SPRINGER BRIEFS IN COMPLEXITY

Eric Bertin

A Concise Introduction to the Statistical Physics of Complex Systems

Springer

SpringerBriefs in Complexity

For further volumes:
http://www.springer.com/series/8907

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "SpringerBriefs in Complexity" which are concise and topical working reports, case-studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Eric Bertin

A Concise Introduction to the Statistical Physics of Complex Systems

Springer

Eric Bertin
Laboratoire de Physique
Ecole Normale Supérieure de Lyon
46 allée d'Italie
69007 Lyon
France
e-mail: eric.bertin@ens-lyon.fr

ISSN 2191-5326 e-ISSN 2191-5334
ISBN 978-3-642-23922-9 e-ISBN 978-3-642-23923-6
DOI 10.1007/978-3-642-23923-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011937256

Cover design: eStudio Calamar, Berlin/Figueres

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In recent years, statistical physics started raising the interest of a broad community of researcher in the field of complex system sciences, ranging from biology to social sciences, economics or computer sciences. More generally, a growing number of graduate students and researchers feel the need for learning some basic concepts and questions coming from other disciplines, leading for instance to the organization of recurrent interdisciplinary summer schools.

The present booklet is partly based on the introductory lecture on statistical physics given at the French Summer School on Complex Systems held both in Lyon and Paris during the summers 2008 and 2009, and jointly organized by two French Complex Systems Institutes, the "Institut des Systèmes Complexes Paris Ile de France" (ISC-PIF) and the "Institut Rhône-Alpin des Systèmes Complexes" (IXXI). This introductory lecture was aimed at providing the participants with a basic knowledge of the concepts and methods of statistical physics so that they could later on follow more advanced lectures on diverse topics in the field of complex systems. The lecture has been further extended in the framework of the second year of Master in "Complex Systems Modelling" of the Ecole Normale Supérieure de Lyon and Université Lyon 1, whose courses take place at IXXI.

It is a pleasure to thank Guillaume Beslon, Tommaso Roscilde and Sébastian Grauwin, who were also involved in some of the lectures mentioned above, as well as Pablo Jensen for his efforts in setting up an interdisciplinary Master course on complex systems, and for the fruitful collaboration we had over the last years.

Contents

1 Equilibrium Statistical Physics 1
1.1 Microscopic Dynamics of a Physical System 1
1.1.1 A Particle Attached to a Spring 1
1.1.2 Many-Particle System 5
1.1.3 Case of Discrete Variables: Spin Models 5
1.2 Statistical Description of an Isolated System at Equilibrium 6
1.2.1 Notion of Statistical Description: A Toy Model 6
1.2.2 Fundamental Postulate of Equilibrium Statistical Physics 7
1.2.3 Computation of $\Omega(E)$ and $S(E)$: Some Simple Examples 8
1.2.4 Distribution of Energy Over Subsystems and Statistical Temperature 10
1.3 Equilibrium System in Contact with its Environment. 12
1.3.1 Exchanges of Energy. 12
1.3.2 Canonical Entropy. 15
1.3.3 Exchanges of Particles with a Reservoir: The Grand-Canonical Ensemble 17
1.4 Phase Transitions. 18
1.4.1 Example of the Ising Model. 18
1.4.2 Ising Model in Fully Connected Geometry. 18
1.4.3 Ising Model with Finite Connectivity 21
1.4.4 Renormalization Group Approach: A Brief Introduction 22
1.5 Disordered Systems and Glass Transition 27
1.5.1 Disorder in Complex Systems: From Social Sciences to Spin-Glasses. 27

1.5.2 Theoretical Spin-Glass Models 28
1.5.3 The Simplest Disordered System: The Random Energy Model 28
References 31

2 Non-Stationary Dynamics and Stochastic Formalism 33
2.1 Markovian Stochastic Processes and Master Equation 33
2.1.1 Definition of Markovian Stochastic Processes 33
2.1.2 Master Equation and Detailed Balance 35
2.1.3 Dynamical Increase of the Entropy 37
2.1.4 A Simple Example: The One-Dimensional Random Walk 38
2.2 Langevin and Fokker–Planck Equations 41
2.2.1 Phenomenological Approach to the Langevin Equation 41
2.2.2 Relation to Random Walks 45
2.2.3 Fokker–Planck Equation 47
2.3 Anomalous Diffusion and Physical Aging 49
2.3.1 Generalized Central Limit Theorem 49
2.3.2 Anomalous Diffusion 52
2.3.3 Aging Dynamics and Trap Models 54
References 57

3 Statistical Physics of Interacting Macroscopic "Entities" 59
3.1 Dynamics of Residential Moves 60
3.1.1 A Simplified Version of the Schelling Model 61
3.1.2 Equilibrium Configurations of the Model 61
3.1.3 Condition for Phase Separation 63
3.2 Condensation Transition 65
3.2.1 Zero Range Process: Definition and Exact Solution 66
3.2.2 Maximal Density and Condensation Phenomenon 66
3.3 Synchronization Transition 67
3.3.1 The Kuramoto Model of Coupled Oscillators 68
3.3.2 Synchronized Steady State 70
3.4 Collective Motion of Active Particles 72
3.4.1 Definition of the Model 73
3.4.2 Description Through a Boltzmann Equation 73
3.4.3 Hydrodynamic Equations and Phase Diagram 74
References 77

Introduction

Generally speaking, the goals of statistical physics may be summarized as follows: on the one hand to study systems composed of a large number of interacting 'entities', and on the other hand to predict the macroscopic (or collective) behavior of the system considered from the microscopic laws ruling the dynamics of the individual 'entities'. These two goals are, to some extent, also shared by what is nowadays called 'complex systems science'. However, the specificity of statistical physics is that:

- The 'entities' considered are in most cases atoms or molecules, for which the individual microscopic laws are known from fundamental physical theories—at variance with other fields like social sciences for example, where little is known about the quantitative behavior of individuals.
- These atoms, or molecules, are often all of the same type, or at most of a few different types—in contrast to biological or social systems for instance, where the individual 'entities' may all differ, or at least belong to a large number of different types.

For these reasons, systems studied in the framework of statistical physics may be considered as among the simplest examples of complex systems. One further specificity of statistical physics with respect to other sciences aiming at describing the collective behavior of complex systems is that it allows for a rather well-developed mathematical treatment.

The present booklet is divided into three chapters. The first one deals with equilibrium statistical physics, trying to expose in a concise way the main concepts of this theory, and paying specific attention to those concepts that could be more generally relevant to complex system sciences. The second part mainly aims at describing time-dependent effects (occurring for instance when a system relaxes to equilibrium from some non-equilibrium initial condition) in the framework of stochastic markovian processes. Emphasis is put on simple and generic models, and some relations to probability theory are also outlined. Finally, the third part presents a few examples of applications of statistical physics to other types of complex systems, beyond the strict realm of physics, with the hope to trigger the interest of readers coming from various disciplinary fields. Simple models of systems composed of a large number of macroscopic ‘entities’ that do not follow the same laws as physical atoms or molecules (for instance sand grains, animals or social agents) are considered. As no general statistical framework exists for such systems, their description relies on the case-by-case adaptation of different techniques borrowed from standard statistical physics, ranging from mappings to effective equilibrium systems, to Boltzmann approaches (a technique early developed in statistical physics to characterize the dynamics of gases) for systems interacting through binary collisions, or to exact solutions when available.

Chapter 1
Equilibrium Statistical Physics

Systems composed of many particles involve a very large number of degrees of freedom, and it is most often uninteresting—not to say hopeless—to try to describe in a detailed way the microscopic state of the system. The aim of statistical physics is rather to restrict the description of the system to a few relevant macroscopic observables, and to predict the average values of these observables, or the relations between them. A standard formalism, called "equilibrium statistical physics", has been developed for systems of physical particles having reached a statistical steady state in the absence of external driving (like heat flux or shearing forces for instance).

In this first part, we shall discuss some of the fundamentals of equilibrium statistical physics. Section 1.1 describes the elementary mechanical notions necessary to describe a system of physical particles. Section 1.2 introduces the basic statistical notions and fundamental postulates required to describe in a statistical way a system that exchanges no energy with its environment. The effect of the environment is then taken into account in Sect. 1.3, in the case where the environment does not generate any sustained energy flux into the system. Applications of this general formalism to the description of collective phenomena and phase transitions are presented in Sect. 1.4. Finally, the influence of disorder and heterogeneities, which are relevant in physical systems, but are also expected to play an essential role in many other types of complex systems, is briefly discussed in Sect. 1.5. For further reading on these topics related to equilibrium statistical physics (especially for Sects. 1.2–1.4), we refer the reader to standard textbooks, like e.g. Refs. [1–4].

1.1 Microscopic Dynamics of a Physical System

1.1.1 A Particle Attached to a Spring

As an elementary example, consider a particle constrained to move on a one-dimensional horizontal axis x, and attached to a spring, the latter being pinned to

E. Bertin, *A Concise Introduction to the Statistical Physics of Complex Systems*,
SpringerBriefs in Complexity, DOI: 10.1007/978-3-642-23923-6_1,

a rigid wall. We consider the position $x(t)$ of the particle at time t, as well as its velocity $v(t)$. The force F exerted by the spring on the particle is given by

$$F = -k(x - x_0), \tag{1.1}$$

where x_0 corresponds to the position of repose of the particle, for which the force vanishes. For convenience, we shall in the following choose the origin of the x axis such that $x_0 = 0$.

From the basic laws of classical mechanics,[1] the motion of the particle is described by the evolution equation:

$$m\frac{dv}{dt} = F \tag{1.2}$$

where m is the mass of the particle. We have neglected all friction forces, so that the force exerted by the spring is the only horizontal force (the gravity force, as well as the reaction force exerted by the support, do not have horizontal components in the absence of friction). In terms of x variable, the equation of motion (1.2) reads

$$m\frac{d^2x}{dt^2} = -kx. \tag{1.3}$$

The generic solution of this equation is

$$x(t) = A\cos(\omega t + \phi), \quad \omega = \sqrt{\frac{k}{m}}. \tag{1.4}$$

The constants A and ϕ are determined by the initial conditions.

Hamiltonian Reformulation of the Problem

Let us introduce the momentum $p = mv$, and the kinetic energy $E_c = \frac{1}{2}mv^2$. In terms of momentum, the kinetic energy reads $E_c = p^2/2m$. The potential energy U of the spring, defined by $F = -dU/dx$, is given by $U = \frac{1}{2}kx^2$. The Hamiltonian $H(x, p)$ is defined as

$$H(x, p) = E_c(p) + U(x). \tag{1.5}$$

In the present case, this definition yields

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2}kx^2. \tag{1.6}$$

---

[1] The fundamental theory describing the dynamics of particles at the atomic scale is actually quantum mechanics rather than classical mechanics. However, classical mechanics is in many cases of interest a reasonable approximation. We shall thus remain in the framework of classical mechanics for the purpose of the present booklet.

In the Hamiltonian formulation, the equations of motion read[2]

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x}. \tag{1.7}$$

On the example of the particle attached to a spring, these equations give

$$\frac{dx}{dt} = \frac{p}{m}, \quad \frac{dp}{dt} = -kx, \tag{1.8}$$

from which one recovers Eq. 1.3 by eliminating p. Hence it is seen on the above example that the Hamiltonian formalism is equivalent to the standard law of motion (1.2).

The Hamiltonian formulation has interesting properties, namely energy conservation and time-reversal invariance. We define the total energy $E(t)$ as $E(t) = H(x(t), p(t)) = E_c(p(t)) + U(x(t))$. It is easily shown that the total energy is conserved during the evolution of the system[3]

$$\frac{dE}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p}\frac{dp}{dt}. \tag{1.9}$$

Using Eq. 1.7, one has

$$\frac{dE}{dt} = \frac{\partial H}{\partial x}\frac{\partial H}{\partial p} + \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial x}\right) = 0, \tag{1.10}$$

so that the energy E is conserved. This is confirmed by a direct calculation on the example of the particle attached to a spring:

$$\begin{aligned} E(t) &= \frac{p(t)^2}{2m} + \frac{1}{2}kx(t)^2 \\ &= \frac{1}{2m}m^2\omega^2 A^2 \sin^2(\omega t + \phi) + \frac{1}{2}kA^2\cos^2(\omega t + \phi). \end{aligned} \tag{1.11}$$

Given that $\omega^2 = k/m$, one finds

$$E(t) = \frac{1}{2}kA^2\left(\sin^2(\omega t + \phi) + \cos^2(\omega t + \phi)\right) = \frac{1}{2}kA^2 \tag{1.12}$$

which is indeed a constant.

---

[2] For a more detailed introduction to the Hamiltonian formalism, see, e.g., Ref. [5].

[3] The concept of energy, introduced here on a specific example, plays a fundamental role in physics. Though any precise definition of the energy is necessarily formal and abstract, the notion of energy can be thought of intuitively as a quantity that can take very different forms (kinetic, electromagnetic or gravitational energy, but also internal energy exchanged through heat transfers) in such a way that the total amount of energy remains constant. Hence an important issue is to describe how energy is transferred from one form to another. For instance, in the case of the particle attached to a spring, the kinetic energy E_c and potential energy U of the spring are continuously exchanged, in a reversible manner. In the presence of friction forces, kinetic energy would also be progressively converted, in an irreversible way, into internal energy, thus raising the temperature of the system.

Phase-Space Representation

Physical space is described in the above example by the coordinate x. The equations of motion (1.7) allow the position and momentum of the particle to be determined at any time once the initial position and momentum are known. So it is interesting to introduce an abstract representation space containing both position and momentum. In this example, it is a two-dimensional space, but it could be of higher dimension in more general situations. This representation space is often called “phase space”. For the particle attached to the spring, the trajectories in this phase space are ellipses. Rescaling the coordinates in an appropriate way, one can transform the ellipse into a circle, and the energy can be identified with the square of the radius of the circle. To illustrate this property, let us define the new phase-space coordinates X and Y as

$$X = \sqrt{\frac{k}{2}}\,x, \quad Y = \frac{p}{\sqrt{2m}}. \tag{1.13}$$

Then the energy E can be written as

$$E = \frac{p^2}{2m} + \frac{1}{2}kx^2 = X^2 + Y^2. \tag{1.14}$$

As the energy is fixed, the trajectory of the particle is a circle of radius $\sqrt{E}$ in the (X,Y)-plane.

Time Reversibility

To illustrate the meaning of time reversibility, let us imagine that we film the motion of the particle with a camera, and that we project it backward. If the backward motion is also a possible motion, meaning that nothing is unphysical in the backward projected movie, then the equations of motion are time-reversible.

More formally, we consider the trajectory $x(t)$, $t = 0, \ldots, t_0$, and define the reversed time $t' = t_0 - t$. Starting from the equations of motion (1.7) expressed with t, x and p, time reversal is implemented by replacing t with $t_0 - t'$, x with x' and p with $-p'$, yielding

$$-\frac{dx}{dt'} = -\frac{\partial H}{\partial p'}, \quad \frac{dp'}{dt'} = -\frac{\partial H}{\partial x'}. \tag{1.15}$$

Changing the overall sign in the first equation, one recovers Eq. 1.7 for the primed variables, meaning that the time-reversed trajectory is also a physical trajectory.

Note that time-reversibility holds only as long as friction forces are neglected. The latter break time reversal invariance, and this explains why our everyday-life experience seems to contradict time reversal invariance. For instance, when a glass falls down onto the floor and breaks into pieces, it is hard to believe that the reverse trajectory, in which pieces would come together and the glass would jump onto the table, is also a possible trajectory, as nobody has ever seen this phenomenon occur. In order to reconcile macroscopic irreversibility and microscopic reversibility of trajectories, the point of view of statistical physics is to consider that the reverse trajectory is possible, but has a very small probability to occur as only very few initial conditions could lead to this trajectory. So in practice, the corresponding trajectory is never observed.

1.1.2 Many-Particle System

In a more general situation, a physical system is composed of N particles in a three-dimensional space. The position of particle i is described by a vector $\mathbf{x}_i$, and its velocity by $\mathbf{v}_i$, $i = 1, \dots, N$. In the Hamiltonian formalism, it is often convenient to introduce generalized coordinates q_j and momenta p_j which are scalar quantities, with $j = 1, \dots, 3N$: (q_1, q_2, q_3) are the components of the vector $\mathbf{x}_1$ describing the position of particle 1, (q_4, q_5, q_6) are the component of $\mathbf{x}_2$, and so on. Similarly, (p_1, p_2, p_3) are the components of the momentum vector $m\mathbf{v}_1$ of particle 1, (p_4, p_5, p_6) are the components of $m\mathbf{v}_2$, etc. With these notations, the Hamiltonian of the N-particle system is defined as

$$H(q_1, \dots, q_{3N}, p_1, \dots, p_{3N}) = \sum_{j=1}^{3N} \frac{p_j^2}{2m} + U(q_1, \dots, q_{3N}). \tag{1.16}$$

The first term in the Hamiltonian is the kinetic energy, and the last one is the potential (or interaction) energy. The equations of motion read

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \quad j = 1, \dots, 3N. \tag{1.17}$$

The properties of energy conservation and time-reversal invariance also hold in this more general formulation, and are derived in the same way as above. As an illustration, typical examples of interaction energy U include

- $U = 0$: case of free particles.
- $U = -\sum_{i=1}^{N} \mathbf{h}_i \mathbf{x}_i$: particles interacting with an external field, for instance the gravity field, or an electric field.
- $U = \sum_{i \neq i'} V(\mathbf{x}_i - \mathbf{x}_{i'})$: pair interaction potential.

