $\mathcal{B}$, of all Borel sets in $\mathbb{R}$:

$$E(\cdot) : \mathcal{B} \longmapsto \mathcal{P} \tag{B.14}$$

associating to $C \in \mathcal{B}$ a projector $E(C)$ on $\mathcal{H}$, such that:

i) $E(\mathbb{R}) = I$.

ii) If $\{C_1, C_2, \ldots\}$ is a sequence (finite or countably infinite) of Borel sets that are pairwise disjoint, i.e.,:

$$C_i \cap C_j = \emptyset \ , \qquad i \neq j$$

then,

$$E(\bigcup_{n=1}^{N} C_n) = \sum_{n=1}^{N} E(C_n) \tag{B.15}$$

Here, N is the number (finite or infinite) of non-empty Borel sets in the sequence $\{C_1, C_2, \ldots\}$ and the series converges in the strong operator sense.

Observe that (B.15) says that, if the Borel sets $\{C_1, C_2, \ldots\}$ are pairwise disjoint, the strong sum $\sum_{n=1}^{N} E(C_n)$ is a projector. This can only happen if the projectors in the sequence $\{E(C_n)\}$ are pairwise orthogonal, which means that

$$E(C_n)E(C_m) = O \ , \qquad n \neq m$$

In particular, if C and D are two disjoint Borel sets, then, $E(C)E(D) = O$.

There are some other consequences that immediately follow from the definition of spectral measure, such as:

i) $E(\emptyset) = O$

ii) If $C \cap D$ are arbitrary Borel sets, we can prove that $E(C \cap D) = E(C)E(D)$.

Let $E(\cdot)$ be a spectral measure on $\mathbb{R}$. The *spectral family* of $E(\cdot)$ is the set of projectors $\{E_\lambda\}$, indexed by $\lambda \in \mathbb{R}$ and defined as:

$$E_\lambda := E(-\infty, \lambda] \tag{B.16}$$

Let $E(\cdot)$ a spectral measure and $\{E_\lambda\}$ its corresponding spectral family. Then, $\{E_\lambda\}$ has the following properties:

i) If $\lambda \leq \mu$, then $E_\lambda \prec E_\mu$.

ii) A spectral family is strongly continuous from the right, i.e.

$$E_\lambda = \lim_{\substack{\mu \mapsto \lambda \\ \mu \geq \lambda}} E_\mu \tag{B.17}$$

and the limit is taken in the strong sense.

iii) The following limits hold:

$$E_{-\infty} = \lim_{\lambda \mapsto -\infty} E_\lambda = O \quad \text{and} \quad E_\infty = \lim_{\lambda \mapsto \infty} E_\lambda = I \tag{B.18}$$

Conversely, if there is a mapping from $\mathbb{R}$ into $\mathcal{P}$, $\lambda \longmapsto E_\lambda$, satisfying the above properties, there is a unique spectral measure $E(\cdot)$ such that its spectral family coincides with $\{E_\lambda\}$.

We will show now that spectral families determine self-adjoint and unitary operators.

Proposition S.37.- Let $\{E_\lambda\}$ be a spectral family and $\psi \in \mathcal{H}$. Then, $\mu(d\lambda) := d\langle\psi|E_\lambda\psi\rangle$ is a measure on $\mathbb{R}$ with $\mu(\mathbb{R}) = ||\psi||^2$.

The next result is one of the most important results on the theory of self-adjoint operators:

Theorem B.3 (Spectral Decomposition Theorem for Self Adjoint Operators) *Let A be a self-adjoint operator on $\mathcal{H}$. Then, there exists a spectral family $\{E_\lambda\}$ such that for any $\psi \in \mathcal{D}(A)$, we have that*

$$\langle\psi|A\psi\rangle = \int_{\mathbb{R}} \lambda \, d\langle\psi|E_\lambda\psi\rangle \tag{B.19}$$

Conversely, for any spectral family $\{E_\lambda\}$, there exists a unique self-adjoint operator A such that (B.19) holds.

Formula (B.19) is often written as:

$$A = \int_{\mathbb{R}} \lambda \, dE_\lambda \tag{B.20}$$

For all $\psi \in \mathcal{D}(A)$, the measure $d\mu(\lambda) := d\langle\psi|E_\lambda\psi\rangle$ vanishes outside the spectrum $\sigma(A)$ of A. For this reason, formula (B.19) can be written as

$$\langle\psi|A\psi\rangle = \int_{\sigma(A)} \lambda \, d\langle\psi|E_\lambda\psi\rangle \tag{B.21}$$

or shortly

$$A = \int_{\sigma(A)} \lambda \, dE_\lambda \tag{B.22}$$

The Stone theorem establishes that if $U(t)$, $t \in \mathbb{R}$ is a one parameter group of unitary operators such that for all $t_0 \in \mathbb{R}$, $\lim_{t \mapsto t_0} U(t) = U(t_0)$ in the strong sense, then, there is a self-adjoint operator A such that $U(t) = e^{itA}$. If A is bounded then the exponential operator can be defined as

$$e^{itA} = \sum_{n=0}^{\infty} \frac{(it)^n A^n}{n!} \,, \tag{B.23}$$

because the series on the right hand side converges uniformly. If A is unbounded the above definition is in general meaningless. However, we can give a meaning to e^{itA} by the means of functional calculus. Its main result is the following:

Theorem S.39.- Let $f(\lambda)$ a bounded measurable complex valued function on $\mathbb{R}$. Then, the symbol $f(A)$ defined as:

$$\langle \psi | f(A) \psi \rangle = \int_{\mathbb{R}} f(\lambda)\, d\langle \psi | E_\lambda \psi \rangle \tag{B.24}$$

represents a bounded operator on $\mathcal{H}$ which is self-adjoint if and only if the function $f(\lambda)$ is real.

Note that, since $f(\lambda)$ is bounded and the measure $d\mu(\lambda) = d\langle \psi | E_\lambda \psi \rangle$ is finite, the integral in (B.24) always exists. If $f(\lambda)$ is not bounded, the integral in (B.24) does not converge in general. If $f(\lambda)$ is real and the integral converges for any ψ in a dense subspace of $\mathcal{H}$, $f(A)$ defines a symmetric operator. This happens for instance if $f(\lambda)$ is a polynomial, $f(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \ldots + \alpha_n$. Then, $f(A) = A^n + \alpha_1 A^{n-1} + \ldots + \alpha_n$. In particular, if $f(\lambda) = e^{it\lambda}$, we have:

$$\langle \psi | e^{itA} \psi \rangle = \int_{\mathbb{R}} e^{it\lambda}\, d\langle \psi | E_\lambda \psi \rangle \tag{B.25}$$

or briefly,

$$e^{itA} = \int_{\mathbb{R}} e^{it\lambda}\, dE_\lambda \tag{B.26}$$

A similar result holds for unitary operators:

Theorem B.4 *(Spectral Decomposition Theorem for Unitary Operators) Let U be a unitary operator. Then, there is a spectral family, $\{E_\lambda\}$, such that for any $\psi \in \mathcal{H}$, we have that*

$$\langle \psi | U \psi \rangle = \int_{\mathbb{R}} e^{i\lambda}\, d\langle \psi | E_\lambda \psi \rangle \tag{B.27}$$

or briefly,

$$U = \int_{\mathbb{R}} e^{i\lambda}\, dE_\lambda \tag{B.28}$$

The spectral family $\{E_\lambda\}$ determines the unitary operator uniquely.

Summarizing, we see that (B.25) or (B.26) determines a group of unitary operators. The spectral family $\{E_\lambda\}$ determines a unique self-adjoint operator A and a unique unitary operator U. Therefore, they must be related and the relation is given by (B.26) with $t = 1$, i.e., $U = e^{iA}$. Thus, for any unitary operator U always exists a self-adjoint operator A with $U = e^{iA}$ and all self-adjoint operator can be exponentiated (multiplied by imaginary unit i) to give a unitary operator.

This spectral decomposition theorem permits a new classification of the spectrum of a self-adjoint operator. This is the subject of our next discussion.

Appendix C
Spectral analysis of dynamical systems

In the previous section we characterized the ergodic properties of dynamical systems in terms of the asymptotic behavior of evolution operators. On the other hand, it is known that the full information about some properties of operators is contain in their spectra. As spectral theory is a powerful tool in studying linear operators on topological vector spaces it would be desirable to express also the ergodic properties in terms of spectral theory.

In this section we shall present the basic facts concerning the relations between ergodic properties of evolution operators and their spectral properties. We focus our attention mainly on Kolmogorov systems as they play a prominent role throughout this book. We shall also elaborate on this subjects in the next sections. Especially in Section 6 which is devoted to exact systems.

If $(\mathcal{X}, \Sigma, \mu, \{S_t\})$ is discrete time dynamical systems then its evolution is described by a single Koopman operator $V = V^1$ or its adjoint - the Frobenius-Perron operator $U = V^*$. The *spectrum of the evolution group* $\{V^n\}$ is the spectrum of the operator V.