1.1.3 Case of Discrete Variables: Spin Models

As a simplified picture, a spin may be thought of as a magnetization $\mathbf{S}$ associated to an atom. The dynamics of spins is ruled by quantum mechanics (the theory that governs particles at the atomic scale), which is far beyond the scope of the present lecture. However, in some situations, the configuration of a spin system can be represented in a simplified way as a set of binary "spin variables" $s_i = \pm 1$, and the corresponding energy takes the form

$$E = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_{i=1}^{N} s_i . \tag{1.18}$$

The parameter J is the coupling constant between spins, while h is the external magnetic field. The first sum corresponds to a sum over nearest neighbor sites on a lattice, but other types of interaction could be considered. This model is called the Ising model. It provides a qualitative description of the phenomenon of ferromagnetism observed in metals like iron, in which a spontaneous macroscopic magnetization appears below a certain critical temperature. In addition, the Ising model turns out to be very useful to illustrate some important concepts of statistical physics.

In what follows, we shall consider the words "energy" and "Hamiltonian" as synonyms, and the corresponding notations E and H as equivalent.

1.2 Statistical Description of an Isolated System at Equilibrium

1.2.1 Notion of Statistical Description: A Toy Model

Let us consider a toy model in which a particle is moving on a ring with L sites. Time is discretized, meaning that for instance every second the particle moves to the next site. The motion is purely deterministic: given the position at time $t = 0$, one can compute the position $i(t)$ at any later time. Now assume that there is an observable ε_i on each site i. It could be for instance the height of the site, or any arbitrary observable that characterizes the state of the particle when it is at site i.

A natural question would be to know what the average value

$$\langle \varepsilon \rangle = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{i(t)} \tag{1.19}$$

is after a large observation time T. Two different answers to this question could be given:

- Simulate the dynamics of the model on a computer, and measure directly $\langle \varepsilon \rangle$.
- Use the concept of probability as a shortcut, and write

$$\langle \varepsilon \rangle = \sum_{i=1}^{L} P_i \varepsilon_i \tag{1.20}$$

where the probability P_i to be on site i is defined as

$$P_i = \frac{\text{time spent on site } i}{\text{total time } T}, \tag{1.21}$$

namely the fraction of time spent on site i.

The probability P_i can be calculated or measured by simulating the dynamics, but it can also be estimated directly: if the particle has turned a lot of times around the ring, the fraction of time spent on each site is the same, $P_i = 1/L$. Hence all positions of the particle are equiprobable, and the average value $\langle \varepsilon \rangle$ is obtained as a flat average over all sites.

1.2.2 Fundamental Postulate of Equilibrium Statistical Physics

We consider a physical system composed of N particles. The microscopic configuration of the system is described by $(\mathbf{x}_i, \mathbf{p}_i = m\mathbf{v}_i)$, $i = 1, \ldots, N$, or $s_i = \pm 1$, $i = 1, \ldots, N$, for spin systems.

The total energy E of the system, given for systems of particles by

$$E = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + U(\mathbf{x}_1, \ldots, \mathbf{x}_N), \tag{1.22}$$

or for spins systems by

$$E = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_{i=1}^{N} s_i, \tag{1.23}$$

is constant as a function of time (or may vary within a tiny interval $[E, E + \delta E]$, in particular for spin systems). Accordingly, starting from an initial condition with energy E, the system can only visit configurations with the same energy. In the absence of further information, it is legitimate to postulate that all configurations with the same energy as the initial one have the same probability to be visited. This leads us to the *fundamental postulate of equilibrium statistical physics*:

All configurations with a given energy E have the same probability. Other configurations have zero probability.

The corresponding probability distribution is called the microcanonical distribution or microcanonical ensemble for historical reasons (a probability distribution can be thought of as describing an infinite set of copies—an ensemble—of a given system).

A quantity that plays an important role is the “volume” $\Omega(E)$ occupied in phase-space by all configurations with energy E. For systems with continuous degrees of freedom, $\Omega(E)$ is the area of the hypersurface defined by fixing the energy E. For systems with discrete configurations (spins), $\Omega(E)$ is the number of configurations with energy E. The Boltzmann entropy is defined as

$$S(E) = k_B \ln \Omega(E), \tag{1.24}$$

where $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant. This constant has been introduced both for historical and practical reasons, but from a theoretical viewpoint, its specific value plays no role, so that we shall set it to $k_B = 1$ in the following (this could be done for instance by working with specific units of temperature and energy such that $k_B = 1$ in these units).

The notion of entropy is a cornerstone of statistical physics. First introduced in the context of thermodynamics (the theory of the balance between mechanical energy transfers and heat exchanges), entropy was later on given a microscopic interpretation in the framework of statistical physics. Basically, entropy is a measure of the number of available microscopic configurations compatible with the macroscopic constraints. More intuitively, entropy can be interpreted as a measure of "disorder" (disordered macroscopic states often correspond to a larger number of microscopic configurations than macroscopically ordered states), though the correspondence between the two notions is not necessarily straightforward. Another popular interpretation, in relation to information theory, is to consider entropy as a measure of the lack of information on the system: the larger the number of accessible microscopic configurations, the less information is available on the system (in an extreme case, if the system can be with equal probability in any microscopic configuration, one has no information on the state of the system).

Let us now give a few simple examples of computation of the entropy.

1.2.3 Computation of $\Omega(E)$ and $S(E)$: Some Simple Examples

Paramagnetic Spin Model

We consider a model of independent spins, interacting only with a uniform external field. The corresponding energy is given by

$$E = -h \sum_{i=1}^{N} s_i, \quad s_i = \pm 1. \tag{1.25}$$

The phase space (or here simply configuration space) is given by the set of values $\{s_i\}_{i=1,\dots,N}$. The question is to know how many configurations there are with a given energy E. In this specific example, it is easily seen that fixing the energy E amounts to fixing the magnetization $M = \sum_{i=1}^{N} s_i$. Let us denote as N_+ the number of spins with value $+1$ ("up" spins). The magnetization is given by $M = N_+ - (N - N_+) = 2N_+ - N$, so that fixing M is in turn equivalent to fixing N_+. From basic combinatorial arguments, the number of configurations with a given number of "up" spins is given by

$$\Omega = \frac{N!}{N_+!(N - N_+)!}. \tag{1.26}$$

Using the relation

$$N_+ = \frac{1}{2}\left(N - \frac{E}{h}\right), \tag{1.27}$$

one can express Ω as a function of E:

$$\Omega(E) = \frac{N!}{\left[\frac{1}{2}(N - E/h)\right]!\left[\frac{1}{2}(N + E/h)\right]!}. \tag{1.28}$$

The entropy *S(E)* is given by

$$\begin{aligned} S(E) &= \ln \Omega(E) \\ &= \ln N! - \ln \left[\frac{1}{2}\left(N - \frac{E}{h}\right)\right]! - \ln \left[\frac{1}{2}\left(N + \frac{E}{h}\right)\right]! \end{aligned} \tag{1.29}$$

Using Stirling's approximation, valid for large N

$$\ln N! \approx N \ln N - N, \tag{1.30}$$

one finds

$$S(E) = N \ln N - \frac{N + E/h}{2} \ln \frac{N + E/h}{2} - \frac{N - E/h}{2} \ln \frac{N - E/h}{2}. \tag{1.31}$$

Perfect Gas of Independent Particles

We consider a gas of independent particles confined into a cubic container of volume $V = L^3$. The generalized coordinates q_j satisfy the constraints

$$0 \le q_j \le L, \quad j = 1, \ldots, L. \tag{1.32}$$

The energy E comes only from the kinetic contribution:

$$E = \sum_{j=1}^{3N} \frac{p_j^2}{2m}. \tag{1.33}$$

The accessible volume in phase space is the product of the accessible volume for each particle, times the area of the hypersphere of radius $\sqrt{2mE}$, embedded in a

3N-dimensional space. The area of the hypersphere of radius R in a D-dimensional space is

$$\mathcal{A}_D(R) = \frac{D\pi^{D/2}}{\Gamma\left(\frac{D}{2}+1\right)} R^{D-1}, \tag{1.34}$$

where $\Gamma(x) = \int_0^\infty dt\, t^{x-1} e^{-t}$ is the Euler Gamma function (a generalization of the factorial to real values, satisfying $\Gamma(n) = (n-1)!$ for integer $n \geq 1$). So the accessible volume $\Omega_V(E)$ is given by

$$\begin{aligned}\Omega_V(E) &= L^{3N} \frac{3N\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)} \sqrt{2mE}^{3N-1} \\ &= \frac{3N\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)} \sqrt{2m}^{3N-1} V^N E^{\frac{3N-1}{2}}.\end{aligned} \tag{1.35}$$

The corresponding entropy reads, assuming $N \gg 1$,

$$S_V(E) = \ln \Omega(E) = S_0 + \frac{3N}{2} \ln E + N \ln V \tag{1.36}$$

with

$$S_0 = \ln\left(\frac{3N\pi^{3N/2}}{\Gamma\left(\frac{3N}{2}+1\right)} \sqrt{2m}^{3N}\right). \tag{1.37}$$

Note that in principle, some corrections need to be included to take into account quantum effects, namely the fact that quantum particles are indistinguishable. This allows in particular $\Omega(E)$ to be made dimensionless, thus rendering the entropy independent of the system of units chosen. Quantum effects are also important in order to recover the extensivity of the entropy, that is, the fact that the entropy is proportional to the number N of particles. In the present form, $N \ln N$ terms are present, making the entropy grow faster than the system size. This is related to the so-called Gibbs paradox. However, we shall not describe these effects in more details here, and refer the reader to standard textbooks.

1.2.4 Distribution of Energy Over Subsystems and Statistical Temperature

Let us consider an isolated system, with fixed energy and number of particles. We then imagine that the system is partitioned into two subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$, the two subsystems being separated by a wall which allows energy exchanges, but not exchanges of matter. The total energy of the system $E = E_1 + E_2$ is fixed, but the energies E_1 and E_2 fluctuate due to thermal exchanges.

For a fixed energy E, let us evaluate the number $\Omega(E_1|E)$ of configurations of the system such that the energy of $\mathcal{S}_1$ has a given value E_1. In the absence of long-range forces in the system, the two subsystems can be considered as statistically independent (apart from the total energy constraint), leading to

$$\Omega(E_1|E) = \Omega_1(E_1)\Omega_2(E - E_1), \tag{1.38}$$

where $\Omega_k(E_k)$ is the number of configurations of $\mathcal{S}_k$.

The most probable value E_1^* of the energy E_1 maximizes by definition $\Omega(E_1|E)$, or equivalently $\ln \Omega(E_1|E)$:

$$\left.\frac{\partial}{\partial E_1}\right|_{E_1^*} \ln \Omega(E_1|E) = 0. \tag{1.39}$$

Combining Eqs. 1.38 and 1.39, one finds

$$\left.\frac{\partial \ln \Omega_1}{\partial E_1}\right|_{E_1^*} = \left.\frac{\partial \ln \Omega_2}{\partial E_2}\right|_{E_2^*=E-E_1^*}. \tag{1.40}$$

Thus it turns out that two quantities defined independently in each subsystem are equal at equilibrium. Namely, defining

$$\beta_k \equiv \left.\frac{\partial \ln \Omega_k}{\partial E_k}\right|_{E_k^*}, \quad k = 1, 2, \tag{1.41}$$

one has $\beta_1 = \beta_2$. This is the reason why the quantity β_k is called the statistical temperature of $\mathcal{S}_k$. In addition, it can be shown that for large systems, the common value of β_1 and β_2 is also equal to

$$\beta = \frac{\partial S}{\partial E} \tag{1.42}$$

computed for the global isolated system.

To identify the precise link between β and the standard thermodynamic temperature, we notice that in thermodynamics, one has for a system that exchanges no work with its environment:

$$dE = T dS, \tag{1.43}$$

which indicates that $\beta = 1/T$ (we recall that we have set $k_B = 1$). This is further confirmed on the example of the perfect gas, for which one finds using Eq. 1.36

$$\beta \equiv \frac{\partial S_V}{\partial E} = \frac{3N}{2E}, \tag{1.44}$$

or equivalently

$$E = \frac{3N}{2\beta}. \tag{1.45}$$

Besides, one has from the kinetic theory of gases

$$E = \frac{3}{2}NT \tag{1.46}$$

(which is nothing but equipartition of energy), leading again to the identification $\beta = 1/T$. Hence, in the microcanonical ensemble, one generically defines temperature T through the relation

$$\frac{1}{T} = \frac{\partial S}{\partial E}. \tag{1.47}$$

We now further illustrate this relation on the example of the paramagnetic crystal that we already encountered earlier. From Eq. 1.31, one has

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{1}{2h} \ln \frac{N - E/h}{N + E/h}. \tag{1.48}$$

This last equation can be inverted to express the energy E as a function of temperature, yielding

$$E = -Nh \tanh \frac{h}{T}. \tag{1.49}$$

This relation has been obtained by noticing that $x = \tanh y$ is equivalent to

$$y = \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right). \tag{1.50}$$

In addition, from the relation $E = -Mh$, where $M = \sum_{i=1}^{N} s_i$ is the total magnetization, one obtains as a byproduct

$$M = N \tanh \frac{h}{T}. \tag{1.51}$$

1.3 Equilibrium System in Contact with its Environment

1.3.1 Exchanges of Energy

Realistic systems are most often not isolated, but they rather exchange energy with their environment. A natural idea is then to describe the system $\mathcal{S}$ of interest as a macroscopic subsystem of a large isolated system $\mathcal{S} \cup \mathcal{R}$, where $\mathcal{R}$ is the environment, or energy reservoir. The total energy $E_{\text{tot}} = E + E_{\mathcal{R}}$ is fixed. A configuration C_{tot} of the total system can be written as $C_{\text{tot}} = (C, C_{\mathcal{R}})$, where C is a configuration of $\mathcal{S}$ and $C_{\mathcal{R}}$ is a configuration of $\mathcal{R}$. The total system $\mathcal{S} \cup \mathcal{R}$ is isolated and at equilibrium, so that it can be described within the macrocanonical framework:

$$P_{\text{tot}}(C_{\text{tot}}) = \frac{1}{\Omega_{\text{tot}}(E_{\text{tot}})}, \quad C_{\text{tot}} = (C, C_{\mathcal{R}}). \tag{1.52}$$

To obtain the probability of a configuration C of $\mathcal{S}$, one needs to sum $P_{\text{tot}}(C_{\text{tot}})$ over all configurations $C_{\mathcal{R}}$ of $\mathcal{R}$ compatible with the total energy E_{tot}, namely

$$P(C) = \sum_{C_{\mathcal{R}}:E_{\mathcal{R}}=E_{\text{tot}}-E(C)} P_{\text{tot}}(C, C_{\mathcal{R}}) = \frac{\Omega_{\mathcal{R}}(E_{\text{tot}} - E(C))}{\Omega_{\text{tot}}(E_{\text{tot}})}. \tag{1.53}$$

We introduce the entropy of the reservoir $S_{\mathcal{R}}(E_{\mathcal{R}}) = \ln \Omega_{\mathcal{R}}(E_{\mathcal{R}})$. Under the assumption that $E(C) \ll E_{\text{tot}}$, one has

$$S_{\mathcal{R}}(E_{\text{tot}} - E(C)) \approx S_{\mathcal{R}}(E_{\text{tot}}) - E(C) \frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\bigg|_{E_{\text{tot}}}. \tag{1.54}$$

One also has

$$\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\bigg|_{E_{\text{tot}}} \approx \frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\bigg|_{E^*_{\mathcal{R}}} = \frac{1}{T} \tag{1.55}$$

where T is the temperature of the reservoir. Altogether, we have

$$P(C) = \frac{\Omega_{\mathcal{R}}(E_{\text{tot}})}{\Omega_{\text{tot}}(E_{\text{tot}})} e^{-E(C)/T}. \tag{1.56}$$

Note that the prefactor $\Omega_{\mathcal{R}}/\Omega_{\text{tot}}$ depends on the total energy E_{tot}, while we would like *P(C)* to depend only on the energy E of the system considered. This problem can however be bypassed by noticing that the distribution $P(C)$ should be normalized to unity, namely, $\sum_C P(C) = 1$. Introducing the partition function

$$Z = \sum_C e^{-E(C)/T}, \tag{1.57}$$

one can then eventually rewrite the distribution $P(C)$ in the form

$$P(C) = \frac{1}{Z} e^{-E(C)/T}, \tag{1.58}$$

which is the standard form of the canonical distribution.

The partition function Z is a useful tool in statistical physics. For instance, the average energy $\langle E \rangle$ can be easily computed from Z:

$$\begin{aligned}\langle E\rangle &= \sum_C P(C)E(C) = \sum_C E(C)\frac{1}{Z}e^{-E(C)/T}\\ &= \frac{1}{Z}\sum_C E(C)e^{-\beta E(C)} \\ &= -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}.\end{aligned} \tag{1.59}$$

Instead of Z, one may also use the "free energy" F defined as

$$F = -T \ln Z. \tag{1.60}$$

Let us give a simple example of computation of Z, in the case of the paramagnetic spin model. The partition function is given by

$$Z = \sum_{\{s_i=\pm 1\}} e^{-\beta E(\{s_i\})}, \tag{1.61}$$

with $E(\{s_i\}) = -h\sum_{i=1}^{N} s_i$. Hence one has

$$\begin{aligned}Z &= \sum_{\{s_i=\pm 1\}} e^{\beta h \sum_{i=1}^{N} s_i}\\ &= \sum_{\{s_i=\pm 1\}} \prod_{i=1}^{N} e^{\beta h s_i} = \prod_{i=1}^{N}\left(\sum_{s=\pm 1} e^{\beta h s}\right)\end{aligned} \tag{1.62}$$

so that one finds

$$Z = \left(e^{\beta h} + e^{-\beta h}\right)^N. \tag{1.63}$$

Turning to the average energy, one has

$$\langle E\rangle = -\frac{\partial \ln Z}{\partial \beta} = -N\frac{\partial}{\partial \beta}\ln\left(e^{\beta h} + e^{-\beta h}\right), \tag{1.64}$$

so that one obtains, recalling that $\beta = 1/T$,

$$\langle E\rangle = -Nh \tanh\frac{h}{T}. \tag{1.65}$$

It is interesting to note that the above equation has exactly the same form as Eq. 1.49, provided that one replaces E, which is fixed in the microcanonical ensemble, by its average value $\langle E\rangle$ in the canonical ensemble. This property is an example of a general property called the "equivalence of ensemble": in the limit of large systems, the relations between macroscopic quantities are the same in the different statistical ensembles, regardless of which quantity is fixed and which one is fluctuating through exchanges with a reservoir. The interpretation of this important property is basically that fluctuating observables actually have very small relative fluctuations for large system sizes. This property is also deeply related to the law of large numbers and to the central limit theorem. Indeed, the relative fluctuations (quantified by the standard deviation normalized by the number of terms) of a sum of independent and identically distributed random variables go to zero when the number of terms in the sum goes to infinity. Note that equivalence of ensemble generally breaks down in the presence of long-range interactions in the systems.

Another example where the computation of Z is straightforward is the perfect gas. In this case, one has

$$\begin{aligned} Z &= \int_0^L dq_1 \ldots \int_0^{3L} dq_{3N} \int_{-\infty}^{\infty} dp_1 \ldots \int_{-\infty}^{\infty} dp_{3N} e^{-\beta \sum_{j=1}^{3N} p_j^2/2m} \\ &= L^{3N} \prod_{j=1}^{3N} \int_{-\infty}^{\infty} dp_j e^{-\beta p_j^2/2m} \\ &= V^N \left(\int_{-\infty}^{\infty} dp e^{-\beta p^2/2m} \right)^{3N}. \end{aligned} \tag{1.66}$$

Given that

$$\int_{-\infty}^{\infty} dp e^{-\beta p^2/2m} = \sqrt{\frac{2\pi m}{\beta}}, \tag{1.67}$$

one finds

$$Z = V^N \left(\frac{2\pi m}{\beta} \right)^{\frac{3N}{2}}. \tag{1.68}$$

Computing the average energy leads to

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{3N}{2\beta} = \frac{3}{2} NT \tag{1.69}$$

yielding another example of ensemble equivalence, as this result has the same form as Eq. 1.45.

1.3.2 Canonical Entropy

As we have seen above, the microcanonical entropy is defined as $S(E) = \ln \Omega(E)$. This definition is clearly related to the equiprobability of accessible microscopic configurations, since it is based on a counting of accessible configurations. A natural question is then to know how to define the entropy in more general situations. A generic definition of entropy has appeared in information theory, namely:

$$S = -\sum_{C} P(C) \ln P(C) \tag{1.70}$$

where the sum is over all accessible configurations of the system. This entropy is called the Boltzmann–Gibbs, von Neumann or Shannon entropy depending on the context. This definition of entropy is moreover consistent with the microcanonical one: if $P(C) = 1/\Omega(E)$ for configurations of energy E, and $P(C) = 0$ otherwise, one finds:

$$S = -\sum_{C:E(C)=E} \frac{1}{\Omega(E)} \ln \frac{1}{\Omega(E)} = \ln \Omega(E). \tag{1.71}$$

In this general framework, the canonical entropy reads

$$\begin{aligned} S_{\mathrm{can}} &= -\sum_{C} P_{\mathrm{can}}(C) \ln P_{\mathrm{can}}(C) \\ &= -\sum_{C} \frac{1}{Z} e^{-\beta E(C)} \ln \left(\frac{1}{Z} e^{-\beta E(C)} \right) \\ &= \sum_{C} \frac{1}{Z} e^{-\beta E(C)} \left(\ln Z + \beta E(C) \right) \\ &= \ln Z + \beta \langle E \rangle. \end{aligned} \tag{1.72}$$

Recalling that the free energy F is defined as $F = -T \ln Z$, one thus has $TS = -F + \langle E \rangle$, which is nothing but the well-known relation $F = \langle E \rangle - TS$. Another standard thermodynamic relation may be found using $\langle E \rangle = -\partial \ln Z / \partial \beta$:

$$\begin{aligned} S_{\mathrm{can}} &= \ln Z - \beta \frac{\partial \ln Z}{\partial \beta} \\ &= \ln Z + T \frac{\partial \ln Z}{\partial T} \\ &= \frac{\partial}{\partial T} (T \ln Z) \end{aligned}$$

so that one finds the standard thermodynamic relation

$$S_{\mathrm{can}} = -\frac{\partial F}{\partial T}. \tag{1.73}$$

1.3.3 Exchanges of Particles with a Reservoir: The Grand-Canonical Ensemble

Similarly to what was done to obtain the canonical ensemble from the microcanonical one by allowing energy exchanges with a reservoir, one can further allow exchanges of particles with a reservoir. The corresponding situation is called the grand-canonical ensemble.

We thus consider a macroscopic system $\mathcal{S}$ exchanging both energy and particles with a reservoir $\mathcal{R}$. The total system $\mathcal{S} \cup \mathcal{R}$ is isolated with total energy E_{tot} and total number of particles N_{tot} fixed:

$$E + E_{\mathcal{R}} = E_{\mathrm{tot}}, \quad N + N_{\mathcal{R}} = N_{\mathrm{tot}}. \tag{1.74}$$

Generalizing the calculations made in the canonical case, one has

$$\begin{aligned} P_{\mathrm{GC}}(C) &= K\Omega_{\mathcal{R}}(E_{\mathcal{R}}, N_{\mathcal{R}}) \\ &= K\Omega_{\mathcal{R}}(E_{\mathrm{tot}} - E(C), N_{\mathrm{tot}} - N(C)) \\ &= K \exp\left[S_{\mathcal{R}}(E_{\mathrm{tot}} - E(C), N_{\mathrm{tot}} - N(C))\right]. \end{aligned}$$

As $E(C) \ll E_{\mathrm{tot}}$ and $N(C) \ll N_{\mathrm{tot}}$, one can expand the entropy $S_{\mathcal{R}}(E_{\mathrm{tot}} - E(C), N_{\mathrm{tot}} - N(C))$ to first order:

$$\begin{aligned} S_{\mathcal{R}}(E_{\mathrm{tot}} - E(C), N_{\mathrm{tot}} - N(C)) &= S_{\mathcal{R}}(E_{\mathrm{tot}}, N_{\mathrm{tot}}) \\ &\quad - E(C)\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\bigg|_{E_{\mathrm{tot}},N_{\mathrm{tot}}} - N(C)\frac{\partial S_{\mathcal{R}}}{\partial N_{\mathcal{R}}}\bigg|_{E_{\mathrm{tot}},N_{\mathrm{tot}}}. \end{aligned} \tag{1.75}$$

As before, the derivative $\partial S_{\mathcal{R}}/\partial E_{\mathcal{R}}$ is identified with $1/T$. We also introduce a new parameter, the chemical potential μ, defined as:

$$\mu = -T\frac{\partial S_{\mathcal{R}}}{\partial N_{\mathcal{R}}}. \tag{1.76}$$

Similarly to the temperature which takes equal values when subsystems exchanging energy have reached equilibrium, the chemical potential takes equal values in subsystems exchanging particles, when equilibrium is attained. Gathering all the above results and notations, one finds that

$$P_{\mathrm{GC}}(C) = \frac{1}{Z_{\mathrm{GC}}} \exp\left(-\frac{1}{T}E(C) + \frac{\mu}{T}N(C)\right) \tag{1.77}$$

which is the standard form of the so-called grand-canonical distribution. The normalization constant Z_{GC}, defined by

$$Z_{\mathrm{GC}} = \sum_{C} \exp\left(-\frac{1}{T}E(C) + \frac{\mu}{T}N(C)\right), \tag{1.78}$$

is called the grand-canonical partition function.

1.4 Phase Transitions

1.4.1 Example of the Ising Model

Phase transitions correspond to a sudden change of behavior of the system when varying an external parameter across a transition point. This could be of interest in complex systems well-beyond physics, and is generically associated with collective effects. To illustrate this last property, let us briefly come back to the paramagnetic model defined in Sect. 1.2.3, for which the mean magnetization per spin is given by

$$\langle m \rangle \equiv \frac{\langle M \rangle}{N} = \tanh\left(\frac{h}{T}\right). \tag{1.79}$$

The magnetization is non-zero only if there is a non-zero external field which tends to align the spins. A natural question is thus to know whether one could obtain a non-zero magnetization by including interactions tending to align spins between them (and not with respect to an external field). In this spirit, let us consider the standard (interaction) energy of the Ising model, in the absence of external field:

$$E_{\mathrm{Ising}} = -J \sum_{\langle i,j \rangle} s_i s_j, \quad J > 0. \tag{1.80}$$

This interaction energy is minimized when all spins are parallel. To compute the mean magnetization per spin, one would need to compute either the partition function in presence of a external magnetic field and take the derivative of the free energy with respect to the field, or to compute directly the mean magnetization from its definition. In any case, this is a very complicated task as soon as the space dimension D is larger than one, and the exact calculation has been achieved only in dimensions one and two. The results can be summarized as follows:

- $D = 1$: $m = 0$ for all $T > 0$, so that there is no phase transition at finite temperature. Calculations are relatively easy.
- $D = 2$: there is a phase transition at a finite critical temperature T_c, namely $m = 0$ for $T \geq T_c$ and $m \neq 0$ for $T < T_c$. Calculations are however very technical.
- $D \geq 3$: no analytical solution is known, but numerical simulations show that there is a phase transition at a finite temperature that depends on D.

1.4.2 Ising Model in Fully Connected Geometry

An interesting benchmark model, which can be shown analytically to exhibit a phase transition, is the fully connected Ising model, whose energy is defined as

$$E_{\mathrm{fc}} = -\frac{J}{N} \sum_{i<j} s_i s_j + E_0, \tag{1.81}$$

where the sum is over all pairs of spins in the system. The $1/N$ prefactor is included in order to keep the energy per spin finite in the large N limit. The term E_0 is added for later convenience, and is arbitrary at this stage (it does not modify the canonical distribution). Considering the magnetization $M = \sum_{i=1}^{N} s_i$, one has, given that $s_i^2 = 1$,

$$M^2 = 2\sum_{i<j} s_i s_j + N \tag{1.82}$$

from which one finds

$$E_{\mathrm{fc}} = -\frac{J}{2N}(M^2 - N) + E_0 = -\frac{J}{2N}M^2 + \frac{J}{2} + E_0. \tag{1.83}$$

Choosing $E_0 = -J/2$, and introducing the magnetization per spin $m = M/N$, one finds

$$E_{\mathrm{fc}} = -\frac{J}{2}Nm^2. \tag{1.84}$$

One possible way to detect the phase transition is to compute the probability distribution $P(m)$ of the magnetization, by summing over all configurations having a given magnetization m:

$$\begin{aligned} P(m) &= \frac{1}{Z}\sum_{C:m(C)=m} e^{-\beta E(C)} \\ &= \frac{1}{Z}\Omega(m)e^{\frac{1}{2}\beta J N m^2} \\ &= \frac{1}{Z}e^{S(m)+\frac{1}{2}\beta J N m^2} \end{aligned} \tag{1.85}$$

where $\Omega(m) = e^{S(m)}$ is the number of configurations with magnetization m. Using the relation

$$\Omega(m) = \frac{N!}{N_+!N_-!} \tag{1.86}$$

with

$$N_+ = \frac{N}{2}(1+m), \quad N_- = \frac{N}{2}(1-m), \tag{1.87}$$

one obtains for $S(m) = \ln \Omega(m)$

$$S(m) = -N\left[\frac{1+m}{2}\ln\left(\frac{1+m}{2}\right) + \frac{1-m}{2}\ln\left(\frac{1-m}{2}\right)\right]. \tag{1.88}$$

Hence from Eq. 1.85 it turns out that $P(m)$ can be written as

$$P(m) = e^{-Nf(m)} \tag{1.89}$$

with $f(m)$ given by

$$f(m) = \frac{1+m}{2} \ln\left(\frac{1+m}{2}\right) + \frac{1-m}{2} \ln\left(\frac{1-m}{2}\right) - \frac{J}{2T} m^2. \tag{1.90}$$

The function $f(m)$ is called a large deviation function, or a Landau free energy function. The magnetization m_0 that maximizes the probability distribution $P(m)$ corresponds to a minimum of $f(m)$. Moreover, fluctuations around m_0 are exponentially suppressed with N. For high temperature T, the term J/T is small, and the entropic contribution to $f(m)$ should dominate, leading to $m_0 = 0$. To understand what happens when temperature is progressively lowered, it is useful to expand $f(m)$ for small values of m, up to order m^4, leading to:

$$f(m) = -\ln 2 + \frac{1}{2}\left(1 - \frac{J}{T}\right) m^2 + \frac{1}{12} m^4 + \mathcal{O}(m^6). \tag{1.91}$$

One can then distinguish two different cases:

- If $T \geq T_c \equiv J$, $f(m)$ has only one minimum, for $m = 0$.
- If $T < T_c$, $f(m)$ has two symmetric minima $\pm m_0$. These minima are obtained as solutions of $df/dm = 0$:

$$\frac{df}{dm} = \left(1 - \frac{J}{T}\right) m + \frac{1}{3} m^3 = -\left|1 - \frac{J}{T}\right| m + \frac{1}{3} m^3 = 0. \tag{1.92}$$

The non-zero solutions are $m = \pm m_0$ with

$$m_0 = \sqrt{3\left(\frac{J}{T} - 1\right)} = \sqrt{3}\left(\frac{T_c - T}{T}\right)^{1/2}. \tag{1.93}$$

It can be checked easily that the solution $m = 0$ corresponds in this case to a local maximum of $f(m)$, and thus to a local minimum of $P(m)$.

Hence, there is a phase transition at $T = T_c \equiv J$, T_c being called the critical temperature. The most probable magnetization m_0 is called the "order parameter of the phase transition", as the phase transition is characterized by the onset of a non-zero value of m_0. In addition, the order parameter varies as $m_0 \sim (T_c - T)^\beta$ for T close to T_c, with $\beta = 1/2$ here. The exponent β is an example of critical exponent, and the value $\beta = 1/2$ is called the "mean-field value" of β, for reasons that will become clear in the next section. The notation β is standard for the critical exponent associated to the order parameter, and should not be confused with the inverse temperature $\beta = 1/T$.

An important remark is that the average value $\langle m \rangle$ of the magnetization is still zero for $T < T_c$, since the two values $\pm m_0$ of the magnetization have the same probability. However, for a large system, the time needed to switch between states m_0 and $-m_0$ becomes very large, and the time-averaged magnetization over a typical observation time window is non-zero, and equal either to m_0 or to $-m_0$.

1.4.3 Ising Model with Finite Connectivity

We now come back to the finite-dimensional Ising model. As mentioned above, the analytical solution is hard to obtain in dimension $D = 2$, and is not known in higher dimensions. However, useful approximations have been developed, the most famous one being called the mean-field approximation.

The reason why the fully connected model can be easily solved analytically is that its energy E is a function of the magnetization m only, as seen in Eq. 1.84. When the model is defined on a finite-dimensional lattice, this property is no longer true, and the energy reads:

$$E = -\frac{J}{2} \sum_{i=1}^{N} s_i \left(\sum_{j \in \mathcal{V}(i)} s_j \right). \tag{1.94}$$

where $\mathcal{V}(i)$ is the set of neighboring sites of site i. The factor $1/2$ comes from the fact that a given link of the lattice now appears twice in the sum. This last expression can be rewritten as

$$E = -DJ \sum_{i=1}^{N} s_i \langle s_j \rangle_{\mathcal{V}(i)}, \tag{1.95}$$

$\langle s_j \rangle_{\mathcal{V}(i)}$ being the local average magnetization of the set of neighbors $\mathcal{V}(i)$:

$$\langle s_j \rangle_{\mathcal{V}(i)} = \frac{1}{2D} \sum_{j \in \mathcal{V}(i)} s_j. \tag{1.96}$$

The parameter D is the space dimension, and the number of neighbors of a given site i is $2D$, given that we consider hypercubic lattices (square lattice in $D = 2$, cubic lattice in $D = 3, \ldots$).

As a first approximation, one could replace the average magnetization over the set of neighbors by the global magnetization per spin of the whole system, $m = N^{-1} \sum_{i=1}^{N} s_i$:

$$\langle s_j \rangle_{\mathcal{V}(i)} \to m. \tag{1.97}$$

This approximation leads to the following expression of the energy

$$E \approx E_{\text{mf}} = -DJm \sum_{i=1}^{N} s_i = -DJNm^2, \tag{1.98}$$

where the subscript “mf” stands for “mean-field” approximation. Then E_{mf} depends only on the magnetization m, and has a form similar to the energy E_{fc} of the fully connected model. One can define an effective coupling $J_{\mathrm{mf}} = 2DJ$ so that the form of the two energies become exactly the same, namely

$$E_{\mathrm{mf}} = -\frac{1}{2} J_{\mathrm{mf}} N m^2. \tag{1.99}$$

Now it is clear that the results of the fully connected model can be applied to the present mean-field approximation, yielding a phase transition at $T_c^{\mathrm{mf}} = J_{\mathrm{mf}} = 2DJ$. For $T > T_c^{\mathrm{mf}}$, $\langle m \rangle = 0$ while for $T < T_c^{\mathrm{mf}}$, but close to T_c^{mf}, $\langle m \rangle \sim (T_c^{\mathrm{mf}} - T)^{1/2}$. Qualitatively, the approximation is expected to be valid for large space dimension D. It can be shown, using more involved arguments, that for $D \geq 4$, the approximation is semi-quantitatively valid, in the sense that the value $\beta = 1/2$ of the critical exponent, obtained from the approximation, is correct. However, the value of the critical temperature T_c^{mf} is not correctly predicted by the mean-field approximation, namely $T_c \neq T_c^{\mathrm{mf}}$. For $D < 4$, the value of β differs from the mean-field value 1/2, and the mean-field approximation breaks down. For $D = 3$, numerical simulations indicate that $\beta \approx 0.31$, and for $D = 2$, the exact solution yields $\beta = 1/8$. Finally, for $D = 1$, $\langle m \rangle = 0$ except for $T = 0$, so that the exponent β is not defined.