If time is continuous and the transformations S_t are invertible then the family $\{V_t\}$, $t \in \mathbb{R}$, forms a group. The *spectrum of the evolution group* $\{V_t\}$ is the spectrum of the self-adjoint generator L, $V_t = e^{itL}$. A similar approach can be applied when S_t

are non-invertible and $\{V_t\}$ is only a semigroup. While the correspondence between unitary groups and self-adjoint generators is well known from the course of functional analysis [Yo] (see also Section 2 and Section 14), the analogues correspondence for evolution semigroup requires some comments.

If $\{V_t\}$ is a continuous semigroup of contraction on L^p, $p \geq 1$, then its generator, called also the *infinitesimal operator* A, is defined as

$$A\rho = \lim_{t \to 0} \frac{V_t \rho - \rho}{t} \,. \tag{C.1}$$

where the limit is considered in the sense of L^p-norm. The domain $D(A)$ of A consists of those $f \in L^p$, for which the limit in (C.1) exists. It can be proved that $D(A)$ is dense in L^p and $\{V_t\}$ is related to A by the relation

$$V_t = e^{-tA} \stackrel{\text{df}}{=} s.\lim_{n\to\infty} \left(1 + \frac{t}{n} A\right)^{-n} \,.$$

In addition the resolvent set of A contains the negative half-axis $(-\infty, 0)$ and

$$\|(A + \lambda I)^{-1}\| \leq \frac{1}{\lambda}, \;\; \lambda > 0$$

Conversely if A in an arbitrary operator densely defined on L^p satisfying the above condition imposed on its resolvent then it is a generator of a semigroup of contractions. The interested reader can find the proof of the above mentioned fact, and many other related results, in the monographs [HiPh,Yo]. A valuable information about the generators of dynamical semigroups can be found in [LM].

Let us concentrate now on the spectral description of ergodic properties of invertible dynamical system. We assume, as usually, that the underlying measure is normalized and invariant with respect to the group transformations of the phase space.

In this case the evolution operator are unitary and the description of spectra introduced in Section 2 can be applied. The spectrum of an arbitrary group of evolution can have all three parts point spectrum, singular continuous and absolutely continuous. However, the ergodic properties imposes additional conditions, which we present below.

(I) Ergodicity

Ergodicity of a dynamical system amounts to imposing additional condition on the discrete part of the spectrum. In the case of discrete time ergodicity is equivalent to the condition that 1 is a simple eigenvalue of the Koopman operator V. Indeed, ergodicity, in terms of evolution operators, means that the only V invariant functions are constants. Thus for ρ constant $V\rho = \rho$, which implies that 1 is an eigenvalue. This eigenvalue is simple because the dimension of the subspace consisting of constant functions is 1. A dynamical system with continuous time is ergodic if and only if 0 is a simple eigenvalue of the infinitesimal generator L.

It is worth to stress that ergodicity is the weakest ergodic property among that introduced in Section 2. Observe also that because 1 is always an eigenvalue of the evolution operators, we can decompose the Hilbert space $L^2_{\mathcal{X}}$ as the orthogonal sum

$L^2_{\mathcal{X}} = \mathcal{H} \oplus [1]$, where $[1]$ denoted the linear space spanned by constants. Now we can ask about the spectral properties of the evolution operators on $\mathcal{H}$. Ergodicity does not imposes any additional condition but, as we shall see below, stronger ergodic properties do. In contrast, the spectrum of the evolution operator restricted to $\mathcal{H}$ is, for stronger ergodic properties, at least continuous. This is the reason that the operator evolution groups and associated time operators are usually considered as acting on space $\mathcal{H}$ - the orthogonal complement of constants. Thus the expression that a dynamical group has, for example, continuous spectrum will always refer to the space $\mathcal{H}$.

(II) Weak mixing

A *dynamical system is weakly mixing if and only if the evolution group has continuous spectrum.*

We shall give a sketch of the proof that weakly mixing systems have continuous spectra. Another, and complete, proof of this implication, as well as, the proof of the converse implication can be found in [Ha].

Consider a discrete time dynamical system and the Koopman operator V. Then consider the sequence

$$c_n(\rho) = (V^n \rho, \rho)\,.$$

Recall that a unitary operator has the representation

$$V^n = \int_0^{2\pi} e^{i\lambda n}\, dE_\lambda, \quad \text{for each } n \in \mathbb{Z}$$

where $\{E_\lambda\}_{\lambda \in [0,2\pi]}$, is a resolution of identity (see Section 2). Thus

$$c_n(\rho) = \langle V^n \rho, \rho\rangle = \int_0^{2\pi} e^{in\lambda}\, d\langle E_\lambda \rho, \rho\rangle$$

are the Fourier-Stieltjes coefficients of the non-decreasing function $G_\rho : \lambda \longmapsto \langle E_\lambda \rho, \rho\rangle$.

If the system is weakly mixing then

$$\lim_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} |c_n(\rho)| = \lim_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} |(V^n \rho, \rho)| = 0$$

for each $\rho \in \mathcal{H}$ (i.e. $\int_{\mathcal{X}} \rho\, d\mu = 0$)

Now by the classical results concerning Fourier-Stieltjes coefficients (see [Zy, III, 9.6]) the function G_p is continuous everywhere on $[0, 2\pi]$. If time is continuous then the unitary group has the spectral representation

$$V_t = \int_{-\infty}^{\infty} e^{it\lambda} dE_\lambda\,.$$

For a given $\rho \in \mathcal{H}$ the expression

$$\langle V_t\rho, \rho\rangle = \int_{-\infty}^{\infty} e^{it\lambda} d\langle E_\lambda \rho, \rho\rangle$$

is nothing but the Fourier-Stieltjes transform of the function $\lambda \longmapsto \langle E_\lambda \rho, \rho\rangle$. Now, if $F(x)$ is a function of bounded variation and $\phi(s)$ its Fourier-Stieltjes transform then the condition

$$\lim_{M\to\infty} \frac{1}{M} \int_{-M}^{M} |\phi(s)|\, ds < \infty$$

is necessary and sufficient for $F(x)$ to be continuous ([Zy, XVI, 4.19]). This proves the equivalence between weak mixing and continuous spectrum for continuous time.

The continuity of the distribution function G_p implies that the measure σ_ρ determined by G_ρ is continuous. This means that in the Jordan decomposition of σ_ρ the discrete component is 0.

Measure σ_ρ is called the *spectral measure* of the “vector" $\rho \in \mathcal{H}$.

(III) Mixing

A dynamical system is mixing if $\lim_{t\to\infty}(\rho_1, V_t\rho_2) = 0$ for each $\rho_1, \rho_2 \in \mathcal{H}$. It is therefore obvious that mixing system is also weakly mixing. That the converse is not true has been shown by Kakutani (LNM318). In consequence the dynamical group $\{V_t\}$ has continuous spectrum on $\mathcal{H}$.

It is easy to see that a sufficient condition for mixing is that $\{V_t\}$ has absolutely continuous spectrum. Indeed if this is the case then the derivative $\frac{d\langle E_\lambda \rho, \rho\rangle}{d\lambda}$ exists for almost all $\lambda \in \mathbb{R}$. Thus applying the Riemann-Lebesgue Lemma we obtain

$$\langle \rho, V_t\rho\rangle = \int_{-\infty}^{\infty} e^{it\lambda} \frac{d\langle E_\lambda \rho, \rho\rangle}{d\lambda}\, d\lambda \to 0 \text{ as } t \to \pm\infty\,.$$

The condition that $\{V_t\}$ has absolutely continuous spectrum is, however not necessary. The complete characterization of the ergodic property of mixing has been obtained in recent years. For this reason we include it here although it is not directly connected with time operators.

Consider, an abstract unitary group $V_t = e^{itL}$ acting on a Hilbert space $\mathcal{H}$, generated by the self-adjoint operator L with spectral family $\{E_\lambda\}$:

$$L = \int_{-\infty}^{\infty} \lambda\, dE_\lambda\,.$$

Denote by $\mathcal{H}_{\mathrm{p}}$ the closed linear hull of all eigenvectors of L. The continuous subspace of L is the orthocomplement of $\mathcal{H}_{\mathrm{p}}$: $\mathcal{H}_{\mathrm{c}} = \mathcal{H} \ominus \mathcal{H}_{\mathrm{p}}$. Recall that the singular continuous subspace $\mathcal{H}_{\mathrm{sc}}$ of $\mathcal{H}_{\mathrm{c}}$ consists of all $f \in \mathcal{H}_{\mathrm{c}}$ for which there exists a Borel set B_0 of Lebesgue measure zero such that $\int_{B_0} dE_\lambda f = f$. By $\mathcal{H}_{\mathrm{ac}} = \mathcal{H}_{\mathrm{c}} \ominus \mathcal{H}_{\mathrm{sc}}$

we shall denote the absolutely continuous subspace of $\mathcal{H}_{\rm c}$. Recall also that $\mathcal{H}_{\rm p}$, $\mathcal{H}_{\rm c}$, $\mathcal{H}_{\rm sc}$ and $\mathcal{H}_{\rm ac}$ are closed linear subspaces of $\mathcal{H}$ which reduce the operator L and that $\mathcal{H} = \mathcal{H}_{\rm p} \oplus \mathcal{H}_{\rm ac} \oplus \mathcal{H}_{\rm sc}$.