The discrepancy mentioned above between mean-field predictions and results obtained in low-dimensional systems mainly comes from the presence of fluctuations of the local magnetization $\sum_{j \in \mathcal{V}(i)} s_j$. Since on the other hand exact solutions are very hard to obtain, there is need for a different approach, that could be generic enough and could be centered on the issue of correlation, which is at the heart of the difficulties encountered. This is precisely the aim of the renormalization group approach.

1.4.4 Renormalization Group Approach: A Brief Introduction

A standard observation on finite dimensional systems exhibiting a continuous phase transition is that the correlation length diverges when the temperature approaches the critical temperature T_c. The correlation length is defined through the correlation function

$$C_{ij} = \langle (s_i - m_0)(s_j - m_0) \rangle = \langle s_i s_j \rangle - m_0^2. \tag{1.100}$$

As soon as the distance $r = d_{ij}$ between sites i and j is large with respect to the lattice spacing a, the correlation function generally becomes isotropic, $C_{ij} = C(r)$. In addition, the large distance behavior of *C(r)* is often of the form

$$C(r) \sim \frac{1}{r^{\alpha}} e^{-r/\xi}, \quad \alpha > 0, \tag{1.101}$$

which defines the correlation length ξ. The latter diverges for $T \to T_c$. This is the reason why direct calculations in the range $T \approx T_c$ are very difficult, due to the strong correlation between spins. A natural idea is to look for an approach that could reduce in some way the intensity of correlations, in order to make calculations tractable.

This is basically the principle of the renormalization group (RG) approach, in which one progressively integrates out small scale degrees of freedom. The idea is that at the critical point, structures are present at all scales, from the lattice spacing to the system size. A RG transform may intuitively be thought of as defocusing the picture of the system, so that fine details become blurred. This method is actually very general, and could be relevant in many fields of complex system sciences, given that issues like large scale correlations and scale invariance or fractals are often involved in complex systems.

For definiteness, let us however consider again the Ising model. To implement the RG ideas in a practical way, one could make blocks of spins and define an effective spin for each block, with effective interactions with the neighboring blocks. The effective interactions are defined in such a way that the large scale properties are the same as for the original (non-renormalized) model. This is done in practice by conserving the partition function, namely $Z' = Z$ (in the present section, the prime denotes renormalized quantities). One would then like to define a renormalized interaction constant J' such that

$$H' = -J' \sum_{\langle B_1, B_2 \rangle} S_{B_1} S_{B_2} \tag{1.102}$$

where B_1 and B_2 are generic labels for the blocks (the sites of the renormalized lattice). The problem is that very often, the RG transform generates new effectives couplings, like next-nearest-neighbor couplings, that were absent in the original model, and the number of couplings keeps increasing with the number of iterations of the RG transform. However, in some simple cases, the transformation can be performed exactly, without increasing the number of coupling constants, as we shall see later on.

Yet, let us first emphasize the practical interest of the RG transform. We already mentioned that one of the main difficulties comes from the presence of long-range correlations close to the critical point. Through the RG transform, the lattice spacing becomes $a' = 2a$ (if one makes blocks of linear size $2a$). On the contrary, the correlation length remains unchanged, since the large scale properties remain unaffected by the RG transform. Hence the correlation length expressed in unit of the lattice spacing, namely ξ/a, decreases by a factor of 2 in the transformation, to become

$$\frac{\xi'}{a'} = \frac{1}{2}\frac{\xi}{a}. \tag{1.103}$$

Thus upon iterations of the RG transform, the effective Hamiltonian becomes such that $\xi' \sim a'$, so that standard approximation schemes (mean-field, ...) can be used. One then needs to follow the evolution of the coupling constant J' under iterations. This is called the renormalization flow.

An explicit example can be given with the one-dimensional Ising chain, using a specific RG transform called decimation procedure [6]. We start with the energy (or Hamiltonian)

$$H = \sum_{i=1}^{N} H_{i,i+1}(s_i, s_{i+1}) \tag{1.104}$$

where the local interaction term $H_{i,i+1}(s_i, s_{i+1})$ is given by

$$H_{i,i+1}(s_i, s_{i+1}) = -J s_i s_{i+1} + c. \tag{1.105}$$

Note that periodic boundary conditions are understood. The constant c plays no role at this stage, but it will be useful later on in the renormalization procedure. The basic idea of the decimation procedure is to perform, in the partition function, a partial sum over the spins of—say—odd indices in order to define renormalized coupling constants J' and h'. Then summing over the values of the spins with even indices yields the partition function Z' of the renormalized model, which is by definition of the renormalization procedure equal to the initial partition function Z. To be more explicit, one can write Z as

$$Z = \sum_{\{s_{2j}\}} \sum_{\{s_{2j+1}\}} \exp\left[-\beta H(\{s_i\})\right] \tag{1.106}$$

where $\sum_{\{s_{2j}\}}$ (resp. $\sum_{\{s_{2j+1}\}}$) indicates a sum over all possible values of the $N/2$ variables $\{s_{2j}\}$ (resp. $\{s_{2j+1}\}$). Equation 1.106 can then be rewritten in the following form:

$$Z = \sum_{\{s_{2j}\}} \exp\left[-\beta H'(\{s_{2j}\})\right] \tag{1.107}$$

where $H'(\{s_{2j}\})$ is the renormalized Hamiltonian, defined by

$$\exp\left[-\beta H'(\{s_{2j}\})\right] = \sum_{\{s_{2j+1}\}} \exp\left[-\beta H(\{s_i\})\right]. \tag{1.108}$$

Assuming that the renormalized Hamiltonian can be decomposed into a sum of local terms

$$H'(\{s_{2j}\}) = \sum_{j=1}^{N/2} H'_{j,j+1}(s_{2j}, s_{2j+2}) \tag{1.109}$$

we get from Eq. 1.108 the relation

$$\begin{aligned}
&\prod_{j=1}^{N/2} \exp\left[-\beta H'_{j,j+1}(s_{2j}, s_{2j+2})\right] \\
&\quad = \sum_{\{s_{2j+1}\}} \prod_{j=1}^{N/2} \exp\left[-\beta H_{2j,2j+1}(s_{2j}, s_{2j+1}) - \beta H_{2j+1,2j+2}(s_{2j+1}, s_{2j+2})\right] \\
&\quad = \prod_{j=1}^{N/2} \sum_{s_{2j+1}} \exp\left[-\beta H_{2j,2j+1}(s_{2j}, s_{2j+1}) - \beta H_{2j+1,2j+2}(s_{2j+1}, s_{2j+2})\right]
\end{aligned} \tag{1.110}$$

where in the last line, the sum runs over the single variable s_{2j+1}, the index j being fixed within the product. This last relation is satisfied if, for any given $j = 1, \ldots, N/2$, and any given values of s_{2j} and s_{2j+2},

$$\begin{aligned}
&\exp\left[-\beta H'_{j,j+1}(s_{2j}, s_{2j+2})\right] \\
&\quad = \sum_{s_{2j+1}=\pm 1} \exp\left[-\beta H_{2j,2j+1}(s_{2j}, s_{2j+1}) - \beta H_{2j+1,2j+2}(s_{2j+1}, s_{2j+2})\right].
\end{aligned} \tag{1.111}$$

Further assuming that $H'_{j,j+1}(s_{2j}, s_{2j+2})$ takes the form

$$H'_{j,j+1}(s_{2j}, s_{2j+2}) = -J' s_{2j} s_{2j+2} + c', \tag{1.112}$$

where J' and c' are the renormalized parameters, one obtains

$$\exp\left[\beta J' s_{2j} s_{2j+2} - \beta c'\right] = \sum_{s_{2j+1}=\pm 1} \exp\left[\beta J (s_{2j} s_{2j+1} + s_{2j+1} s_{2j+2}) - 2\beta c\right]. \tag{1.113}$$

Introducing the reduced variable[4]

$$u = e^{-4\beta J}, \tag{1.114}$$

Equation 1.113 leads to the following recursion relation:

$$u' = \frac{4u}{(1+u)^2}. \tag{1.115}$$

Let us denote as ξ_{nd} the dimensionless correlation length

$$\xi_{\text{nd}} = \frac{\xi}{a}. \tag{1.116}$$

Then from Eq. 1.103 the recursion relation for ξ_{nd} reads

---

[4] We do not follow here the evolution of the constant c under renormalization, and rather focus on the evolution of the physically relevant coupling constant J.

$$\xi'_{\mathrm{nd}} = \frac{1}{2}\xi_{\mathrm{nd}}, \tag{1.117}$$

from which one deduces that the fixed points of the renormalization procedure, that satisfy $\xi'_{\mathrm{nd}} = \xi_{\mathrm{nd}}$, can only be $\xi_{\mathrm{nd}} = \infty$ or $\xi_{\mathrm{nd}} = 0$. The latter is called the trivial fixed point, as it corresponds to the limit situation where no correlation is present in the system. In contrast, the fixed point $\xi_{\mathrm{nd}} = \infty$ corresponds to the critical fixed point, where correlation extends over the whole system size. As ξ_{nd} decreases through iteration of the RG transform, the critical fixed point $\xi_{\mathrm{nd}} = \infty$ is unstable, while the trivial fixed point $\xi_{\mathrm{nd}} = 0$ is stable.

Coming back to the iteration relation Eq. 1.115, let us first look for the fixed points of this equation, namely the solutions of

$$u = \frac{4u}{(1+u)^2}. \tag{1.118}$$

The value $u = 0$ is obviously a solution, and it is easy to check that $u = 1$ is the other positive solution ($u = -3$ is the third solution, but in view of Eq. 1.114, we are seeking for positive solutions only). Then to identify which one of the two fixed points is the critical point, we need to investigate the stability of each fixed point under iteration. The stability is studied by introducing a small variation δu around a given fixed point u_1, namely $u = u_1 \pm \delta u$, and writing the evolution equation for δu to leading order. For $u_1 = 0$, one finds, with $u = \delta u$,

$$\delta u' = \frac{4\delta u}{(1+\delta u)^2} \approx 4\delta u, \quad \delta u > 0, \tag{1.119}$$

so that δu increases upon iteration: the fixed point $u_1 = 0$ is unstable, and thus corresponds to the critical fixed point. Besides, the fixed point $u_1 = 1$ is easily checked to be stable. Using $u = 1 - \delta u$, we have

$$1 - \delta u' = \frac{4(1-\delta u)}{(2-\delta u)^2}, \tag{1.120}$$

leading after a second order expansion in δu to

$$\delta u' \approx \frac{1}{4}\delta u^2. \tag{1.121}$$

Hence δu converges to 0 upon iteration, confirming the stability of the fixed point $u_1 = 1$. Coming back to the critical fixed point, and recalling the definition Eq. 1.114, one sees that $u_1 = 0$ corresponds to an infinite value of J/T. In the above framework, this case is interpreted as an infinite coupling limit, as the iteration was made on J. However, the fixed point can also be interpreted as a zero-temperature fixed point, keeping the coupling constant J fixed.

This one-dimensional example is of course only a very simple case, which can be solved through other more direct methods. However, it is a good illustration of the way the concept of RG can be implemented in practice. In two- or three-dimensional models, exact treatments like the above one are most often not available. Yet, many approaches based on different approximation schemes have been developed. A typical situation in dimension $D > 1$ is that there is a finite value K_c of the ratio $K = J/T$ which corresponds to a critical fixed point, and both values $K = 0$ and $K = \infty$ correspond to trivial fixed points, where no correlation is present. Quite importantly, linearizing the iteration equation in the vicinity of the critical fixed point allows the determination of the critical exponent β, as well as other critical exponents. In the Ising chain studied above, this is not possible because the critical temperature is zero, so that there is no extended temperature region where the magnetization is non-zero. But this approach turns out to be relevant in dimension higher than one.

1.5 Disordered Systems and Glass Transition

1.5.1 Disorder in Complex Systems: From Social Sciences to Spin-Glasses

In the general framework of complex systems, disordered systems are systems where each particle or agent has specific properties, which are qualitatively the same for all of them, but differ quantitatively from one to the other. In theoretical models, these quantitative properties are most often drawn from a given probability distribution for each particle or agent, and remain constant in time. Disordered systems should be very relevant in complex systems studies, like social science for instance, as each human being has its specific skills or tastes.

In a physical context, the concept of disordered system could seem less natural, since all particles within a given class are identical. In this case, disorder rather comes from the possibly random position of the particles. A standard example is that of magnetic impurities (that carry a magnetic moment, or spin) diluted in a non-magnetic material. The interaction between magnetic atoms (which have to be described in the framework of quantum mechanics) is mediated by the non-magnetic atoms, and acquires an oscillatory behavior, depending on the distance r_{ij} between the two spins:

$$H = -\sum_{i,j} J(r_{ij}) s_i s_j. \tag{1.122}$$

The interaction constant $J(r_{ij})$ is a given function of the distance r_{ij}, which oscillates around 0, thus taking both positive and negative values. The amplitude of the oscillations decays as a power-law of the distance. As the distances between atoms are random, the interactions between atoms have a random sign, which is the basic property of spin-glasses.

1.5.2 Theoretical Spin-Glass Models

In order to propose a simplified model for spin-glass materials, it has been proposed to replace the position disorder by an interaction disorder, the magnetic atoms being now situated on a regular lattice. To this purpose, one considers an Ising-like model in D-dimensions, where the spins are placed at each node of a lattice. Spins on neighboring sites (i,j) interact through a coupling constant J_{ij}, drawn from a distribution $P(J)$. As the couplings J_{ij} are kept constant in time, one speaks about quenched disorder. This model is called the Edwards–Anderson model [7]. In $D = 1$, the Edwards–Anderson model is equivalent to the standard Ising model, up to a redefinition of the spins. In $D > 1$, analytical solutions are not known, and results have thus been obtained through numerical simulations. A fully connected version, called the Sherrington–Kirkpatrick model, has been proposed and solved, but the techniques involved are already rather difficult, even at the level of the fully connected model. The main qualitative picture emerging from these models is that below a given energy level, the phase space decomposes into a lot of valleys, or metastable states, from which it takes a very long time to escape.

1.5.3 The Simplest Disordered System: The Random Energy Model

A very simple disordered model, which already captures a lot of the phenomenology of realistic disordered systems, has been proposed by Derrida in 1980, and has been called the Random Energy Model (REM) [8, 9]. The model has 2^N configurations, labeled by an index $\alpha = 1, \ldots, 2^N$ (it can be thought of as a spin model, with N spins $s_i = \pm 1$.) To each configuration α is attached a time-independent energy E_α, chosen at random from the distribution

$$P(E) = \frac{1}{\sqrt{N\pi J^2}} \exp\left(-\frac{E^2}{NJ^2}\right). \tag{1.123}$$

All the energies E_α are independent random variables. We denote as $n(E)dE$ the number of configurations with energy in the interval $[E, E + dE]$, so that $n(E)$ is the density of configurations with energy E. The density $n(E)$ is a random quantity, but its fluctuations are small if $n(E)$ is large, namely $n(E) \approx \langle n(E) \rangle$. By definition, $P(E) = \langle n(E) \rangle / 2^N$, so that $\langle n(E) \rangle = 2^N P(E)$, leading to

$$\begin{aligned} \langle n(E) \rangle &= \exp\left(N \ln 2 - \frac{E^2}{NJ^2}\right) \\ &= \exp\left[N\left(\ln 2 - \frac{\varepsilon^2}{J^2}\right)\right] \end{aligned} \tag{1.124}$$

where the energy density $\varepsilon = E/N$ has been introduced. One sees that if $\ln 2 - \varepsilon^2/J^2 > 0$, corresponding to $|\varepsilon| < \varepsilon_0 = J\sqrt{\ln 2}$, $\langle n(E) \rangle$ is exponentially large with N, so that there is a large number of configurations at energy density ε, and the assumption $n(E) \approx \langle n(E) \rangle$ is justified. In contrast, if $\ln 2 - \varepsilon^2/J^2 < 0$, which corresponds to $|\varepsilon| > \varepsilon_0$, $\langle n(E) \rangle$ is exponentially small with N. This means that in most samples, there are no configurations at energy density $|\varepsilon| > \varepsilon_0$. The non-zero, but small value of $\langle n(E) \rangle$ comes from the contribution to the average value of very rare samples, which include some configurations with exceptionally low (or high) energy.