The spectra of the corresponding reductions of L will be called respectively point, continuous, singular continuous and absolutely continuous spectrum of L, and will be denoted by $\sigma_{\rm p}(L)$, $\sigma_{\rm c}(L)$, $\sigma_{\rm sc}(L)$ and $\sigma_{\rm ac}(L)$ correspondingly [Wa] (for more details see [Weid,Kato].

Let $\mu = \mu_h$ denotes, for a given $h \in \mathcal{H}$, the spectral measure on $\sigma(L)$ determined by the nondecreasing function

$$F_h(\lambda) = \langle h, E_\lambda h \rangle \ , \quad \text{for} \ \ \lambda \in \mathbb{R} \ .$$

Let $h = h_{\rm p} + h_{\rm ac} + h_{\rm sc}$ be the decomposition of h corresponding to the direct sum $\mathcal{H}_{\rm p} \oplus \mathcal{H}_{\rm ac} \oplus \mathcal{H}_{\rm sc}$. Putting $\mu_{\rm p} = \mu_{h_{\rm p}}$, $\mu_{\rm ac} = \mu_{h_{\rm ac}}$ and $\mu_{\rm sc} = \mu_{h_{\rm sc}}$ we obtain the Jordan decomposition of μ

$$\mu = \mu_{\rm p} + \mu_{\rm sc} + \mu_{\rm ac} \tag{C.2}$$

onto the point, singular continuous and absolutely continuous component.

It turns out that it is possible to obtain further refinement of the spectral decomposition of L or, equivalently, the group $\{V_t\}$. Let us call the *decaying elements* those elements $h \in \mathcal{H}$ which satisfy

$$(h, V_t h) \longrightarrow 0 \, , \ \ \text{as} \, t \to \infty \, .$$

and denote by $\mathcal{H}_{\rm sc}^{\rm D}$ the set of all decaying elements in the singular continuous subspace $\mathcal{H}_{\rm sc}$. The space $\mathcal{H}_{\rm sc}^{\rm D}$ consists of all vectors $h \in \mathcal{H}_{\rm sc}$ such that the corresponding measure $\mu = \mu_h$ is singular with respect to the Lebesgue measure and its Fourier transform is 0 in infinity. An important result is that $\mathcal{H}_{\rm sc}^{\rm D}$ is a closed linear subspace $\mathcal{H}_{\rm sc}$. Moreover $\mathcal{H}_{\rm sc}$ reduces L, i.e. the action of L does not lead out of the space $\mathcal{H}_{\rm sc}^{\rm D}$:

$$L(D(L) \cap \mathcal{H}_{\rm sc}^{\rm D}) \subset \mathcal{H}_{\rm sc}^{\rm D} \, ,$$

where $D(L)$ is the domain of L.

Let us also introduce the space $\mathcal{H}_{\rm sc}^{\rm ND}$ of all $h \in \mathcal{H}_{\rm sc}$ such that any measure ν which is absolutely continuous with respect to μ_h does not decay. The space $\mathcal{H}_{\rm sc}^{\rm ND}$ will be called the space of *non decaying singular elements*. The space $\mathcal{H}_{\rm sc}^{\rm ND}$ also reduces L. This leads to the direct sum decomposition $\mathcal{H}_{\rm sc} = \mathcal{H}_{\rm sc}^{\rm D} \oplus \mathcal{H}_{\rm sc}^{\rm ND}$ and, consequently, to the following direct sum decomposition of the whole space $\mathcal{H}$:

$$\mathcal{H} = \mathcal{H}_{\rm p} \oplus \mathcal{H}_{\rm ac} \oplus \mathcal{H}_{\rm sc}^{\rm D} \oplus \mathcal{H}_{\rm sc}^{\rm ND} \, . \tag{C.3}$$

Therefore denoting the corresponding spectra of reduced operators by $\sigma_{\rm p}$, $\sigma_{\rm ac}$, $\sigma_{\rm sc}^{\rm D}$ and $\sigma_{\rm sc}^{\rm ND}$ respectively we obtain a new decomposition of the spectrum σ of any self-adjoint operator which is the missing necessary fact to describe mixing and decay:

$$\sigma = \sigma_{\rm p} \cup \sigma_{\rm ac} \cup \sigma_{\rm sc}^{\rm D} \cup \sigma_{\rm sc}^{\rm ND} \, .$$

We are therefore led to the following definitions:

$\mathcal{H}^{\mathrm{D}} = \mathcal{H}_{\mathrm{ac}} \oplus \mathcal{H}_{\mathrm{sc}}^{\mathrm{D}}$ the space of decaying elements with respect to V_t

$\mathcal{H}^{\mathrm{ND}} = \mathcal{H}_{\mathrm{p}} \oplus \mathcal{H}_{\mathrm{sc}}^{\mathrm{ND}}$ the space of non-decaying elements with respect to V_t

$\sigma^{\mathrm{D}} = \sigma_{\mathrm{ac}} \cup \sigma_{\mathrm{sc}}^{\mathrm{D}}$ the decay spectrum of L or V_t

$\sigma^{\mathrm{ND}} = \sigma_{\mathrm{p}} \cup \sigma_{\mathrm{sc}}^{\mathrm{ND}}$ the non-decay spectrum of L or V_t

and to the following spectral characterization of mixing:

Theorem C.1 *A dynamical system is mixing if and only if its group of evolution* $\{V_t\}_{t\in\mathbb{R}}$ *on* $\mathcal{H} = L^2 \ominus \{1\}$ *has purely decaying spectrum, i.e.*

$$\sigma(L) = \sigma^{\mathrm{D}} = \sigma_{\mathrm{ac}} \cup \sigma_{\mathrm{sc}}^{\mathrm{D}} \,. \tag{C.4}$$

(IV) Kolmogorov systems

We begin with K-systems, i.e. discrete time Kolmogorov systems. Let $S : \mathcal{X} \to \mathcal{X}$ be an invertible measure preserving transformation and Σ_0 a distinguish σ-algebra. Then the family of σ-algebras $\Sigma_n = S(\Sigma_0),\ n \in \mathbb{Z}$, satisfies the three conditions characterizing K-system. Let us denote by $L^2_{\mathcal{X}}(\Sigma_n)$ the subspace of $L^2_{\mathcal{X}}$ $(= L^2_{\mathcal{X}}(\Sigma))$ consisting of all functions $f \in L^2_{\mathcal{X}}$ which are Σ_n measurable.

For K-systems it is more convenient to consider the Frobenius-Perron operator U, $U\rho(\omega) = f(S^{-1}\omega)$, because the increasing order.

$$U(L^2_{\mathcal{X}}(\Sigma_0)) = L^2_{\mathcal{X}}(\Sigma_1)$$

is preserved and because we will be interested in the evolution of densities as time goes to ∞.

Denote by $\mathcal{H}$ the orthogonal complement of constants in $L^2_{\mathcal{X}}$ and by $\mathcal{H}_0$ the space

$$\mathcal{H}_o = L^2_{\mathcal{X}}(\Sigma_0) \ominus L^2_{\mathcal{X}}(\Sigma_{-1}) \,.$$

In particular, $\mathcal{H}_0$ is also orthogonal to constants. Let

$$\mathcal{H}_n \stackrel{\mathrm{df}}{=} U(\mathcal{H}_0) \,,$$

then

$$L^2_{\mathcal{X}} = \mathcal{H} \oplus [1] = \bigoplus_{n=-\infty}^{\infty} \mathcal{H}_n \oplus [1] \,.$$

Thus the operator U is a bilateral shift on the space $\mathcal{H}$ (for more detailed information about shift operators see Section 5). The space $\mathcal{H}_0$ is the generating space for this shift. For any orthonormal basis $\{\phi_k\}_{k\in K}$ in $\mathcal{H}$ the shifts $\{U^n\phi_k\}_{k\in K}$ form an orthonormal basis in $\mathcal{H}_k$. The whole family $U^n\phi_n,\ n \in \mathbb{Z},\ k \in K$, is an orthogonal basis of $\mathcal{H}$. We, of course, assume that all considered Hilbert space are separable. In the case of an L^2-space its separability is equivalent to separability of the underlying measure space.