We can now evaluate the partition function of the REM, defined as

$$Z = \sum_{\alpha=1}^{2^N} e^{-E_\alpha/T}. \tag{1.125}$$

As all the energies E_α are random variables, the partition function Z is also a random variable, which fluctuates from one realization of the disorder to another. Yet, we can evaluate the typical value of Z as follows:

$$Z \approx Z_{\text{typ}} = \int_{-\varepsilon_0}^{\varepsilon_0} d\varepsilon \langle \tilde{n}(\varepsilon) \rangle e^{-N\varepsilon/T}, \tag{1.126}$$

with the notation $\tilde{n}(\varepsilon) = Nn(N\varepsilon)$. In Eq. 1.126, we have replaced $\tilde{n}(\varepsilon)$ by $\langle \tilde{n}(\varepsilon) \rangle$ for $|\varepsilon| < \varepsilon_0$, and by 0 for $|\varepsilon| > \varepsilon_0$, consistently with the above discussion. We can then write, using Eqs. 1.124 and 1.126,

$$Z_{\text{typ}} = \int_{-\varepsilon_0}^{\varepsilon_0} d\varepsilon e^{Ng(\varepsilon)} \tag{1.127}$$

with

$$g(\varepsilon) = \ln 2 - \frac{\varepsilon^2}{J^2} - \frac{\varepsilon}{T}. \tag{1.128}$$

In the large N limit, we can evaluate Z_{typ} through a saddle point calculation, namely

$$Z_{\text{typ}} \sim e^{Ng_{\max}(\varepsilon_0)} \tag{1.129}$$

where $g_{\max}(\varepsilon_0)$ is the maximum value of $g(\varepsilon)$ over the interval $[-\varepsilon_0, \varepsilon_0]$. Let us first consider the maximum ε^* of $g(\varepsilon)$ over the entire real line. Taking the derivative of $g(\varepsilon)$, one has

$$g'(\varepsilon) = -\frac{2\varepsilon}{J^2} - \frac{1}{T}. \tag{1.130}$$

From $g'(\varepsilon) = 0$, we find

$$\varepsilon^* = -\frac{J^2}{2T}. \tag{1.131}$$

As $g(\varepsilon)$ is a parabola, it is increasing for $\varepsilon < \varepsilon*$ and decreasing for $\varepsilon > \varepsilon^*$. If $\varepsilon^* > -\varepsilon_0$, then $g_{\max}(\varepsilon_0) = g(\varepsilon^*)$, so that

$$Z_{\text{typ}} \sim e^{Ng(\varepsilon^*)}. \tag{1.132}$$

The condition $\varepsilon^* > -\varepsilon_0$ translates into $T > T_g$, where the so-called glass transition temperature T_g is defined as

$$T_g = \frac{J}{2\sqrt{\ln 2}}. \tag{1.133}$$

For $\varepsilon^* < -\varepsilon_0$, or equivalently $T < T_g$, $g(\varepsilon)$ is a decreasing function of ε over the entire interval $[-\varepsilon_0, \varepsilon_0]$, so that $g_{\max}(\varepsilon_0) = g(-\varepsilon_0)$, and

$$Z_{\text{typ}} \sim e^{Ng(-\varepsilon_0)}. \tag{1.134}$$

From these estimates of Z_{typ}, one can compute the free energy $F = -T \ln Z_{\text{typ}}$, and the entropy $S = -\partial F/\partial T$. For $T > T_g$, one finds

$$F = -N\left(T \ln 2 + \frac{J^2}{4T}\right), \tag{1.135}$$

leading for the entropy to

$$S = N\left(\ln 2 - \frac{J^2}{4T^2}\right). \tag{1.136}$$

For $T < T_g$, we have

$$F = -TNg(-\varepsilon_0) = -TN\left(\ln 2 - \ln 2 + \frac{J}{T}\sqrt{\ln 2}\right) = -NJ\sqrt{\ln 2}. \tag{1.137}$$

The free energy does not depend on temperature in this range, so that the corresponding entropy vanishes:

$$S = 0, \quad T < T_g. \tag{1.138}$$

It can also be checked that the entropy given in Eq. 1.136 for $T > T_g$ vanishes continuously for $T \to T_g$. Hence the temperature T_g corresponds to a glass transition temperature, where the entropy goes to zero when lowering temperature down to T_g, and remains zero below T_g. Actually, to make the statement sharper, only the entropy density S/N goes to zero for $T < T_g$, in the infinite N limit. Computing subleading corrections to the entropy, one finds that the entropy S is independent of N, but non-zero, for $T < T_g$. The entropy is then intensive in this temperature range, meaning that only a finite number of configurations, among the 2^N ones a priori available, are effectively occupied: the system is quenched in the lowest energy configurations.

References

1. R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975)
2. D.A. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976)
3. T.L. Hill, *An Introduction to Statistical Thermodynamics* (Dover, New York, 1986)
4. D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987)
5. H. Goldstein, *Classical Mechanics*, 2nd edn. (Addison-Wesley, Reading, 1980)
6. D.R. Nelson, M.E. Fisher, Ann. Phys. **91**, 226 (1975)
7. S.F. Edwards, P.W. Anderson, J. Phys. F **5**, 965 (1975)
8. B. Derrida, Phys. Rev. Lett. **45**, 79 (1980)
9. B. Derrida, Phys. Rev. B **24**, 2613 (1981)

Chapter 2
Non-stationary Dynamics and Stochastic Formalism

In the first part of this book, we have considered the stationary properties of physical systems composed of a large number of particles, using as fundamental statistical object the joint distribution of all the degrees of freedom of the system (for instance positions and velocities, or spin variables). This steady state is expected to be reached after a transient regime, during which the macroscopic properties of the system evolve with time. Describing the statistical state of the system during this transient regime is also certainly of interest.

However, there is no known simple postulate (similar to the postulate of equiprobability of configurations having a given energy) to characterize the N-particle probability distribution in this time-dependent regime. Still, one can resort to the generic mathematical formalism of stochastic processes in order to describe statistically the time evolution of some specific variables of interest, like the position or velocity of a probe particle immersed in a fluid. This formalism is presented in Sect. 2.1, in the simplest case of Markov processes. The example of the random evolution of a single degree of freedom in a noisy environment (diffusive motion), leading to the Langevin and Fokker–Planck equations, is discussed in Sect. 2.2. In addition, there exists situations in which this random evolution can be much faster or much slower than a priori expected, leading to anomalous diffusion. It may also happen that some systems fail to reach an equilibrium state, and thus keep relaxing for a very long time, as in the case of glasses. These situations, that bear some interesting relations to the central limit theorem and its generalization, are described in Sect. 2.3.

2.1 Markovian Stochastic Processes and Master Equation

2.1.1 Definition of Markovian Stochastic Processes

Let us start with some basic considerations on stochastic processes. For more advance reading on this topic, we refer the reader to [1] for instance. Roughly speaking, a

E. Bertin, *A Concise Introduction to the Statistical Physics of Complex Systems*,
SpringerBriefs in Complexity, DOI: 10.1007/978-3-642-23923-6_2,

stochastic process is a dynamical process whose evolution is random, and depends on the presently occupied state and possibly on the history of the system.

Considering first a discrete time process ($t = 0, 1, 2, \ldots$), with a finite number N of configurations C, we denote as $T(C_{t+1}|C_t, C_{t-1}, \ldots, C_0)$ the probability for the process to jump to a new configuration C_{t+1} between times t and $t+1$, given the whole history $(C_t, C_{t-1}, \ldots, C_0)$. Note that C_{t+1} can a priori be any of the N possible configurations, including the configuration C_t itself. The transition probability $T(C_{t+1}|C_t, C_{t-1}, \ldots, C_0)$ can be considered as a conditional probability, so that the following normalization condition holds

$$\sum_{C_{t+1}} T(C_{t+1}|C_t, C_{t-1}, \ldots, C_0) = 1. \tag{2.139}$$

Such a stochastic process is said to be Markovian if the transition probability $T(C_{t+1}|C_t, C_{t-1}, \ldots, C_0)$ depends only on the configuration C_t occupied at time t, and not on previously occupied configurations. In short, markovian processes are said to be "memoryless".

The above definition of markovian stochastic processes can be rather straightforwardly extended to several other cases of practical importance. First, the number of discrete configurations can be infinite, and this case can often be recovered by tacking the limit $N \to \infty$ in the above definition. If configurations are no longer discrete, but are defined by a continuous variable y_t, a probability density $\tilde{T}(y_{t+1}|y_t)$ needs to be introduced, in such a way that $\tilde{T}(y_{t+1}|y_t)dy_{t+1}$ is the probability to choose a new configuration in the interval $[y_{t+1}, y_{t+1}+dy_{t+1}]$, starting from a given configuration y_t at time t. The equivalent of the normalization condition Eq. 2.139 now reads

$$\int_{-\infty}^{\infty} \tilde{T}(y_{t+1}|y_t)dy_{t+1} = 1. \tag{2.140}$$

Another generalization consists in replacing the discrete time steps by a continuous time evolution. Interestingly, continuous time dynamics can be obtained from the discrete time dynamics in the limit of a vanishing time step. Hence instead of using a time step $\Delta t = 1$ as above, we now take an infinitesimal step $\Delta t = dt$. In order to obtain a meaningful limit when $dt \to 0$, the transition probabilities $T(C'|C)$ from configuration C to configuration C' have to scale with dt in the following way:

$$T(C'|C) = W(C'|C)dt + \mathcal{O}(dt^2) \quad \text{if } C' \neq C \tag{2.141}$$

$$T(C|C) = 1 - \sum_{C''(\neq C)} W(C''|C)dt + \mathcal{O}(dt^2) \tag{2.142}$$

where $W(C'|C)$ is independent of dt. In other words, the evolution of continuous time Markovian stochastic processes is characterized by transition rates $W(C'|C)$, such that $W(C'|C)dt$ is the probability for the process to go from configuration C to a new configuration C' in a time interval $[t, t+dt]$.

Finally, in the case of a continuous time process represented by a continuous variable y, a density of transition rate $w(y'|y)$ should be defined, a such a way that $w(y'|y)dy'dt$ is the probability for the process to reach a value in the interval $[y', y' + dy']$ at time $t + dt$, starting from a value y at time t.

2.1.2 Master Equation and Detailed Balance

The master equation describes the time evolution of the probability to occupy a given configuration. The simplest situation corresponds to discrete time and discrete configurations. The evolution of the probability $P_t(C)$ to occupy configuration C at time t is given by

$$P_{t+1}(C) = \sum_{C'} T(C|C')P_t(C'). \tag{2.143}$$

The probability $P_{t+1}(C)$ is thus simply a sum over all possible configurations C' of the probability to go from C' to C, weighted by the probability to occupy the configuration C' at time t. It is easy to check, by summing over all configurations C and using the normalization equation (2.139), that Eq. 2.143 conserves the normalization of the probability $P_t(C)$; namely, if $\sum_C P_t(C) = 1$, then $\sum_C P_{t+1}(C) = 1$.

For continuous configurations y, a density $p_t(y)$ has to be introduced (namely, $p_t(y)dy$ is the probability that the configuration at time t belongs to the interval $[y, y + dy]$), and the evolution equation reads:

$$p_{t+1}(y) = \int_{-\infty}^{\infty} \tilde{T}(y|y')p_t(y')dy'. \tag{2.144}$$

The evolution of continuous time processes can be derived from this discrete time equation, using again the limit of a vanishing time step dt. Considering a continuous time process with discrete configurations, we denote as $P(C, t)$ the probability to be in configuration C at time t. Combining Eqs. 2.141 and 2.143, we get

$$P(C, t + dt) = \sum_{C'(\neq C)} W(C|C')dt\,P(C', t) + \left(1 - \sum_{C'(\neq C)} W(C'|C)dt\right) P(C, t). \tag{2.145}$$

Expanding the left-hand-side of this last equation to first order in dt, as

$$P(C, t + dt) = P(C, t) + \frac{dP}{dt}(C, t)dt + \mathcal{O}(dt^2) \tag{2.146}$$

we eventually find, in the limit $dt \to 0$, that the probability $P(C, t)$ evolves according to the master equation:

$$\frac{dP}{dt}(C,t) = -P(C,t) \sum_{C'(\neq C)} W(C'|C) + \sum_{C'(\neq C)} W(C|C')P(C',t). \tag{2.147}$$

The first term in the right-hand-side can be interpreted as a "loss" term (i.e., the sum of all the possibilities to exit configuration *C*), while the second term can be thought of as a "gain" term (the sum of all the possibilities to arrive at configuration *C*, starting from any other configuration). A similar equation is obtained in the case of continuous configurations y for the probability density $p(y,t)$, by basically replacing discrete sums by integrals in Eq. 2.147:

$$\frac{\partial p}{\partial t}(y,t) = -p(y,t) \int_{-\infty}^{\infty} dy' w(y'|y) + \int_{-\infty}^{\infty} dy' w(y|y') p(y',t). \tag{2.148}$$

From now on, we will work mainly with discrete configurations as far as formal and generic calculations are concerned, keeping in mind that the continuous variable case can be obtained by switching from discrete to continuous notations.

An interesting property of continuous time master equations is the notion of detailed balance, which is related to the steady-state (or time-independent) solution of the master equation. From Eq. 2.147, a time-independent solution *P(C)* satisfies, for all configurations *C*

$$\sum_{C'(\neq C)} [-W(C'|C)P(C) + W(C|C')P(C')] = 0. \tag{2.149}$$

It may happen, for some specific stochastic processes, that the term between bracket vanishes for all C', namely

$$\forall (C, C'), \qquad W(C'|C)P(C) = W(C|C')P(C'). \tag{2.150}$$

This situation is referred to as detailed balance. Processes satisfying detailed balance are much easier to handle analytically. Besides this practical advantage, detailed balance also plays an important role in the stochastic modeling of microscopic physical processes (i.e., at the molecular scale). This is due to the fact that detailed balance can be interpreted as the stochastic counterpart of the microreversibility property satisfied by the Hamiltonian dynamics—see Sect. 1.1. Indeed, the probability to observe, once a statistical steady-state is reached, an elementary trajectory from C at time t to C' at time $t + dt$ is $W(C'|C)dt P(C)$, while the probability to observe the reverse trajectory is $W(C|C')dt P(C')$. The equality of these two probabilities, to be thought of as a statistical microreversibility, precisely yields the detailed balance relation (2.150). Hence in order to model, at a coarse-grained level, the dynamics of a microscopic physical system through a markovian stochastic process, it is natural to assume that the process satisfies detailed balance (in addition to the appropriate conservation laws, like energy conservation).

2.1.3 Dynamical Increase of the Entropy

It is interesting to introduce a time-dependent entropy defined as

$$S(t) = -\sum_C P(C,t)\ln P(C,t). \tag{2.151}$$

This definition closely follows the definition (1.70). Under the assumption $W(C'|C) = W(C|C')$, which is a specific form of the detailed balance relation (2.150), one can show that *S(t)* is increasing with time. Let us start by computing the time-derivative of the entropy:

$$\frac{dS}{dt} = -\sum_C \frac{dP}{dt}(C,t)\ln P(C,t) - \sum_C \frac{dP}{dt}(C,t). \tag{2.152}$$

The last term cancels due to the normalization condition $\sum_C P(C,t) = 1$. Using the master equation, one has:

$$\begin{aligned}\frac{dS}{dt} &= -\sum_C \ln P(C,t) \sum_{C'(\neq C)} \big(-W(C'|C)P(C,t) + W(C|C')P(C',t)\big)\\ &= \sum_{C,C'(C\neq C')} \ln P(C,t)\big(W(C'|C)P(C,t) - W(C|C')P(C',t)\big).\end{aligned} \tag{2.153}$$

Exchanging the notations C and C' in the last equation, we also have

$$\frac{dS}{dt} = \sum_{C,C'(C\neq C')} \ln P(C',t)\big(W(C|C')P(C',t) - W(C'|C)P(C,t)\big). \tag{2.154}$$

Summing Eqs. 2.153 and 2.154, and using the detailed balance property $W(C'|C) = W(C|C')$, we obtain

$$\frac{dS}{dt} = \frac{1}{2}\sum_{C,C'(C\neq C')} \big(\ln P(C',t) - \ln P(C,t)\big)\big(P(C',t) - P(C,t)\big)\,W(C|C'). \tag{2.155}$$

As $[\ln P(C',t) - \ln P(C,t)]$ and $[P(C',t) - P(C,t)]$ have the same sign, one concludes that

$$\frac{dS}{dt} \geq 0. \tag{2.156}$$

This is one possible statement, in the context of stochastic processes, of the second law of thermodynamics. Moreover, in the stationary state, $dS/dt = 0$, and one necessarily has for all pairs (C, C') either $P_{\mathrm{st}}(C) = P_{\mathrm{st}}(C')$ or $W(C|C') = 0$, where $P_{\mathrm{st}}(C)$ is the stationary probability distribution. One then recovers, in a stochastic framework, the fundamental postulate of equilibrium statistical mechanics, stating that mutually accessible configurations have the same probability.

More generally, for Markovian stochastic processes described by the master equation (2.147), it is always possible to define a functional $\tilde{S}(\{P(C,t)\})$ that increases with time, without need for detailed balance or microreversibility properties [1]. The general definition of $\tilde{S}$ is

$$\tilde{S}(t) = -\sum_C P(C,t) \ln \left(\frac{P(C,t)}{P_{\mathrm{st}}(C)} \right). \tag{2.157}$$

However, it turns out that the stationary distribution needs to be known in order to define $\tilde{S}$, which in many cases restricts the usefulness of the functional $\tilde{S}$.

2.1.4 A Simple Example: The One-Dimensional Random Walk

A simple and illustrative example of stochastic process is the one-dimensional random walk, where a "particle" moves at random on a discretized line. Let us consider first the discrete time case: a particle can take only discrete positions $x = \ldots, -2, -1, 0, 1, 2, \ldots$ on a line. Between times t and $t+1$, the particle randomly jumps to one of the two neighboring sites, so that $x_{t+1} = x_t + \epsilon_t$, with $\epsilon_t = \pm 1$ with equal probabilities. The random variables ϵ_t and $\epsilon_{t'}$, with $t \neq t'$, are independent and identically distributed.

The average value and the variance of this process can be derived straightforwardly. We first note that $\langle x_{t+1} \rangle = \langle x_t \rangle$, so that $\langle x_t \rangle = \langle x_0 \rangle$ for all t (the notation $\langle \ldots \rangle$ denotes an ensemble average, that is an average over a very large number of samples of the same process; it may thus depend on time). For instance, if the walk starts with probability 1 from $x_0 = 0$, then all subsequent averages $\langle x_t \rangle = 0$.

Let us now compute the variance of the process, defined as

$$\mathrm{Var}(x_t) = \langle x_t^2 \rangle - \langle x_t \rangle^2. \tag{2.158}$$

We assume for simplicity that $\langle x_t \rangle = 0$, so that $\mathrm{Var}(x_t) = \langle x_t^2 \rangle$ (the generalization to $\langle x_t \rangle \neq 0$ is however straightforward). From $x_{t+1} = x_t + \epsilon_t$, we get

$$x_{t+1}^2 = x_t^2 + 2x_t\epsilon_t + 1, \tag{2.159}$$

taking into account that $\epsilon_t^2 = 1$. Computing the ensemble average of Eq. 2.159 yields

$$\langle x_{t+1}^2 \rangle = \langle x_t^2 \rangle + 2\langle x_t \rangle \langle \epsilon_t \rangle + 1, \tag{2.160}$$

using the fact that x_t depends only on $\epsilon_{t'}$ with $t' < t$, so that x_t and ϵ_t are independent random variables. As $\langle \epsilon_t \rangle = 0$, it follows that $\langle x_{t+1}^2 \rangle = \langle x_t^2 \rangle + 1$, so that $\langle x_t^2 \rangle =$

$\langle x_0^2 \rangle + t$. If $x_0 = 0$ with probability 1, one has $\langle x_0^2 \rangle = 0$, and $\langle x_t^2 \rangle = t$. This means that the typical position reached by the walk after t steps is of the order of $\sqrt{t}$.

The present random walk problem bears a direct relationship to the central limit theorem [2, 3]. The central limit theorem states that a sum of independent and identically distributed (iid) random variables, once conveniently rescaled to have zero mean and unit variance, converges to a Gaussian distribution. Namely, if $y = \sum_{i=1}^{N} u_i$ is a sum of iid random variables of variance σ_u^2, the rescaled variable

$$z = \frac{y - N\langle u \rangle}{\sigma_u \sqrt{N}} \tag{2.161}$$

converges to the normal distribution

$$P(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}. \tag{2.162}$$

As the position x_t of the random walk can be expressed as $x_t = \sum_{t'=0}^{t-1} \epsilon_{t'}$, the distribution of the position of the random walk can be approximated for a large time t, using the central limit theorem, as

$$P(x, t) \approx \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}. \tag{2.163}$$

Another popular formulation of the random walk problem is the continuous time dynamics. Labeling with an integer n the sites of the lattice, the transition rate $W(n'|n)$ from site n to site n' is given by

$$W(n'|n) = \begin{cases} \frac{\nu}{2} & \text{if } n' = n \pm 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.164}$$

where ν is a characteristic frequency (the inverse of a time scale) of the process. The master equation reads

$$\frac{dP_n}{dt} = -\sum_{n'(\neq n)} W(n'|n) P_n(t) + \sum_{n'(\neq n)} W(n|n') P_{n'}(t). \tag{2.165}$$

Replacing the transition rates by their expression given in Eq. 2.164, one finds

$$\frac{dP_n}{dt} = -\nu P_n(t) + \frac{\nu}{2} P_{n+1}(t) + \frac{\nu}{2} P_{n-1}(t). \tag{2.166}$$

The evolution of the probability distribution $P_n(t)$ can be evaluated from Eq. 2.166, for instance by integrating it numerically. However, one may be interested in making analytical predictions in the large time limit, and such a discrete-space equation is not easy to handle in this case. To this aim, it is thus useful to use a procedure called "continuous limit", through which the discrete-space equation (2.166) can be approximated by a continuous space equation, namely, a partial differential equation. To be more specific, let us call a the lattice spacing (which was set above to $a = 1$). At large time $t \gg 1/\nu$, the distribution $P_n(t)$ is expected to vary over spatial scales much larger than the lattice spacing a; in other words, one has

$$|P_{n+1}(t) - P_n(t)| \ll P_n(t). \tag{2.167}$$

Plotting $P_n(t)$ as a function of space, it thus appear essentially continuous. We thus postulate the existence of a distribution $p(x,t)$ of the continuous variable x, such that the discrete-space distribution can be approximated as $P_n(t) \approx ap(na, t)$. The prefactor a is included to ensure a correct normalization, $\sum_n P_n(t) = 1$. Indeed, one has for $a \to 0$

$$\sum_n P_n(t) = a \sum_n p(na, t) \to \int_{-\infty}^{\infty} p(x, t)dx. \tag{2.168}$$

For consistency, it is thus necessary to assume that $p(x, t)$ is normalized such that $\int_{-\infty}^{\infty} p(x, t)dx = 1$.

Replacing $P_n(t)$ by $ap(na, t)$ in the master equation (2.166), one obtains

$$\frac{\partial p}{\partial t}(x, t) = -\nu p(x, t) + \frac{\nu}{2} p(x + a, t) + \frac{\nu}{2} p(x - a, t). \tag{2.169}$$

As a is small, one can expand $p(x \pm a, t)$ to second order in a, leading to

$$p(x \pm a, t) = p(x, t) \pm a \frac{\partial p}{\partial x}(x, t) + \frac{a^2}{2} \frac{\partial^2 p}{\partial x^2}(x, t) + \mathcal{O}(a^3). \tag{2.170}$$

The linear terms in a appearing in Eq. 2.169 cancel out, so that this equation reduces to

$$\frac{\partial p}{\partial t}(x, t) = \frac{\nu a^2}{2} \frac{\partial^2 p}{\partial x^2}(x, t) \tag{2.171}$$

which is called the diffusion equation. This equation appears in numerous problems in physics, like the diffusion of heat in a material, or the diffusion of dye in water for instance. The coefficient $\frac{1}{2}\nu a^2$ has to take a finite value $D > 0$ for the equation to be well-defined. As the lattice spacing a goes to zero, it is thus necessary that ν simultaneously goes to infinity, which means that the 'microscopic' process appears very fast on the scale of observation.

Equation 2.171 has several simple solutions of interest. For instance, if the diffusing particle is bound to stay on a segment $[-L, L]$, the long-time limit distribution is a flat and time-independent distribution over the segment, $p(x) = (2L)^{-1}$. In other words, diffusion tends to flatten, or smoothen, the distribution. In contrast, if the particle can diffuse on the entire line without bound, the distribution $p(x, t)$

never reaches a steady-state regime, but rather enters a scaling regime in which the distribution keeps broadening with time, with a well-define Gaussian shape:

$$p(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}. \tag{2.172}$$

Note the analogy with the result obtained from the central limit theorem in the discrete time case—see Eq. 2.163.

2.2 Langevin and Fokker–Planck Equations

2.2.1 Phenomenological Approach to the Langevin Equation

The above random walk example was quite simple to investigate, but had little explicit connection with physical systems. We now present another standard example based on a physical phenomenology. Let us imagine a probe particle immersed in a fluid, such that the size of the particle is small at the macroscopic scale, but still much larger than the typical size of the molecules of the fluid. For the sake of simplicity, we restrict the presentation to a one-dimensional system, but the more realistic three-dimensional situation would follow the same line.

We choose the mass of the probe particle as the unit mass. The acceleration of the particle is then governed by the force F_{coll} exerted by the collisions with the other particles:

$$\frac{dv}{dt} = F_{\text{coll}}, \tag{2.173}$$

where v is the velocity of the probe particle. Since the collisional force is strongly fluctuating, the basic idea is to decompose F_{coll} into an average force which depends on the velocity, and a purely fluctuating (or noise) part:

$$F_{\text{coll}} = \langle F_{\text{coll}} \rangle + \xi(t). \tag{2.174}$$

Here also, the ensemble average $\langle \ldots \rangle$ is computed as an average over a large number of samples of the process, so that F_{coll} a priori depends on time. By definition, the noise $\xi(t)$ has zero mean, $\langle \xi(t) \rangle = 0$. To proceed further, it is necessary to choose a specific model for both $\langle F_{\text{coll}} \rangle$ and $\xi(t)$. The average force $\langle F_{\text{coll}} \rangle$ can be interpreted as an effective friction force, which slows down the probe particle; it is thus natural to choose, as a first approximation, a linear friction force $\langle F_{\text{coll}} \rangle = -\gamma v$, with $\gamma > 0$ a friction coefficient.

Then, a model of the noise should be given. Beside the property $\langle \xi(t) \rangle = 0$, its two-time correlation should be specified. Intuitively, one expects collisions occurring at different times to be essentially uncorrelated, so that one should have $\langle \xi(t)\xi(t') \rangle = 0$

for $|t - t'| \gg \tau_{\mathrm{col}}$, where τ_{col} is the typical duration of a collision. To emphasize that $\xi(t)$ decorrelates over a time scale τ_{col}, we write the correlation in the form

$$\langle \xi(t)\xi(t') \rangle = C\left(\frac{|t - t'|}{\tau_{\mathrm{col}}}\right), \tag{2.175}$$

where the function $C(u)$ converges rapidly to zero when $u \to \infty$.

Altogether, Eq. 2.173 can be rewritten as:

$$\frac{dv}{dt} = -\gamma v + \xi(t). \tag{2.176}$$

Such an equation is called a linear Langevin equation. For simplicity, we take as initial condition a fixed value $v(0) = v_0$. We first note, computing the ensemble average of Eq. 2.176:

$$\frac{d}{dt}\langle v(t) \rangle = -\gamma \langle v(t) \rangle, \tag{2.177}$$

that the ensemble-averaged velocity $\langle v(t) \rangle$ obeys the same equation as the non-averaged velocity, except that noise is now absent. This property is specific to the linear Langevin equation, and would not be present if we had included a non-linear dependence on v in the friction force—e.g., $\langle F_{\mathrm{coll}} \rangle = -\gamma v - \gamma_3 v^3$. The solution of Eq. 2.177 is a decaying exponential:

$$\langle v(t) \rangle = v_0 e^{-\gamma t}. \tag{2.178}$$

More interestingly, the effect of the noise has a deep impact on the evolution of the variance of the velocity, $\mathrm{Var}[v(t)] = \langle v(t)^2 \rangle - \langle v(t) \rangle^2$. In order to compute $\langle v(t)^2 \rangle$, we first determine the explicit time-dependence of $v(t)$, considering $\xi(t)$ as an arbitrary given function. Following standard mathematical methods, the general solution of Eq. 2.176 is given by the sum of the general solution of the homogeneous equation (i.e., the noiseless equation) and of a particular solution of the full equation. The general solution of the homogeneous equation is $v_{\mathrm{h}}(t) = A_0 e^{-\gamma t}$, where A_0 is an arbitrary constant. In order to determine a particular solution, one can use the so-called "variation of the constant" method, which indicates that such a solution should be searched in the form $v_{\mathrm{p}}(t) = A(t)e^{-\gamma t}$, that is, simply replacing the constant A_0 in the solution $v_{\mathrm{h}}(t)$ of the homogeneous equation by a function $A(t)$ to be determined. Inserting $v_{\mathrm{p}}(t)$ in Eq. 2.176, we get

$$\frac{dA}{dt} e^{-\gamma t} = \xi(t) \tag{2.179}$$

whence the solution

$$A(t) = \int_0^t e^{\gamma t'} \xi(t') dt' \tag{2.180}$$

follows—since we look for a particular solution at this stage, there is no need to add a constant term to Eq. 2.180. Altogether, one finds for $v(t)$, taking into account the initial condition $v(0) = v_0$,

$$v(t) = v_0 e^{-\gamma t} + e^{-\gamma t} \int_0^t e^{\gamma t'} \xi(t') dt'. \tag{2.181}$$

Computing $v(t)^2$ yields

$$v(t)^2 = v_0^2 e^{-2\gamma t} + e^{-2\gamma t} \left(\int_0^t e^{\gamma t'} \xi(t') dt' \right)^2 + 2v_0 e^{-2\gamma t} \int_0^t e^{\gamma t'} \xi(t') dt'. \tag{2.182}$$

Now taking an ensemble average, the last term vanishes because $\langle \xi(t) \rangle = 0$, and we get

$$\langle v(t)^2 \rangle = v_0^2 e^{-2\gamma t} + \left\langle e^{-2\gamma t} \left(\int_0^t e^{\gamma t'} \xi(t') dt' \right)^2 \right\rangle. \tag{2.183}$$

The first term on the right hand side is precisely $\langle v(t) \rangle^2$, so that

$$\mathrm{Var}[v(t)] = \left\langle e^{-2\gamma t} \left(\int_0^t e^{\gamma t'} \xi(t') dt' \right)^2 \right\rangle. \tag{2.184}$$

The square of the integral can be expanded as a product of two integrals, which in turn can be converted into a double integral:

$$\left(\int_0^t e^{\gamma t'} \xi(t') dt' \right)^2 = \int_0^t e^{\gamma t'} \xi(t') dt' \int_0^t e^{\gamma t''} \xi(t'') dt'' \tag{2.185}$$

$$= \int_0^t dt' \int_0^t dt'' e^{\gamma (t'+t'')} \xi(t') \xi(t'') \tag{2.186}$$

so that Eq. 2.184 eventually turns into

$$\mathrm{Var}[v(t)] = e^{-2\gamma t} \int_0^t dt' \int_0^t dt'' e^{\gamma (t'+t'')} \langle \xi(t') \xi(t'') \rangle \tag{2.187}$$

(we recall that the ensemble average can be interchanged with linear operations like integrals or derivatives). Using the expression (2.175) of $\langle \xi(t') \xi(t'') \rangle$, we get

$$\mathrm{Var}[v(t)] = e^{-2\gamma t} \int_0^t dt' \int_0^t dt'' e^{\gamma(t'+t'')} C\left(\frac{|t'-t''|}{\tau_{\mathrm{col}}}\right). \tag{2.188}$$

It is useful to make a change of variable here, replacing t'' by the variable $y = t'' - t'$ in the second integral, which yields

$$\mathrm{Var}[v(t)] = e^{-2\gamma t} \int_0^t dt' e^{2\gamma t'} \int_{-t'}^{t-t'} dy e^{\gamma y} C\left(\frac{|y|}{\tau_{\mathrm{col}}}\right). \tag{2.189}$$

Considering the limit where $\tau_{\mathrm{col}} \ll \gamma^{-1}$ (limit of very short collision time), the exponential term in the integral can be regarded as a constant, $e^{\gamma y} \approx 1$, on the range of y over which $C(|y|/\tau_{\mathrm{col}})$ takes significant values. Note that we have assumed above a “rapid” decay of $C(u)$ for large u; this assumption can now be made more precise, as the decay of $C(u)$ has to dominate over the divergence of the term $e^{\gamma y}$ for $y \to \infty$. A Gaussian tail $C(u) \sim e^{-\alpha u^2}$ ($\alpha > 0$) would for instance be suitable here.

Coming back to the evaluation of the integral, the above approximation leads to

$$\int_{-t'}^{t-t'} dy e^{\gamma y} C\left(\frac{|y|}{\tau_{\mathrm{col}}}\right) \approx \int_{-\infty}^{\infty} dy C\left(\frac{|y|}{\tau_{\mathrm{col}}}\right), \tag{2.190}$$

where we have also approximated the original integral by its extension over the whole real axis, which is justified by fast decay of $C(u)$. Denoting as Γ the integral[1]

$$\Gamma \equiv \int_{-\infty}^{\infty} dy C\left(\frac{|y|}{\tau_{\mathrm{col}}}\right), \tag{2.191}$$

we eventually find

$$\mathrm{Var}[v(t)] = e^{-2\gamma t} \Gamma \int_0^t dt' e^{2\gamma t'} \tag{2.192}$$

which is readily integrated into

$$\mathrm{Var}[v(t)] = \frac{\Gamma}{2\gamma}\left(1 - e^{-2\gamma t}\right). \tag{2.193}$$

Hence the variance starts from a zero value at $t = 0$ (the value v_0 at $t = 0$ is non-random), and progressively grows until reaching the asymptotic limit $\Gamma/(2\gamma)$. As

---

[1] Note that from this definition, Γ is of the order of $C(0)\tau_{\mathrm{col}}$. For Γ to be finite in the limit of a small τ_{col}, one also needs to assume that $C(0) \propto 1/\tau_{\mathrm{col}}$.

$\langle v(t)\rangle \to 0$ when $t \to \infty$, the variance reduces to $\langle v^2\rangle$ at large time, and this value can be identified with the equilibrium average. It is known from equilibrium statistical physics (see Sect. 1.2.4) that $\langle \frac{1}{2}v^2\rangle_{\text{eq}} = \frac{1}{2}k_BT$ (equipartition relation), where T is the temperature of the surrounding liquid –we recall that the mass of the probe particle was set to unity. Hence equilibrium statistical physics imposes a relation between the two phenomenologically introduced coefficients Γ and γ, namely $\Gamma = 2\gamma k_BT$.

For pedagogical purposes, we have used here a correlation $\langle \xi(t)\xi(t')\rangle$ of the noise that decays over a time scale τ_{col}, considered to be small with respect to γ^{-1}. However, in standard presentations of the Langevin equation, use is made of a mathematical tool called the Dirac distribution $\delta(x)$, which can be thought of as a function being equal to zero for all $x \neq 0$, and being infinite for $x = 0$, and such that $\int_{-\infty}^{\infty} \delta(x)dx = 1$. The correlation $\langle \xi(t)\xi(t')\rangle$ is then usually written as

$$\langle \xi(t)\xi(t')\rangle = \Gamma\delta(t - t'). \tag{2.194}$$

The main interest of the Dirac distribution is that for an arbitrary function f,

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0). \tag{2.195}$$

In other words, once inserted in an integral, the Dirac distribution precisely picks up the value of the integrand associated to the value of the variable around which it is peaked. Hence the Dirac distribution is a tool which gives a mathematical meaning to the limit $\tau_{\text{col}} \to 0$, and which thus simplifies formal calculations. In the above presentation, we chose to keep a small but finite τ_{col} in order to give the reader a better intuition of the physics behind the calculations, and to avoid possible difficulties arising from the use of a new mathematical object.