The dimension of the orthogonal basis of the space $\mathcal{H}_0$ is called the *multiplicity* of the shift U. We shall show now that for K-systems the multiplicity of the shift is always (countably) infinite. First let us note, however, that the measure space $(\mathcal{X}, \Sigma, \mu)$ is non-atomic. Recall that the set $A \in \Sigma$ is called an atom of μ if $A \neq \emptyset$ and for each $B \subset A$ either $B = \emptyset$ or $B = A$. Suppose then that A is an atom and consider the sets $A \cap S^{-n}(A),\ n = 1, 2, \ldots$. It must be either $A \cap S^{-n}(A) = \emptyset$ for each n or $A \cap S^{-n_0}(A) = A$ for some n_0. The first case must be excluded, since otherwise we would have

$$\emptyset = S^{-m}(\emptyset) = S^{-m}(A \cap S^{-n}(A)) = S^{-m}(A) \cap S^{-(m+n)}(A)\,,$$

for each $m = 1, 2, \ldots$. This would mean that the sets $S^{-k}(A),\ k = 1, 2, \ldots$ are pointwise disjoint, which would imply

$$\begin{aligned} 1 \quad &= \mu(\mathcal{X}) \geq \mu(S^{-1}(A) \cup S^{-2}(A) \cup \ldots) \\ &= \mu(S^{-1}(A)) + \mu(S^{-2}(A)) + \ldots \\ &= \mu(A) + \mu(A) + \ldots \\ &= \infty\,. \end{aligned}$$

The second case means that $S^{-n_0}(A) \supset A$. Thus, by the invariance of S,

$$S^{-n_0}(A) = A \cup N\,,$$

where $\mu(N) = 0$. Similarly, $S^{-kn_0}(A) = A \cup N_k$, where $k = 2, 3, \ldots$ and $\mu(N_k) = 0$. Since we assume that considered measures are complete, $A \in \bigcap_n \Sigma_n$. As the latter intersection is the trivial σ-algebra, $\mu(A) = 1$ but that would mean that the whole $\mathcal{X}$ is an atom and that $(\mathcal{X}, \Sigma, \mu)$ is trivial.

If a measure space is non-atomic then the Hilbert space of all square integrable functions is infinite dimensional. We can prove now the following:

Lemma. Let $(\mathcal{X}, \mathcal{A}, \mu)$ be a probability space and $\mathcal{B}$ a proper non-atomic sub-σ-algebra of $\mathcal{A}$. Then the space $L^2_{\mathcal{X}}(\mathcal{A}) \ominus L^2_{\mathcal{X}}(\mathcal{B})$ is infinite dimensional.

Proof. Since $\mathcal{B}$ is a proper sub-σ-algebra of $\mathcal{A}$ there exists a function $f \in L^2_{\mathcal{X}}(\mathcal{A})$ which is not $\mathcal{B}$-measurable, i.e. $E(f|\mathcal{B}) \neq f$ on some set $B \in \mathcal{B},\ \mu(B) > 0$. Therefore the function

$$g \stackrel{\mathrm{df}}{=} [f - E(f|\mathcal{B})]\mathbb{1}_B$$

is correctly defined and orthogonal to $L^2_{\mathcal{X}}(\mathcal{B})$. Indeed, if $h \in L^2_{\mathcal{X}}(\mathcal{B})$ then

$$\begin{aligned} \int_{\mathcal{X}} gh d\mu \quad &= \int_{\mathcal{X}} E(gh|\mathcal{B})\, d\mu \\ &= \int_{\mathcal{X}} h\mathbb{1}_B E(f - E(f|B))\, d\mu \\ &= 0\,. \end{aligned}$$

Consider now the space $L^2_B(\mathcal{B})$ of all functions defined on B, measurable with respect to the σ-algebra $\mathcal{B}$ restricted to B and square integrable. Because the $\mathcal{B}$ is non-atomic there exists a sequence $C_n,\ n = 1, 2, \ldots$, of pairwise disjoint and $\mathcal{B}$-measurable subsets of B such that $\mu(C_n) > 0$, for each n. Then putting $g_n = g\mathbb{1}_{C_n}$ we obtain infinitely many linearly independent functions in the space $L^2_{\mathcal{X}}(\mathcal{A}) \ominus L^2_{\mathcal{X}}(\mathcal{B})$.

Applying the above Lemma to the family of Hilbert spaces $\mathcal{H}_n = L^2_{\mathcal{X}}(\Sigma_n) \ominus L^2_{\mathcal{X}}(\Sigma_{n-1})$, we conclude that the multiplicity of the bilateral shift U is infinite. The K-system has therefore the following property:

> *There exists an orthogonal basis of $L^2_{\mathcal{X}}$ formed by the function 1 and by the functions $f_{k,n}$, where $k = 1,\ 2, \ldots$, $n \in \mathbb{Z}$ such that*
>
> $$U f_{k,n} = f_{k,n+1}\,,\ \textit{for every}\ k, n\,.$$

In general, it is said that a dynamical system having the above property has the *uniform Lebesgue spectrum with (countably) infinite multiplicity.*

A similar spectral characterization can be obtained for K-flows. We have the following:

Theorem The unitary group $\{U_t\}$ on the space $L^2_{\mathcal{X}}(\Sigma) \ominus [1]$ has a uniform countable Lebesgue spectrum.

The term spectrum of $\{U_t\}$ in the above theorem refers to each $U_t,\ t \neq 0$, separately. For the proof of this theorem see [KSF]. Here we would like to comment on the terminology Lebesgue spectrum.

Let $U_t = \int_{-\infty}^{\infty} e^{it\lambda} dG_\lambda$ be the spectral resolution of the unitary group $\{U_t\}$. Consider an arbitrary $\rho \in \mathcal{H}$ and the cyclic space $\mathcal{C}(\rho)$ spanned by all $U_t\rho,\ t \in \mathbb{R}$. Vector ρ corresponds to the spectral measure σ_ρ on the Borel σ-algebra on $\mathbb{R}$ determined by the function $\lambda \longmapsto \langle G_\lambda \rho, \rho\rangle$. Any other vector $\rho_1 \in \mathcal{C}(\rho)$ corresponds to the measure σ_{ρ_1}, which is absolutely continuous with respect to σ_ρ. Let E_s be the orthogonal projection on the space $L^2_{\mathcal{X}}(\Sigma_s) \ominus [1],\ s \in \mathbb{R}$. Then, as it was shown in Proposition 1, we have the imprimitivity condition:

$$E_{s+t} = U_t E_s U_{-t}\,.$$

Note that for each $a \in \mathbb{R}$ the translation of the measure σ_ρ by a, i.e. $\sigma_\rho(\cdot + a)$ corresponds to the function

$$\lambda \longmapsto \langle E_{\lambda+a}\rho, \rho\rangle\,.$$

But, by the imprimitivity condition

$$\langle E_{\lambda+a}\rho, \rho\rangle = \langle E_\lambda U_{-a}\rho, U_{-a}\rho\rangle\,.$$

Since $U_{-a}\rho \in \mathcal{C}(\rho)$, for each $a \in \mathbb{R}$, all the translations $\sigma_\rho(\cdot + a)$ are absolutely continuous with respect to σ_ρ. This implies that σ_ρ is equivalent to the Lebesgue measure. This justifies the use of the term Lebesgue spectrum.

References

1. Aaronson, J., M. Lin, and B. Weiss (1979), Mixing properties of Markov operators and ergodic transformations, and ergodicity of Cartesian products, *Israel J. Math.* **33** 198-224.

2. Antoine, J.-P. (1969), Dirac formalism and symmetry problems in quantum mechanics. I. General Dirac formalism, *J. Math. Phys.* **10** 53-69; II. Symmetry problems, *ibidem* 2276-2290.

3. Antoine, J.-P., G. Epifanio and C. Trapani (1983), Complete sets of unbounded observables, *Helv. Phys. Acta* **56** 1175-1186.

4. I. Antoniou (1988), *Internal Time and Irreversibility of Relativistic Dynamical Systems*, Thesis, Free University of Brussels.

5. Antoniou, I. (2001), The time operator of the cusp map, *Chaos Solitons and Fractals* **12** 1619-1627.

6. Antoniou, I., V. Basios and F. Bosco (1996), Probabilistic control of chaos: The β-adic Renyi map under control, *Int. J. Bifurcations and Chaos* **6** 1563-1573.

7. Antoniou, I., and Bi Qiao (1996), Spectral decomposition of the tent maps and the isomorphism of dynamical systems, *Physics Letters A* **215** 280-290.

8. Antoniou, I., Bi Qiao and Z. Suchanecki (1997), Generalized spectral decomposition and intrinsic irreversibility of the Arnold Cat map, *Chaos Solitons and Fractals* **8** 77-90.

9. Antoniou, I., L. Dmitrieva, Yu. Kuperin and Yu. Melnikov (1997), Resonances and the extension of dynamics to rigged Hilbert spaces, *Computers Math. Applic.* **34** 399-425.

10. Antoniou, I., M. Gadella, and Z. Suchanecki (1998), General properties of the Liouville operator, *Int. J. Theor. Phys.* **37** 1641-1654.

11. Antoniou, I., and K. Gustafson (1993), From probabilistic description to deterministic dynamics, *Physica* **A 197** 153-166.