2.2.2 *Relation to Random Walks*

After having introduced the Langevin equation from a physical perspective (that of a probe particle immersed in a fluid), it is interesting to present the Langevin equation from another perspective, that of random walks. To this aim, we come back to the random walk model introduced in Sect. 2.1.4 and generalize it by including a small bias in the displacements. We consider a discrete time dynamics with a time step Δt, and we call a the lattice spacing. At time $t + \Delta t$, the new position $x_{t+\Delta t}$ is chosen according to $x_{t+\Delta t} = x_t + \epsilon_t$, where ϵ_t is given by

$$\epsilon_t = \begin{cases} a & \text{with prob.} \quad \frac{\nu}{2}\big(1 + aq(x_t)\big)\Delta t \\ -a & \text{with prob.} \quad \frac{\nu}{2}\big(1 - aq(x_t)\big)\Delta t \\ 0 & \text{with prob.} \quad 1 - \nu\Delta t \end{cases} \tag{2.196}$$

Note that the above dynamical rules can be interpreted as a discretized version of a continuous time dynamics, as seen from the presence of the time step Δt and from the allowed value $\epsilon_t = 0$. Let us define $\Delta x_t \equiv x_{t+\Delta t} - x_t$. The dynamical rules $x_{t+\Delta t} = x_t + \epsilon_t$ can be rewritten as

$$\frac{\Delta x_t}{\Delta t} = \frac{\epsilon_t}{\Delta t} \tag{2.197}$$

which is the analog of Eq. 2.173, provided that x_t is interpreted as a velocity; $\epsilon_t / \Delta t$ then plays the role of a random force. Computing the average value of this 'force', we find using Eq. 2.196

$$\left\langle \frac{\epsilon_t}{\Delta t} \right\rangle = a^2 \nu q(x_t). \tag{2.198}$$

Note that the average is taken over ϵ_t, for a fixed value of x_t. Let us now consider the fluctuating part of the 'force', and define

$$\xi_t = \frac{1}{\Delta t}(\epsilon_t - \langle \epsilon_t \rangle), \tag{2.199}$$

which is thus the discrete-time analog of $\xi(t)$ introduced in Sect. 2.2.1. We wish to evaluate the correlation of ξ_t, given by

$$\langle \xi_t \xi_{t'} \rangle = \frac{1}{(\Delta t)^2} \Big\langle (\epsilon_t - \langle \epsilon_t \rangle)(\epsilon_{t'} - \langle \epsilon_{t'} \rangle) \Big\rangle. \tag{2.200}$$

For $t \neq t'$, $\langle \xi_t \xi_{t'} \rangle$ is thus equal to zero, as ϵ_t and $\epsilon_{t'}$ are independent random variables. If $t = t'$, one has $\langle \xi_t \xi_{t'} \rangle = \mathrm{Var}(\epsilon_t)/(\Delta t)^2$. Both cases can be encompassed in a single expression, introducing (k, k') through $t = k\Delta t$ and $t' = k'\Delta t$:

$$\langle \xi_t \xi_{t'} \rangle = \frac{1}{(\Delta t)^2} \mathrm{Var}(\epsilon_t) \delta_{k,k'} \tag{2.201}$$

where $\delta_{k,k'}$ is the Kronecker delta symbol, equal to 1 if $k = k'$ and to zero otherwise. Evaluating the variance of ϵ_t, we find

$$\mathrm{Var}(\epsilon_t) = a^2 \nu \Delta t + \mathcal{O}(\Delta t^2), \tag{2.202}$$

so that to leading order in Δt,

$$\langle \xi_t \xi_{t'} \rangle = a^2 \nu \frac{\delta_{k,k'}}{\Delta t}. \tag{2.203}$$

This expression is the analog of Eq. 2.194, and the role played by τ_{col} in the physical approach to the Langevin equation is now played by Δt. To give further evidence for this correspondence, we point out that $\delta_{k,k'} / \Delta t$ can be interpreted as the discretized version of the Dirac distribution. Indeed, from the definition of the Kronecker delta symbol, one can write for an arbitrary function f

$$\sum_{k'=-\infty}^{\infty} \Delta t f(k'\Delta t)\frac{\delta_{k,k'}}{\Delta t} = f(k\Delta t), \tag{2.204}$$

which is precisely the discretized version of the fundamental property (2.195) of the Dirac delta function. Hence taking the limit $\Delta t \to 0$ and $a \to 0$, one can reformulate the above biased random walk problem as a Langevin equation, namely

$$\frac{dx}{dt} = Q(x) + \xi(t) \tag{2.205}$$

where $Q(x) \equiv a^2 \nu q(x)$, and where the noise $\xi(t)$ satisfies

$$\langle \xi(t) \rangle = 0, \qquad \langle \xi(t)\xi(t') \rangle = \Gamma \delta(t - t'). \tag{2.206}$$

2.2.3 Fokker–Planck Equation

The Fokker–Planck equation basically describes the evolution of the probability distribution $p(x, t)$ of a variable x obeying a Langevin equation. It can be derived in several ways, one of the simplest being to start from the above biased random walk problem, and to derive the continuous limit of the master equation, following the same lines as for the derivation of the diffusion equation—see Sect. 2.1.4.

Starting from the biased random walk model of Sect. 2.2.2, we consider the continuous time version of the model, and write the corresponding transition rates $W(n'|n)$, where $n = x/a$ is a integer labeling the sites of the one-dimensional lattice:

$$W(n'|n) = \begin{cases} \frac{\nu}{2}(1 + aq_n) & \text{if } n' = n + 1 \\ \frac{\nu}{2}(1 - aq_n) & \text{if } n' = n - 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.207}$$

To lighten the notations, we have denoted $q(na)$ as q_n. Formally, one can write the transition rates as

$$W(n'|n) = \frac{\nu}{2}(1 + aq_n)\delta_{n',\, n+1} + \frac{\nu}{2}(1 - aq_n)\delta_{n',\, n-1}. \tag{2.208}$$

The master equation then reads

$$\frac{dP_n}{dt} = -\nu P_n(t) + \frac{\nu}{2}(1 + aq_{n-1})P_{n-1}(t) + \frac{\nu}{2}(1 - aq_{n+1})P_{n+1}(t). \tag{2.209}$$

We now take the continuous limit of this master equation. Writing, as in Sect. 2.1.4, $P_n(t) = ap(na, t)$, where $p(x, t)$ is a distribution of the continuous variable x satisfying $\int_{-\infty}^{\infty} p(x, t)dx = 1$, we have

$$\frac{\partial p}{\partial t}(x, t) = -\nu p(x, t) + \frac{\nu}{2}[1 + aq(x - a)]p(x - a, t) + \frac{\nu}{2}[1 - aq(x + a)]p(x + a, t). \tag{2.210}$$

Expanding $p(x \pm a, t)$ and $q(x \pm a)$ to second order in a, we get

$$p(x \pm a, t) = p(x,t) \pm a \frac{\partial p}{\partial x}(x,t) + \frac{a^2}{2} \frac{\partial^2 p}{\partial x^2}(x,t) + \mathcal{O}(a^2), \tag{2.211}$$

$$q(x \pm a) = q(x) \pm a q'(x) + \frac{a^2}{2} q''(x) + \mathcal{O}(a^2). \tag{2.212}$$

One then finds, keeping only terms up to order a^2 in Eq. 2.210:

$$\frac{\partial p}{\partial t}(x,t) = -a^2 \nu q(x) \frac{\partial p}{\partial x} - a^2 \nu q'(x) p(x,t) + \frac{a^2 \nu}{2} \frac{\partial^2 p}{\partial x^2}. \tag{2.213}$$

We note that $a^2\nu$ is related both to the diffusion coefficient D introduced in Sect. 2.1.4, and to the coefficient Γ characterizing the correlation of the noise in Sect. 2.2.1:

$$a^2 \nu = 2D = \Gamma. \tag{2.214}$$

In order to have a well-defined continuous limit, one must here again take the limits $a \to 0$ and $\nu \to \infty$ in such a way that $a^2\nu$ converges to a finite value. Defining $Q(x) = \Gamma q(x)$, Eq. 2.213 can be rewritten as

$$\frac{\partial p}{\partial t}(x,t) = -\frac{\partial}{\partial x}\Big(Q(x) p(x,t) \Big) + \frac{\Gamma}{2} \frac{\partial^2 p}{\partial x^2}. \tag{2.215}$$

This equation is called a Fokker–Planck equation. It describes, from another perspective, the same random process as the Langevin equation (2.205).

As an example of application of the Fokker–Planck equation, we come back to the probe particle studied in Sect. 2.2.1. In this case, the variable x is replaced by the velocity v, and the bias function is given by $Q(v) = -\gamma v$. The Fokker–Planck equation reads

$$\frac{\partial p}{\partial t}(v,t) = \gamma \frac{\partial}{\partial v}\Big(v p(v,t) \Big) + \frac{\Gamma}{2} \frac{\partial^2 p}{\partial v^2}, \tag{2.216}$$

where the coefficients Γ and γ are related through $\Gamma = 2\gamma k_B T$. It can be checked that the solution of this equation, with initial condition $p(v, t = 0) = \delta(v - v_0)$—that is, the initial velocity is non-random and equal to v_0—is given by

$$p(v,t) = \left[2\pi k_B T \left(1 - e^{-2\gamma t} \right) \right]^{-1/2} \exp\left[-\frac{(v - v_0 e^{-\gamma t})^2}{2 k_B T (1 - e^{-2\gamma t})} \right]. \tag{2.217}$$

This process, namely a random walk confined by a quadratic potential, is also called Ornstein–Uhlenbeck process.

2.3 Anomalous Diffusion and Physical Aging

2.3.1 Generalized Central Limit Theorem

We have already briefly discussed in Sect. 2.1.4 the central limit theorem which, loosely speaking, states that the distribution of a sum of independent and identically distributed random variables converges to a Gaussian distribution in the limit of a large number of terms. This theorem actually relies on the important assumption that the summed random variables have a finite variance. When this condition is not satisfied, the central limit theorem breaks down.

Let us first examine which type of law can have an infinite variance. We start by the following simple example of power-law distribution (sometimes called Pareto distribution in the literature):

$$p(x) = \frac{\alpha x_0^{\alpha}}{x^{1+\alpha}}, \qquad x \geq x_0 \tag{2.218}$$

and $p(x) = 0$ for $x < x_0$. This distribution is well-defined for all $\alpha > 0$, in the sense that $\int_{x_0}^{\infty} p(x)dx = 1$. The second moment $\langle x^2 \rangle$ reads

$$\langle x^2 \rangle = \int_{x_0}^{\infty} x^2 p(x) dx = \int_{x_0}^{\infty} \frac{\alpha x_0^{\alpha}}{x^{\alpha-1}} dx. \tag{2.219}$$

The integral in Eq. 2.219 converges only for $\alpha > 2$, and diverges for $\alpha \leq 2$. Hence, considering sums of the type $S_N = \sum_{i=1}^{N} x_i$ where the variables x_i are distributed according to the distribution $p(x)$ given in Eq. 2.218, the central limit theorem does not apply if $\alpha \leq 2$, and such sums do not have a Gaussian distribution for large N. The aim of the Generalized Central Limit Theorem, that we shall discuss below, is precisely to deal with this type of situations.

However, before presenting the theorem, we would like first to give a flavor, in intuitive terms, of why the central limit theorem breaks down. Basically, when α is lowered, the distribution (2.218) becomes broader and broader, with a 'heavy tail' that contains a significant part of the probability weight. In other words, very large values of x have a significant probability to be drawn from the distribution, and such large values play an essential role in the sum.

We focus on the regime where this effect is the strongest, which corresponds to $\alpha < 1$. Indeed, in this range of α, the average value $\langle x \rangle$ itself becomes infinite. We slightly generalize the distribution given in Eq. 2.218 by assuming that only the tail of the distribution has a power-law behavior, namely

$$p(x) \sim \frac{c}{x^{1+\alpha}}, \qquad x \to \infty. \tag{2.220}$$

The variable x is defined in the range (x_0, ∞). Considering N random values x_i, $i = 1, \ldots, N$ drawn from the distribution $p(x)$, we wish to compare the largest value in the set $\{x_i\}$ to the sum $\sum_{i=1}^{N} x_i$. The typical value of the maximum $\max(x_i)$ can be evaluated as follows. Let us define

$$F_N^{\max}(z) \equiv \mathrm{Prob}\Big(\max(x_1, \ldots, x_N) < z\Big). \tag{2.221}$$

From the independence property of the x_i's, one has

$$F_N^{\max}(z) = \left(\int_{x_0}^{z} p(x)dx\right)^N = \Big(1 - \tilde{F}(z)\Big)^N, \tag{2.222}$$

where we have defined the complementary cumulative distribution $\tilde{F}(z) \equiv \int_z^\infty p(x)\,dx$. As the typical value of $\max(x_1, \ldots, x_N)$ is large for large N, we can approximate $\tilde{F}(z)$ by its asymptotic behavior at large z:

$$\tilde{F}(z) \sim \frac{\tilde{c}}{z^\alpha}, \qquad z \to \infty \tag{2.223}$$

with $\tilde{c} = c/\alpha$. It follows that

$$\ln\Big(1 - \tilde{F}(z)\Big)^N \sim -\frac{\tilde{c}N}{z^\alpha} \tag{2.224}$$

so that

$$F_N^{\max}(z) \sim e^{-\tilde{c}N/z^\alpha}. \tag{2.225}$$

In other words, $F_N^{\max}(z)$ can be rewritten in the scaling form

$$F_N^{\max}(z) \sim \Phi\left(\frac{z}{N^{1/\alpha}}\right), \tag{2.226}$$

with $\Phi(u) = e^{-\tilde{c}u^{-\alpha}}$, which indicates that the typical value of $\max(x_i)$ is of the order of $N^{1/\alpha}$, as $F_N^{\max}(z)$ increases from 0 to 1 around $z \approx N^{1/\alpha}$.

This observation has important consequences on the sum $\sum_{i=1}^{N} x_i$. Intuitively, one expects the typical value of the sum to be proportional to the number N of terms. If $\alpha > 1,\ N^{1/\alpha} \ll N$ for large N, so that the largest term remains much smaller than the sum. In contrast, if $\alpha < 1,\ N^{1/\alpha} \gg N$, and the assumption that $\sum_{i=1}^{N} x_i$ is of the order of N breaks down, as the sum is necessarily greater than its largest term. A more involved study shows in this case that the sum is of the order of the largest term itself, namely

$$\sum_{i=1}^{N} x_i = \mathcal{O}(N^{1/\alpha}). \tag{2.227}$$

It is customary to say that the largest term 'dominates' the sum.

For $1 < \alpha < 2$, the situation is slightly more subtle: the largest term remains much smaller than the sum, consistently with the finiteness of $\langle x \rangle$ which implies $\sum_{i=1}^{N} x_i \sim N\langle x \rangle$. However, the fluctuations of x remain large, as witnessed by the divergence of the variance of x, which prevents the central limit theorem from being applicable.

Let us now formulate the Generalized Central Limit Theorem [2, 4]. One considers a set of N independent and identically distributed random variables, drawn from a distribution $p(x)$ such that $\int_{-\infty}^{\infty} x^2 p(x)dx$ is infinite. The cumulative distribution function is denoted as $F(x) \equiv \int_{-\infty}^{x} p(x')dx'$. The Generalized Central Limit Theorem states that the distribution of the rescaled variable

$$z_N = \frac{1}{b_N}\left(\sum_{i=1}^{N} x_i - a_N\right) \tag{2.228}$$

converges, for a suitable choice of the rescaling parameters a_N and b_N, to the Lévy distribution $L(z; \alpha, \beta)$ with $0 < \alpha \leq 2$ and $-1 \leq \beta \leq 1$, if the following conditions are satisfied

$$\lim_{x\to\infty} \frac{F(-x)}{1 - F(x)} = \frac{1-\beta}{1+\beta} \tag{2.229}$$

$$\forall r > 0, \quad \lim_{x\to\infty} \frac{1 - F(x) + F(-x)}{1 - F(rx) + F(-rx)} = r^{\alpha}. \tag{2.230}$$

The Lévy distribution $L(z; \alpha, \beta)$ is defined through its characteristic function (Fourier transform of the probability density)

$$\begin{aligned} \hat{L}(k; \alpha, \beta) &\equiv \int_{-\infty}^{\infty} L(z; \alpha, \beta) e^{-ikz} dz \\ &= \exp\Big[-|k|^{\alpha} \Big(1 + i\beta \ \mathrm{sgn}(k)\, \varphi(k, \alpha)\Big)\Big] \end{aligned} \tag{2.231}$$

with

$$\varphi(k, \alpha) = \begin{cases} \tan \frac{\pi\alpha}{2} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \ln |k| & \text{if } \alpha = 1 \end{cases} \tag{2.232}$$

A few remarks are in order here:

- The parameter α has the same interpretation here as in the example above, namely the tail of the distribution $p(x)$ typically decays with an exponent $1 + \alpha$. More precisely, α is related to the 'heaviest' tail: if $p(x)$ behaves as

$$p(x) \sim \frac{c_1}{x^{1+\alpha_1}}, \qquad x \to +\infty \tag{2.233}$$

$$p(x) \sim \frac{c_2}{|x|^{1+\alpha_2}}, \qquad x \to -\infty \tag{2.234}$$

with $0 < \alpha_1, \alpha_2 \leq 2$, then $\alpha = \min(\alpha_1, \alpha_2)$.