12. Antoniou, I., and K. Gustafson (1997), From irreversible Markov semigroups to chaotic dynamics, *Physica* **A 236** 296-308.

13. Antoniou, I., and K. Gustafson (2000), The time operator of wavelets, *Chaos, Solitons and Fractals***11** 443-452.

14. Antoniou, I., K. Gustafson and Z. Suchanecki (1998), On the inverse problem of statistical physics: from irreversible semigroups to chaotic dynamics *Physica A* **252** 345-361.

15. Antoniou, I., Yu. Melnikov, S. Shkarin and Z. Suchanecki (2000), "Extended spectral decomposition of the Renyi map" *Chaos Solitons and Fractals* **11**, 393-421.

16. Antoniou, I., and B. Misra (1989), The relativistic internal time algebra, *Nuclear Physics, Proceed. Suppl. Sect.* **6** 240-242.

17. Antoniou, I., and B. Misra (1991), Characterization of semidirect sum Lie algebras, *J. Math. Phys.* **32** 864-868.

18. Antoniou, I., and B. Misra (1991), Non-unitary transformation of conservative to dissipative evolutions, *J. Phys. A: Math. Gen.* **24** 2723-2729.

19. Antoniou, I., and I. Prigogine (1989), *Intrinsic Irreversibility in Classical and Quantum Mechanics* in E. Bitsakis, C. Nicolaides eds. The Concept of Probability, Kluwer, Dordrecht pp. 223-234.

20. Antoniou, I., and I. Prigogine (1993), Intrinsic irreversibility and integrability of dynamics, *Physica* **A 192** 443-464.

21. Antoniou, I., I. Prigogine, V.A. Sadovnichii and S.A. Shkarin (2000), Time operator for diffusion, *Chaos, Solitons and Fractals* **11** 465-477.

22. Antoniou, I., V.A. Sadovnichii and S.A. Shkarin (1999), Time operators and shift representation of dynamical systems *Physica* **A 269** 299-313.

23. Antoniou, I., and S. Shkarin (2001), Decaying measures on locally compact Abelian topological groups *Proc. Royal Soc. Edinburgh* **131A** 1257-1273.

24. Antoniou, I., and S. Shkarin (2001), Decay spectrum and decay subspace of normal operators *Proc. Royal Soc. Edinburgh* **131A** 1245-1255.

25. Antoniou, I., S. Shkarin, and Z. Suchanecki (1999), The spectrum of the Liouville-von Neumann operator in the Hilbert-Schmidt space *J. Math. Phys.* **40** 4106-4118.

26. Antoniou, I., and Z. Suchanecki (1994), The logic of quantum systems with diagonal singularity, *Found. Phys.* **24** 1439-1457.

27. Antoniou, I., and Z. Suchanecki (1995), Quantum systems with diagonal singularity: observables, states, logic and intrinsic irreversibility, in *Nonlinear, Deformed and Irreversible Systems*, eds. H.-D. Doebner, V.K. Dobrev & P. Nattermann, World Scientific Publishing, pp. 22-52.

28. Antoniou, I., and Z. Suchanecki (1997), Quantum systems with diagonal singularity, *Adv. Chem. Phys.* **99** , ed. I. Prigogine and S. Rice, Wiley, pp. 299-332.

29. Antoniou, I., and Z. Suchanecki (1997), Logics of unstable systems, *Foundations of Physics* **27** 333-362.

30. Antoniou, I., and Z. Suchanecki (1998), Extension of the dynamics of unstable systems, *Infinite Dimensional Analysis, Quantum Probability and Rel. Topics* **1** 127-165.

31. Antoniou, I., and Z. Suchanecki (1999), Densities of singular measures and generalized spectral decomposition, *Generalized Functions, Operator Theory, and Dynamical Systems*, ed. I. Antoniou and G. Lumer, Chapman and Hall/CCR Research Notes in Mathematics **399**, pp. 56-67.

32. Antoniou, I., and Z. Suchanecki (2000), Non-uniform time operator, chaos and wavelets on the interval, *Chaos Solitons and Fractals* **11**, 423-435.

33. Antoniou, I., Z. Suchanecki, and S. Tasaki (1999), Series representations of the complex delta function, *Generalized Functions, Operator Theory, and Dynamical Systems*, ed. I. Antoniou and G. Lumer, Chapman and Hall/CCR Research Notes in Mathematics **399**, pp. 130-143.

34. Antoniou, I., Z. Suchanecki, R. Laura, and S. Tasaki (1997), Intrinsic irreversibility of quantum systems with diagonal singularity, *Physica A* **241** 737-772.

35. Antoniou, I., and S. Tasaki (1992), Generalized spectral decomposition of the β-adic baker's transformation and intrinsic irreversibility, *Physica* **A 190** 303-329.

36. Antoniou, I., and S. Tasaki (1993), Generalized spectral decomposition of mixing dynamical systems, *Int. J. Quantum Chemistry* **46** 425-474.

37. Antoniou, I., and S. Tasaki (1993), Spectral decomposition of the Renyi map, *J. Phys. A: Math.* Gen. **26** 73-94.

38. Arnold, V.I. (1978), *Mathematical Methods of Classical Mechanics*, Springer, Berlin.

39. Arnold, V.I., and A. Avez (1968), *Ergodic Problems of Classical Mechanics*, Benjamin, New York.

40. Ash, R.B., and M.F. Gardner (1975), *Topics in Stochastic Processes*, Academic Press Inc.

41. Atlee Jackson, E. (1989), *Perspectives of Nonlinear Dynamics*, Vol.1,2, Cambridge University Press.

42. Banach, S. (1932), *Théorie des Opérations Lineaires*, Monografje Matematyczne, Warszawa.

43. Bandtlow, O. F., I. Antoniou, and Z. Suchanecki, Resonances of dynamical systems and Fredholm-Riesz operators on rigged Hilbert spaces, *Computers Math. Applic.* **34** 95-102 (1997).

44. Barrow, J.D. (1982), Chaotic behaviour in general relativity, *Phys. Rep.* **85** 1-49.

45. Barut, A.O., and A. Wightman (1959), Relativistic invariance and quantum mechanics, *Nuovo Cim. Suppl.* **15** 81-91.

46. Beals, R. (1971), *Topics in Operator Theory*, The University of Chicago Press.

47. Berezanski, J. M. (1968), *Expansions in Eigenfunctions of Self Adjoint Operators*, Translations of Mathematical Monographs, **17**, AMS.

48. Billingsley, P. (1985), *Probability and Measure*, Wiley, New York.

49. Boas, R.P., and R.C. Buck (1964), *Polynomial Expansions of Analytic Functions*, Springer, Berlin.

50. Böhm, A. (1993), *Quantum Mechanics, Foundations and Applications* 3^{d}ed, Springer, Berlin.

51. Böhm, A., and M. Gadella (1989), *Dirac Kets, Gamow Vectors and Gelfand Triplets*, Springer Lect. Notes on Physics **348**, Berlin.

52. Bourbaki, N. (1987), *Topological Vector Spaces*, Springer Verlag, Berlin.

53. Braunss, G. (1985), Intrinsic stochasticity of dynamical systems, *Acta Applicandae Math.* **3** 1-21.

54. Brout, R., and I. Prigogine (1956), Statistical mechanics of irreversible processes, part VII: a general theory of weakly coupled systems, *Physica* **22** 621-636.

55. Brown, J.R. (1976), *Ergodic Theory and Topological Dynamics*, Academic Press, New York.

56. Chui, C.K. (1992), *An Introduction to Wavelets*, Academic Press, Inc., San Diego.

57. Ciesielski, Z. (1959), On Haar functions and the Schauder basis of the space $C_{\langle 0,1\rangle}$, *Bull. Acad. Polon. Sci., Sér. sci. math., astr. et phys.* **7** 227-232.

58. Ciesielski, Z. (1963), Properties of the orthonormal Franklin system, *Studia Math.* **23** 141-157.

59. Ciesielski, Z. (1966), Properties of the orthonormal Franklin system II, *Studia Math.* **27** 289-323.

60. Ciesielski, Z., and S. Kwapień(1979), Some properties of the Haar, Walsh–Paley, Franklin and the bounded polygonal orthonormal bases in L_p spaces, *Comment. Math.* Spec. Vol. **II**, Dedic. L. Orlicz, 37-42.

61. Cornfeld, I., S. Fomin and Ya. Sinai (1982), *Ergodic Theory*, Springer-Verlag, Berlin.

62. da Costa, N., and F. Doria (1991), Undecidability and incompleteness in classical mechanics, *Int. J. Theor. Phys.* **30** 1041.

63. Courbage, M. (1980), On necessary and sufficient conditions for the existence of time and entropy operators in quantum mechanics, *Lett. Math. Phys.* **4** 425-432.

64. Courbage, M. (1983), Intrinsic irreversibility of Kolmogorov dynamical systems, *Physica* **A 122** 459-482.

65. Courbage, M., and B. Misra (1980), On the equivalence between Bernoulli dynamical systems and stochastic Markov processes, *Physica* **A 104** 359-377.

66. Daubechies, I. (1992), *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

67. Davies, E.B. (1976), *Quantum Theory of Open Systems*, Academic Press, London.

68. E.B. Davies (1980), *One-Parameter Semigroups*, London Math. Soc. Monographs No.15, Academic Press.

69. Davies, P. (1974), *The Physics of Time Asymmetry*, Surrey Univ. Press, U.K.

70. Dellacherie, C., and P.A. Meyer (1978), *Probabilities and Potentiel A*, North-Holland, Amsterdam.

71. de la Llave, R. (1982), Rates of convergence to equilibrium in the Prigogine-Misra-Courbage theory of irreversibility, *J. Stat. Physics* **29** 17-31.

72. Diestel, J., and J. J. Uhl, Jr. (1977), *Vector Measures*, American Mathematical Society, Providence.

73. Dirac, P.A.M. (1958), *The Principles of Quantum Mechanics*, Oxford University Press, London.

74. Doob, J.L. (1953), *Stochastic Processes*, Wiley, New York.

75. Dörfle, M. (1985), Spectrum and eigenfunctions of the Frobenius-Perron operator of the tent map, *J. Stat. Phys.* **40** 93-103.

76. de Dormale, B., and H.-F. Gautrin (1975), Spectral representation and decomposition of self-adjoint operators, *J. Math. Phys.* **16** 2328-2332.

77. Dunford, N., and J. Schwartz (1988), *Linear Operators I : General Theory*, Wiley, New York.

78. Dunford, N., and J. Schwartz (1963), *Linear Operators II : Spectral Theory*, Wiley Interscience.

79. Duren, P.L. (1970), *Theory of H^p Spaces*, Academic Press, New York.

80. Dynkin, E. (1965), *Markov Processes* Springer, Berlin.

81. Eckmann , J.-P. (1989), *Resonances in dynamical systems* Proc. IX Int.Congr. on Mathematical Physics ed. B Simon, 192-207, Adam Hilder, Bristol.

82. Eckmann, J.-P., and D. Ruelle (1985), Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.* **57** 617-656.

83. Elskens, Y., and I. Prigogine (1986), From instability to irreversibility, *Proc. Natl. Acad. Sci.* **83** 5756-5760.

84. Emch, G. (1972), *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*, Wiley-Interscience.

85. Emch, G. (1976), Generalized K-flows, *Comm. Math. Phys.* **49** 191-215.

86. Exner, P. (1985), *Open Quantum Systems and Feynman Integrals*, Reidel, The Netherlands.

87. Falconer, K. (1990), *Fractal geometry: mathematical foundations and applications*, Chichester, Wiley.

88. Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, vol.I, Wiley, New York.

89. Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, vol.II, Wiley, New York.

90. Feynman, R.P. (1948), Space-time approach to non-relativistic quantum mechanics, *Rev. Mod. Phys.* **20** 367-387.

91. Floret, K., and J. Wloka (1968), *Einführung in die Theorie der lokalkonvexen Räume*, Lecture Notes in Mathematics 56, Berlin, Springer.

92. Foguel, S. (1969), *The Ergodic Theory of Markov Processes*, Van Nostrand, New York.

93. Fredricks, D. (1975) Tight riggings for a complete set of commuting observables, *Reports Math. Physics* **8** 277-293.

94. Gallavotti, G., and D. Ornstein (1974), Billiards and Bernoulli schemes, *Comm. Math Phys.* **38** 83-101.

95. GelfandI., and G. Shilov (1967), *Generalized Functions Vol.3: Theory of Differential Equations*, Academic Press, New York-San Francisco-London.

96. Gelfand, I., and N. Vilenkin (1964), *Generalized Functions Vol.4: Applications of Harmonic Analysis*, Academic Press, New York.

97. Gevorkjan, G.G. (1985), On the unboundedness of the shift operator for the Franklin system in L^1, *Mat. Zametki* **38** 523-533.

98. Glansdorff, P., and I. Prigogine (1971), *Thermodynamic Theory of Structure, Stability, and Fluctuations* Willey-Interscience New York.

99. Gleason, A.M. (1957), Measures on the closed subspaces of a Hilbert space, *J. Rat. Mech. and Analysis* **6** 885-893.

100. Goldstein, S. (1981), Entropy increase in dynamical systems, *Israel J. Math.* **38** 241-256.

101. Goldstein, S., B. Misra, and M. Courbage (1981), On intrinsic randomness of dynamical systems, *J. Stat. Phys.* **25** 111-126.

102. Goldstein, S., and O. Penrose (1981), A nonequilibrium entropy for dynamical systems, *J. Stat. Phys.* **24** 325-343.

103. Goodrich, K. K. Gustafson, and B. Misra (1980), On converse to Koopman's lemma, *Physica* **102A** 379-388.

104. Goodrich,R., K. Gustafson, and B. Misra (1986), On K-flows and irreversibility, *J. Stat. Phys.* **43** 317-320.

105. Halmos, P.R. (1956), *Lectures on Ergodic Theory*, Chelsea Publishing Company, New York.

106. Halmos, P.R. (1982), *A Hilbert Space Problem Book* (2nd edn), New York; Springer.

107. Hardin, Jr., C.D. (1981), Isometries on subspaces of L^p, *Indiana Univ. Math. J.* **30** 449-465.

108. Hartley, R.V. (1928), Transmission of information, *Bell System Technical Journal* **7** 535-563.

109. Hida, T. (1970), *Stationary Stochastic Processes*, Princeton University Press

110. Hida, T. (1980), *Brownian Motion*, Springer Verlag.

111. Hille, E., and R.S. Phillips (1968), *Functional Analysis and Semi-groups*, American Mathematical Society, Providence.

112. Hofbauer, F., and G. Keller (1982), Ergodic properties of invariant measures for piecewise monotonic transformations, *Math. Z.* **180** 119-140.

113. Hoffman, K. (1988), *Banach Spaces of Analytic Functions*, Dover Publications, New York.

114. Ionescu Tulcea, C.T. (1949), Mesures dans les espaces produits, *Lincei – Rend. Sc. fis. mat. e nat.* **7** 208-211.

115. Jauch, J.M. (1973), *Foundations of Quantum Mechanics* Addison-Wesley Pub. Co.

116. Jauch, J.M., and B. Misra (1965), The spectral representation, *Helv. Phys. Acta* **38** 30-52.

117. Jost, R. (1976), Measures on the finite dimensional subspaces of Hilbert space: Remarks to a theorem by A. M. Gleason in *Studies in Mathematical Physics*, eds. E.H. Lieb, B. Simon and A.S. Wightman, Princeton University, Princeton 1976, p.209-228.

118. Kaczmarz, S., and H. Steinhaus (1935), *Theorie der Orthogonalreihen*, Monografje Matematyczne, Warszawa.

119. Kakutani, S. (1961), Spectral analysis of stationary Gaussian processes, *Proc. 4th Berkeley Symp. Math. Stat. Probab.* **2** 239-247.

120. Kato, T. (1984), *Perturbation Theory for Linear Operators*, Second corrected printing of the second edition, Springer Verlag, Berlin.

121. Keller, G. (1984) On the rate of convergence to equilibrium in one-dimensional systems, *Comm. Math. Phys.* **96** 181-193.

122. Kern, M., R. Nagel and G. Palm (1977), Dilations of positive operators: Construction and ergodic theory, *Math. Z.* **156** 265-277.

123. Kolmogorov, A.N. (1958), A new metric invariant for transitive dynamical systems and automorphisms of Lebesgue space (Russian), *Dokl. Akad. Nauk SSSR* **119** 861-864.

124. Koopman, B. (1931), Hamiltonian systems and transformations in Hilbert spaces, *Proc. Nat. Acad. Sci. USA* **17** 315-318 .

125. Koopman, B., and J. von Neumann (1932), Dynamical systems of continuous spectra, *Proc. Nat. Acad. Sci. USA* **18** 255-266.

126. Kümmerer, B. (1985), Markov dilations on W^*-algebras, *J. Funct. Anal.* **63** 139-177.

127. Kwapień, S., and W.A. Woyczyński (1992), *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, Boston.

128. Lamperti, J. (1958), On the isometries of certain function spaces, *Pacific J. Math.* **8** 459-466.

129. Lasota, A., and M. Mackey (1994), *Chaos, Fractals, and Noise*, Springer-Verlag New York.

130. Lax, P.D., and R.S. Phillips (1989), *Scattering Theory* Academic Press, Boston.

131. Lewis, J. (1981), Quantum stochastic processes, *Phys. Reports* **77** 339-349.

132. Lin, M. (1971), Mixing for Markov operators, *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **19**, 231-242.

133. Lindbland, G. (1976), On the generators of quantum dynamical semigroups, *Comm. Math. Phys.* **48** 119-130.

134. Lindbland, G. (1979), Non-Markovian quantum stochastic processes and their entropy, *Comm. Math. Phys.* **65** 281-294.

135. Lindenstrauss, J., and L. Tzafriri (1977), *Classical Banach Spaces I and II*, Springer-Verlag.

136. Lockhart, C.M. (1981), *Time Operators in Classical and Quantum Systems*, Thesis, University of Texas at Austin.

137. Lockhart, C.M., and B. Misra (1986), Irreversibility and measurement in quantum mechanics, *Physica* **136A** 47-76.

138. Lockhart, C.M., B. Misra, and I. Prigogine (1982), Geodesic instability and internal time in relativistic cosmology, *Phys. Rev.* **D 25** 921-929.

139. Loève, M. (1977-78), *Probability Theory*, vol.I,II, Springer Verlag.

140. Louis, A.K., P. Maaß, and A. Rieder (1997), *Wavelets Theory and Applications*, John Wiley & Sons.

141. Mackey, G.W. (1963), *The Mathematical Foundations of Quantum Mechanics*, W.A. Benjamin, Inc.

142. Mandelbrot, B.B. (1982), *The Fractal Geometry of Nature* Freeman, San Francisco.

143. Martinez, S., and E. Tirapegui (1983), Markov processes and entropy associated to a reversible evolution, *Phys. Lett.* **95A** 143-145.

144. Martinez, S., and E. Tirapegui (1984), Description of a class of Markov processes "equivalent" to K-shifts, *J. Stat. Phys.* **37** 173-186.

145. Martinez, S., and E. Tirapegui (1985), A possible physical interpretation of the Λ operator in the Prigogine theory of irreversibility, *Phys. Lett.* **110A** 81-83.

146. Maurin, K. (1968), *General Eigenfunction Expansions and Unitary Representations of Topological Groups*, Polish Scientific Publishers, Warsaw.

147. Maurin, K. (1972), *Methods of Hilbert Spaces*, Polish Scientific Publishers, Warsaw.

148. Mejlbo, L.C. (1963), On the solution of the commutation relation $PQ - QP = -iI$, *Math. Scand.* **13** 129-139.

149. Melsheimer, O. (1974), Rigged Hilbert space formalism as extended mathematical formalism for quantum systems. I. General theory, *J. Math. Phys.* **15** 902-916; *ibid* II. Transformation theory in nonrelativistic quantum mechanics, 917-925.

150. Misra, B. (1978), Nonequilibrium entropy, Lyapunov variables, and ergodic properties of classical systems, *Proc. Natl. Acad. USA* **75** 1627-1631.

151. Misra, B. (1987), Fields as Kolmogorov flows, *J. Stat. Phys.* **48** 1295-1320.

152. Misra, B., and I. Prigogine (1980), On the foundations of kinetic theory, *Suppl. Progress of Theor. Phys.* **69** 101-110.

153. Misra, B., and I. Prigogine (1983), Time Probability and Dynamics, in *Long Time Predictions in Dynamical Systems*, ed. C. Horton, L. Reichl, V. Szebehely, Wiley, N.Y. 21-43.

154. Misra, B., and I. Prigogine (1983), Irreversibility and nonlocality, *Letters in Mathematical Physics* **7** 421-429.

155. Misra, B., I. Prigogine, and M. Courbage (1979), From deterministic dynamics to probabilistic descriptions, *Physica* **98 A** 1-26.

156. Misra, B., I. Prigogine, and M. Courbage (1979), From deterministic dynamics to probabilistic descriptions, *Proc. Natl. Acad. Sci. USA* **76** 3607-3611.

157. Misra, B., I. Prigogine, and M. Courbage (1979), Liapunov variable: entropy and measurements in quantum mechanics, *Proc. Natl. Acad. Sci. USA* **76** 4768-4772.

158. Misra, B., and E.C.G. Sudarshan(1976), The Zeno's paradox in quantum theory, *J. Math. Phys.* **18** 756-763 .

159. Naimark, M.A. (1972), *Normed Algebra*, Wolfer-Noordhoff, Groningen.

160. Narnhofer, H., and W. Thirring (1990), Algebraic K-systems, *Lett. Mat. Phys.* **20** 231-250 .

161. Narnhofer, H., M. Requardt, and W. Thirring (1983), Quasi particles at finite temperatures, *Comm. Math. Phys.* **92** 247-268.

162. Neveu ,J. (1965), *Mathematical Foundations of the Calculus of Probability*, Holden–Day Inc.

163. Nicolis, G., and I. Prigogine (1977), *Self-organization in Non-equilibrium Systems*, Wiley, New York.

164. Nicolis, G., and I. Prigogine (1989), *Exploring Complexity*, Freeman, New York.

165. Ornstein, D.S. (1970), Imbedding Bernoulli shifts in flows, *Springer Lecture Notes in Math.* **160** 178-218.

166. Ornstein, D.S. (1973), The isomorphism theorem for Bernoulli flows, *Adv. in Math* **10** 124-142.

167. Parry, W. (1981), *Topics in Ergodic Theory*, Cambridge University Press.

168. Pauli, W. (1980), *General Principles of Quantum Mechanics*, Springer-Verlag.

169. Petrosky, T., and I. Prigogine (1988), Poincaré's theorem and unitary transformations for classical and quantum systems, *Physica* **A147** 439-460.

170. Petrosky, T., and I. Prigogine (1991), Alternative formulation of classical and quantum dynamics for non-integrable systems, *Physica* **A175** 146-209.

171. Petrosky, T., and I. Prigogine (1993), Poincaré resonances and the limits of quantum mechanics, *Phys. Lett.* **A182** 5-15.

172. Petrosky, T., and I. Prigogine (1993), Poincaré resonances and the limits of trajectory dynamics, *Proc. Natl. Acad. Sci. USA* **90** 9393-9397.

173. Petrosky, T., and I. Prigogine (1994), Quantum chaos, complex spectral representations and time-symmetry breaking, *Chaos Solitons and Fractals* **4** 311-359.

174. Petrosky, T., and I. Prigogine (1996), Poincaré resonances and the extension of classical dynamics, *Chaos Solitons and Fractals* **7** 441-497.

175. Petrosky, T., and I. Prigogine (1996), Extension of classical dynamics. The case of anharmonic lattices, in *Gravity, Particles and Space-Time* , eds. P. Pronin, G. Sardanashvily, World Scientific, Singapore – New Jersey.

176. Petrosky, T., and I. Prigogine (1997), The Liouville space extension of quantum mechanics, *Adv. Chem. Phys.* **99** 1-120, ed. I. Prigogine and S. Rice, Wiley.

177. Petrosky, T., I. Prigogine, and S. Tasaki (1991), Quantum theory of nonintegrable systems, *Physica* **A173** 175-242.

178. Plesner, A. (1929), Eine Kennzeichnung der totalstetigen Funktionen, *J. für Math.* **160** 26-32.

179. Podgorski, K. (1989), Martingale approach to Boltzmann's entropy and exact transformations, in *Probability Theory on Vector Spaces*, S. Cambanis and A. Weron eds., Springer lecture Notes in Mathematics vol. 1391, Springer, Berlin, p. 321-328.

180. Poincaré, H. (1957), *Les Methodes Nouvelles de la Mécanique Céleste*, Vol. I (1892), Dover, New York.

181. Pollicott, M. (1990), The differential zeta function for axiom A attractors, *Ann. Math.* **131** 335-405.

182. Pour-el, M. and J. Richards (1989), *Computability in Analysis and Physics*, Springer, Berlin.

183. Prigogine, I. (1962), *Non-Equilibrium Statistical Mechanics*, Wiley, New York.

184. Prigogine, I. (1968), *Introduction to Thermodynamics of Irreversible Processes*, Interscience Publishers, New York.

185. Prigogine, I. (1973), The statistical interpretation of non-equilibrium entropy, *Acta Phys. Austriaca Suppl. X* 401-450, Proc. Int. Symp. "100 years Boltzmann eq." Vienna Sept. 1972.

186. Prigogine, I. (1980), *From Being to Becoming*, Freeman, New York.

187. Prigogine, I. (1992), Dissipative processes in quantum theory. With discussion, *Phys. Reports* **219** 93-120.

188. Prigogine, I., and M. Courbage (1983), Intrinsic randomness and intrinsic irreversibility in classical dynamical systems, *Proc. Natl. Acad. Sci. USA* **80** 2412-2416.

189. Prigogine, I., and C. George (1983), The second law as a selection principle: The microscopic theory of dissipative processes in quantum systems, *Proc. Natl. Acad. Sci. USA* **80** 4590-4594.

190. Prigogine, I., C George, and F. Henin (1969), Dynamical and statistical description of N-body systems, *Physica* **45** 418-434.

191. Prigogine, I., C George, F. Henin, and L. Rosenfeld (1973), A unified formulation of dynamics and thermodynamics, *Chemica Scripta* **4** 5-32.

192. Prigogine, I., and T. Petrosky (1987), Intrinsic irreversibility in quantum theory, *Physica* **147 A** 33-47.

193. Prigogine, I., and P. Resibois (1961), On the kinetic of the approach to equilibrium, *Physica* **27** 629-646.

194. Prigogine I., and I. Stengers (1984), *Order out of Chaos – Man's New Dialogue with Nature*, Bantam Books, Toronto.

195. Prugovecki, E. (1981), *Quantum Mechanics in Hilbert Space*, Academic Press, New York.

196. Putnam, C.R. (1967), *Commutation Properties of Hilbert Space Operators and Related Topics*, Springer Verlag.

197. Radjavi, H., and P. Rosenthal (1973), *Invariant Subspaces*, Springer-Verlag, Berlin, New York.

198. Reed, M., and B. Simon (1980), *Methods of Modern Mathematical Physics Vol I: Functional Analysis*, Academic Press, New York.

199. Reed, M., and B. Simon (1975), *Methods of Modern Mathematical Physics Vol II: Fourier Analysis, Self-Adjointness*, Academic Press, London.

200. Reed, M., and B. Simon (1979), *Methods of Modern Mathematical Physics Vol III: Scattering Theory*, Academic Press, London.

201. Reed, M., and B. Simon (1978), *Methods of Modern Mathematical Physics Vol IV: Analysis of Operators*, Academic Press, London.

202. Rényi, A. (1957), Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hungary* **8** 477-493.

203. Riesz, F., and B. Sz.-Nagy (1972), *Leçons d'Analyse Fonctionnelle* Académiai Kiadó, Budapest.

204. Rokhlin, V. (1964), Exact endomorphisms of Lebesgue spaces, *Am. Math. Soc. Transl.* **39** 1-3.

205. Rozanov, Y. (1977), *Innovation Processes*, Wiley New York.

206. Ruelle, D. (1986), Resonances of chaotic dynamical systems, *Phys. Rev. Lett.* **56** 405-407.

207. Ruelle, D. (1989), The thermodynamic formalism for expanding maps, *Commun. Math. Phys.* **125** 239-262.

208. Ruelle, D. (1990), An extension of the theory of Fredholm determinants, *Publ. Math. IHES* **72** 175-193.

209. Rybaczuk, M., K. Weron, Z. Suchanecki, and A. Weron (1990), Kinetic equation in Prigogine's theory of irreversibility, in *Stochastic Methods in Math. and Physics*, p. 267-274, eds. R. Gierelak, W. Karwowski - World Scientific, Singapore.

210. Samorodnitsky, G., and M.S. Taqqu (1994), *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall.

211. Schaefer, H.H. (1970), *Topological Vector Spaces*, Springer Verlag.

212. Schipp, F., W.R. Wade, P. Simon, and J. Pal (1990), *Walsh Series – an Introduction to Dyadic Harmonic Analysis,* Adam Hilger, Bristol and New York.

213. Schwartz, L. (1950-51), *Theorie des Distributions*, vol.I, II, Hermann, Paris.

214. Schwartz, L. (1973), *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*, Oxford University Press.

215. Semadeni, Z. (1982), *Schauder Bases in Banach Spaces of Continuous Functions*, Lecture Notes in Math. **918**, Springer-Verlag.

216. Shannon, C.E. (1948), A mathematical theory of communication, *Bell Sys. Tech. J.* **27**, 379-423, 623-653.

217. Shannon, C.E., and W. Weaver (1949), *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, Ill.

218. Shields, P. (1973), *The Theory of Bernoulli Shifts*, Chicago Univ. Press.

219. Sinha, K. (1977), *Scattering Theory in Quantum Mechanics* Benjamin, Massachusetts. .

220. Suchanecki, Z. (1989), An L^1 extension of stochastic dynamics for irreversible systems, *Lecture Notes in Math.* No.**1391** 367-374.

221. Suchanecki, Z. (1991), Exactness an irreversibility in classical and quantum dynamical systems, *Quantum Probab. and Rel. Topics* **6** 437-451.

222. Suchanecki, Z. (1992a), Entropy and asymptotic stability of stochastic operators, *Europhysics Letters* **17** 93-96.

223. Suchanecki, Z. (1992b), On lambda and internal time operators, *Physica* **A 187** 249-266.

224. Suchanecki, Z., I. Antoniou and S. Tasaki (1994), Nonlocality of the Misra-Prigogine-Courbage Semigroup, *J. Stat. Physics* **75** 919-928.

225. Suchanecki, Z., I. Antoniou, S. Tasaki, and O.F. Bandtlow (1996), Rigged Hilbert spaces for chaotic dynamical systems, *J. Math. Phys.* **37** 5837-5847.

226. Suchanecki, Z., and A. Weron (1984), Decomposability of cylindrical martingales and absolutely summing operators, *Math. Z.* **185** 271-280.

227. Suchanecki, Z., and A. Weron (1990a), Applications of an operator stochastic integral in Prigogine's theory of irreversible dynamical systems, *Expo. Math.* **8** 67-79.

228. Suchanecki, Z., and A. Weron (1990b), Characterization of intrinsically random dynamical systems, *Physica* **A 166** 220-228.

229. Suchanecki, Z., A. Weron, M. Rybaczuk, and K. Weron (1989), An explicit approach to the Λ-operator and H-Theorem in the Prigogine's theory of irreversibility, *Physica* **A 155** 385-400.

230. Sudarshan, E.C.G., C. Chiu, and V. Gorini (1978), Decaying states as complex energy eigenvectors in generalized quantum mechanics, *Phys. Rev.* **D18** 2914-2929.

231. Sz-Nagy, B., and C Foias (1970), *Harmonic Analysis of Operators in Hilbert Space*, North-Holland, Amsterdam.

232. Takesaki, M. (1979), *Theory of Operator algebras I*, Springer, Berlin.

233. Tasaki, S., I. Antoniou, and Z. Suchanecki (1994), Spectral decomposition and fractal eigenvectors for a class of piecewise linear maps, *Chaos Solitons and Fractals* **4** 227-254.

234. Tasaki S., I. Antoniou, and Z. Suchanecki (1993), Deterministic diffusion, De Rham equation and fractal eigenvectors, *Physics Letters A* **179** 97-102.

235. Tasaki, S., Z. Suchanecki, and I. Antoniou (1993), Ergodic properties of piecewise linear maps on fractal repellers, *Physics Letters A* **179** 103-110.

236. Treves, F. (1967), *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York.

237. van Hove, L. (1955), Energy corrections and persistent perturbation effects in continuous spectra, *Physica* **21** 901-923.

238. van Hove, L. (1956), Energy corrections and persistent perturbation effects in continuous spectra. II. The perturbed stationary states, *Physica* **22** 343-354.

239. van Hove, L. (1957), The approach to equilibrium in quantum statistics, *Physica* **23** 441-480.

240. van Hove, L. (1959), The ergodic behaviour of quantum many-body systems, *Physica* **25** 268-276.

241. van Kampen, N. (1981), *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam.

242. van Winter, C. (1971), Fredholm equations on a Hilbert space of analytic functions, *Trans. Amer. Math. Soc.* **162** 103-139.

243. Varadarajan, V.S. (1961), Measures on Topological Spaces, *Mat. Sbornik* **55** 35-103.

244. von Neumann, J. (1955), *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton.

245. Wehrl, A. (1978), General properties of entropy, *Rev. Mod. Phys.* **50** 221-260.

246. Weidmann, J. (1980), *Linear Operators in Hilbert Spaces*, Springer-Verlag.

247. Whittaker, E. (1937), *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed. Cambridge Univ. Press, London.

248. Wiener, N. (1938), The homogeneous chaos, *Amer. J. Math.* **60** 897-936.

249. Wiener, N. (1948), *Cybernetics, or Control and Communication in the Animal and the Machine*, John Wiley, New York 1948.

250. Wiener, N., and E.J. Akutowicz (1957), The definition and ergodic properties of the stochastic adjoint of a unitary transformation, *Rend. Circ. Mat. Palermo (2)* **6** 205-217.

251. Wigner, E.P. (1939), Unitary representations of the inhomogenous Lorenz group, Ann. Math., **40**, 149-204.

252. Wigner, E.P. (1956), Relativistic invariance in quantum mechanics, *Nuovo Cim.* **3** 517-532.

253. Wigner, E.P. (1959), *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York.

254. Yosida, K. (1980), *Functional Analysis* 6th ed. Springer, Berlin.

255. Yosida, K., and S. Kakutani (1941), Operator-theoretical treatment of Markoff process and mean ergodic theorem, *Annals of Mathematics* **42** 188-229.

256. Zwanziger, D. (1963), Representation of the Lorentz group corresponding to unstable particles, *Phys. Rev.* **131** 2818-2819.

257. Zygmund, A. (1968), *Trigonometric Series* vol. I,II, Cambridge, University Press.