- The parameter β characterizes the asymmetry of the Lévy distribution: $\beta = 0$ corresponds to a symmetric distribution $L(-z;\alpha,\beta) = L(z;\alpha,\beta)$, while for $\beta > 0$ (resp. $\beta < 0$) positive (resp. negative) values of z have a higher probability. In the above example, one has $\beta = 1$ if $\alpha_1 < \alpha_2$ and $\beta = -1$ if $\alpha_1 > \alpha_2$, while if $\alpha_1 = \alpha_2$, β is given by

$$\beta = \frac{c_1 - c_2}{c_1 + c_2}\,. \tag{2.235}$$

- A “suitable choice of a_N and b_N” means that there exists a family of parameters $\{a_N\}$ and $\{b_N\}$ such that the theorem applies. Though the theorem does not specify the values of a_N and b_N, typical choices are $a_N = 0$ $(0 < \alpha < 1)$ or $a_N = N\langle x\rangle + a_0$ $(1 < \alpha < 2)$, and $b_N = b_1 N^{1/\alpha}$. Logarithmic corrections may be needed for $\alpha = 1$.
- For general values of α and β, the Lévy distribution is known only through its Fourier transform. However, for a few specific values of the parameters, explicit expressions of $L(z;\alpha,\beta)$ are available. For instance, the Gaussian distribution is recovered for $\alpha = 2$, whatever the value of β:

$$L(z;2,\beta) = \frac{1}{\sqrt{4\pi}} e^{-z^2/4}. \tag{2.236}$$

Note that due to the parameterization used in Eq. 2.231, the variance of the Gaussian distribution (2.236) is not equal to one. For $\alpha = 1$ and $\beta = 0$, one gets the Cauchy distribution:

$$L(z;1,0) = \frac{1}{\pi(1+z^2)}. \tag{2.237}$$

2.3.2 *Anomalous Diffusion*

The above results on the statistics of sums of broadly distributed random variables have important consequences for anomalous diffusion processes. Such processes are characterized by the fact that the typical displacement of the random walk grows with time t either faster than $t^{1/2}$ (superdiffusion) or slower than $t^{1/2}$ (subdiffusion).

Let us start with the superdiffusive case, which corresponds to a broad distribution of jump sizes. We consider a discrete time random walk evolving according to $x_{t+1} = x_t + u_t$, where u_t is drawn from a symmetric distribution $p(u)$. We assume that space is continuous, and that the variables $\{u_t\}$ are independent and identically distributed random variables. Accordingly, the position x_t is given by

$$x_t = \sum_{t'=0}^{t-1} u_{t'} \tag{2.238}$$

where we have assumed that $x_0 = 0$. The present problem is thus directly related to problems of random sums. The symmetry of the distribution $p(u)$ implies that $\langle u_t \rangle = 0$, from which $\langle x_t \rangle = 0$ follows. If $\langle u^2 \rangle$ is finite, one has

$$\langle x_t^2 \rangle = \sum_{t',\,t''} \langle u_{t'} u_{t''} \rangle = t \langle u^2 \rangle \tag{2.239}$$

where we have used the fact that the variables $u_{t'}$ and $u_{t''}$ are statistically independent. Hence the mean-square displacement $\langle x_t^2 \rangle$ is linear in t, which corresponds to a normal diffusive process. In contrast, if the distribution $p(u)$ is broad, with an infinite variance, the above reasoning fails, and one needs to use the Generalized Central Limit Theorem. Considering a distribution $p(u)$ such that

$$p(u) \sim \frac{c}{|u|^{1+\alpha}}, \qquad u \to \pm\infty \tag{2.240}$$

with $\alpha < 2$, the distribution of the rescaled variable

$$z_t = \frac{x_t}{b_1 t^{1/\alpha}}, \tag{2.241}$$

where b_1 is a suitable scale factor, converges to the Lévy distribution $L(z; \alpha, 0)$. This means that the typical displacement x_{typ} after a time t is of the order of $t^{1/\alpha}$. As $\alpha < 2$, one has $\frac{1}{\alpha} > \frac{1}{2}$, which indeed corresponds to a superdiffusive process.

On the contrary, subdiffusive walks correspond to a strong local trapping effect, so that the sojourn times on a given site become broadly distributed, instead of being fixed to a value Δt as the above superdiffusive example. We thus consider a random walk process in which the time lag τ between two jumps is itself a random variable τ following a distribution $p(\tau)$, with a tail $p(\tau) \sim c/\tau^{1+\alpha}$ when $\tau \to \infty$ $(0 < \alpha < 1)$. The walker jumps to one of the two neighboring sites, namely $x_{t+\tau} = x_t + \epsilon_t$, where $\epsilon_t = \pm 1$ with equal probabilities.

This case is not a direct application of the Generalized Central Limit Theorem. However, the behavior of the random walk can be understood through a simple scaling argument. After N steps, the typical displacement x_N^{typ} of the walker is of the order of $\sqrt{N}$. To relate N to the actual time t, one can observe that time t is the sum of the N sojourn times τ_i at the ith position. Hence,

$$t = \sum_{i=1}^{N} \tau_i = \mathcal{O}(N^{1/\alpha}) \tag{2.242}$$

where we have used the estimation given in Eq. 2.227. It follows that $N = \mathcal{O}(t^{\alpha})$, so that $x_N^{\mathrm{typ}} = \mathcal{O}(t^{\alpha/2})$, which is slower than normal diffusion (we recall that $\alpha < 1$). Note that there also exists rigorous methods to prove the above scaling arguments.

2.3.3 Aging Dynamics and Trap Models

Although many systems converge to a stationary state on times shorter than or comparable to the observation time, it turns out that some systems do not reach a steady state and keep evolving on time scales that can be very large compared to standard observation times. This is the case for instance of glasses, which keep aging for years or more [5]. The same type of mechanism is at play in laser cooling experiments [5]. It is also likely that aging mechanisms, or slow relaxation effects, play a significant role in many different types of complex systems. Even though the aging mechanisms may differ from one situation to the other, it is certainly of interest to investigate one of the simplest known aging phenomena, illustrated by the trap model, which we describe here within a generic formalism that does not rely on a specific physical realization.

Let us consider a model system in which to each configuration C is associated a given lifetime τ. This lifetime τ is the mean time spent in configuration C before moving to another configuration. As we consider only temporal aspects of the dynamics, and not other types of observables (energy, magnetization,...), we simply label the configurations by their lifetime τ. We then choose a simple form for the transition rate $W(\tau'|\tau)$, namely:

$$W(\tau'|\tau) = \frac{1}{\tau}\psi(\tau'). \tag{2.243}$$

The function $\psi(\tau')$ is the a priori probability distribution of the configurations τ', meaning that the selected new configuration is chosen completely at random. From the normalization condition $\int_0^\infty d\tau'\psi(\tau') = 1$, we have

$$\int_0^\infty d\tau' W(\tau'|\tau) = \frac{1}{\tau}, \tag{2.244}$$

so that the characteristic escape time from a configuration with lifetime τ is precisely τ. For simplicity, we also assume that all lifetimes τ are greater than a value τ_0, that we set to $\tau_0 = 1$ in the following. The master equation then reads:

$$\begin{aligned}\frac{\partial P}{\partial t}(\tau, t) &= -\, P(\tau, t)\int_1^\infty d\tau' W(\tau'|\tau) + \int_1^\infty d\tau' W(\tau|\tau')P(\tau', t)\\ &= -\,\frac{1}{\tau}P(\tau, t) + \psi(\tau)\int_1^\infty \frac{d\tau'}{\tau'}P(\tau', t).\end{aligned} \tag{2.245}$$

At equilibrium, the probability to be in a configuration with lifetime τ is proportional to τ and to the a priori distribution $\psi(\tau)$ of configurations:

$$P_{\text{eq}}(\tau) = \frac{1}{\langle\tau\rangle}\tau\psi(\tau), \tag{2.246}$$

where $\langle\tau\rangle$ is defined as

$$\langle\tau\rangle = \int_1^\infty d\tau\,\tau\psi(\tau). \tag{2.247}$$

An interesting situation appears when the distribution $\psi(\tau)$ takes a power-law form, namely

$$\psi(\tau) = \frac{\alpha}{\tau^{1+\alpha}}, \qquad \tau > 1, \quad \alpha > 0. \tag{2.248}$$

This is realized for instance in the case of a particle trapped into potential wells of random depth E, with an exponential distribution

$$\rho(E) = \frac{1}{E_0}e^{-E/E_0}. \tag{2.249}$$

The lifetime τ is given by the standard Arrhenius law

$$\tau = \tau_0 e^{E/T}, \tag{2.250}$$

where $\tau_0 = 1$ is a microscopic time scale. Using the relation $\psi(\tau)|d\tau| = \rho(E)|dE|$, one precisely finds the form (2.248) for $\psi(\tau)$, with $\alpha = T/E_0$.

In the case $\alpha > 1$, $\langle\tau\rangle$ is finite, but if $\alpha \leq 1$ then $\langle\tau\rangle$ is infinite, so that the equilibrium distribution (2.246) does not exist, as it is not normalizable. As a result, no stationary state can be reached, and the system keeps drifting toward configurations with larger and larger lifetimes τ.

It is then of interest to determine the time-dependent probability distribution $P(\tau, t)$ in the long-time regime. We postulate the following scaling form

$$P(\tau, t) = \frac{1}{t}\phi\left(\frac{\tau}{t}\right). \tag{2.251}$$

From the normalization condition of $P(\tau, t)$, one has

$$\int_1^\infty d\tau P(\tau, t) = \frac{1}{t}\int_1^\infty d\tau\phi\left(\frac{\tau}{t}\right) = 1, \tag{2.252}$$

from which one gets, with the change of variable $u = \tau/t$,

$$\int_{1/t}^\infty du\phi(u) = 1. \tag{2.253}$$

As $\phi(u)$ does not depend explicitly on time t, the above condition cannot be satisfied for all t. But we are looking for an asymptotic large-t solution, so that we impose that Eq. 2.253 is satisfied in the infinite t limit, namely

$$\int_0^\infty du\phi(u) = 1. \tag{2.254}$$

As a result, the scaling form (2.251) is an approximate solution that becomes exact when $t \to \infty$. From Eq. 2.251, one obtains for the time derivative of $P(\tau, t)$:

$$\frac{\partial P}{\partial t} = -\frac{1}{t^2}\phi\left(\frac{\tau}{t}\right) - \frac{\tau}{t^3}\phi'\left(\frac{\tau}{t}\right), \tag{2.255}$$

where ϕ' is the derivative of ϕ. Multiplying Eq. 2.255 by t^2, one obtains, with the notations $u = \tau/t$ and $v = \tau'/t$,

$$-\phi(u) - u\phi'(u) = -\frac{1}{u}\phi(u) + \psi(ut)t\int_{1/t}^\infty \frac{dv}{v}\phi(v). \tag{2.256}$$

Using the specific form (2.248) of $\psi(\tau)$, we find

$$\left(1 - \frac{1}{u}\right)\phi(u) + u\phi'(u) + \frac{\alpha}{u^{1+\alpha}}t^{-\alpha}\int_{1/t}^\infty \frac{dv}{v}\phi(v) = 0. \tag{2.257}$$

For the above equation to be well-defined in the infinite t limit in which it is supposed to be valid, the explicit t-dependence has to cancel out. One thus needs to have

$$\int_{1/t}^\infty \frac{dv}{v}\phi(v) \sim t^\alpha, \qquad t \to \infty, \tag{2.258}$$

which requires that $\phi(v)$ has the following asymptotic form at small v:

$$\phi(v) \approx \frac{\phi_0}{v^\alpha}, \qquad v \to 0. \tag{2.259}$$

Here, ϕ_0 is an arbitrary constant, to be determined later on from the normalization condition of $\phi(u)$. The master equation is then finally written as the following differential equation:

$$\left(1 - \frac{1}{u}\right)\phi(u) + u\phi'(u) + \frac{\phi_0}{u^{1+\alpha}} = 0. \tag{2.260}$$

This equation is a linear differential equation, and its solution can be found using standard techniques. The solution of Eq. 2.260 satisfying the normalization condition (2.254) reads [6]

$$\phi(u) = \frac{\sin(\pi\alpha)}{\Gamma(\alpha)} \frac{1}{u} e^{-1/u} \int_0^{1/u} dv v^{\alpha-1} e^v, \tag{2.261}$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ is the Euler Gamma function. It is rather easy to show that $\phi(u) \sim u^{-\alpha}$ for $u \to 0$ as expected, and that $\phi(u) \sim u^{-1-\alpha}$ for $u \to \infty$, leading for $P(\tau, t)$ to

$$P(\tau, t) \propto \tau\psi(\tau), \qquad \tau \ll t, \tag{2.262}$$

$$P(\tau, t) \propto \psi(\tau), \qquad \tau \gg t. \tag{2.263}$$

These asymptotic behaviors can be interpreted rather easily: configurations with lifetimes $\tau \ll t$ have been visited a large number of times, so that they are quasi-equilibrated; in contrast, configurations with lifetimes $\tau \gg t$ have been visited at most once, and the precise value of τ is not yet felt by the dynamics (τ appears as essentially infinite).

References

1. N.G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1992)
2. B.V. Gnedenko, A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Reading, 1954)
3. V.V. Petrov, *Sums of Independent Random Variables* (Springer, Berlin, 1975)
4. W. Feller, *An Introduction to Probability Theory and its Applications*. vol. I, (Wiley, New York, 1966)
5. E. Bertin, F. Bardou, Am. J. Phys. **76**, 630 (2008)
6. C. Monthus, J.-P. Bouchaud, J. Phys. A **29**, 3847 (1996)

Chapter 3
Statistical Physics of Interacting Macroscopic "Entities"

Until now, we have mainly considered physical systems, in which elementary entities are implicitly atoms or molecules. In this case, the laws of motion of the individual particles are known, and the main difficulty consists in being able to change the scale of description, going from the scale of particles to the system size.

However, our everydaylife experience tells us that there exist many familiar systems that are composed of interacting macroscopic "entities", that thus behave very differently from atoms or molecules: examples range from sand piles, foams, bacteria colonies, animal flocks, or road traffic, to quote only a few examples. In such cases, it is clear that the interacting objects, or individuals, cannot be described in the same way as molecules, and precise dynamical laws at the individual scale are most often not known.

The difficulties encountered when trying to apply a statistical physics approach to such assemblies of macroscopic "entities" are then two-fold. On the one hand, a model should be given for the dynamics of individual, and it is often not clear how relevant or reliable such modeling is to describe realistic systems. On the other hand, reasonable models of individual dynamics usually do not have similar conservation laws and time-reversal symmetry as the Hamiltonian dynamics of molecular systems. Hence it is hard, even in specific cases, to build a statistical physics approach from a postulate similar to the hypothesis of equiprobability of configurations having the same energy. Interesting attempts in this direction, notably in the context of granular matter, have however been proposed [1].

In this last part, we illustrate on several examples how different statistical physics techniques can be devised, in specific cases, to describe assemblies of interacting agents. In the first example (the dynamics of residential moves in a city, Sect. 3.1), a mapping can be performed to an effective equilibrium system, yielding interesting insights. In the second example (condensation transition, Sect. 3.2), an explicit stationary solution of the master equation can be found. In the third example (synchronization transition, Sect. 3.3), a mean-field approach is used to predict the collective behavior of coupled oscillators. Finally, the last example (collective motion, Sect. 3.4) is studied through the so-called Boltzmann equation, a generic approach that can be

E. Bertin, *A Concise Introduction to the Statistical Physics of Complex Systems*,
SpringerBriefs in Complexity, DOI: 10.1007/978-3-642-23923-6_3,

used when interactions are limited to binary "collisions", that are extremely localized in space and time.

3.1 Dynamics of Residential Moves

A standard example of complex system dynamics is the Schelling model which represents in a schematic way the dynamics of residential moves in a city [2, 3]. The city is modeled as a checkerboard, divided into cells. Two types of agents (say red and green) live in the city. They reside in the cells of the checkerboard, with at most one agent in each cell. Agents characterize their degree of satisfaction regarding their environment by a utility, which is a given function (the same for all agents) of the number of agents of the same type in their neighborhood. The neighborhood can be defined in different ways. One possibility would be to consider the set of nearest neighbors. However, most studies rather use the Moore neighborhood, that is the 3×3 (or sometimes 5×5) square surrounding the current cell.

Before moving, an agent chooses at random an empty cell, and evaluates the utility u_{new} associated to this new location. The agent compares this quantity to the utility u_{old} of his present location, by computing the utility difference $\Delta u = u_{\text{new}} - u_{\text{old}}$. The move is then accepted with probability $1/(1 + e^{-\Delta u/T})$. Here, T is a parameter analogous to the temperature in physical systems, that characterizes the influence of other factors, like the presence of facilities, shops, or friends, that are not explicitly taken into account in the model, but could bias the decision of moving or not. At low T, and for a large class of utility functions such that agents have a (possibly slight) preference for being with agents of the same type, a segregation phenomenon is observed when simulating the model numerically: two types of domains form, namely domains with a majority of red agents and domains with a majority of green agents. Quite surprisingly, this segregation phenomenon seems quite robust, and is also observed in the case where agents have a marked preference for mixed neighborhood.

The Schelling model in its standard form is very hard to solve analytically, and solutions are not presently known. The reason for these difficulties is mainly that the neighborhoods of two neighboring cells overlap, generating complicated correlations in the system. In order to find an analytical solution, a standard strategy is to define a variant of the model on a specific geometry that avoids these correlations. This strategy was for instance successful in the Ising model, by introducing a fully connected version of the model (see Sect. 1.4.2): assuming that all spins interact together, the phase transition could be obtained analytically in a simple way.

A straightforward application of this idea to the Schelling model a priori seems to lead to a deadlock. If an agent evaluates his utility by considering the whole city as his neighborhood, this utility will not change when moving within the city. A more interesting strategy is then to divide the city into a large number of blocks, so that agents evaluate their utility within blocks, and move from blocks to blocks. In this way, correlations between blocks may be suppressed.

3.1.1 A Simplified Version of the Schelling Model

In order to implement this strategy, we consider the following model, with a single type of agent to further simplify the derivation (the case of two different types of agents can be dealt with in the same way). The segregation phenomenon then corresponds to the formation of domains of different densities. The city is divided into a large number Q of blocks, each block containing H cells (a cell may be thought of as representing a flat). We assume that each cell can contain at most one agent, so that the number n_q of agents in a given block q ($q = 1, \dots, Q$) satisfies $n_q \leq H$. A microscopic configuration C of the city corresponds to the knowledge of the state (empty or occupied) of each cell. For each block q, we also introduce the density of agents $\rho_q = n_q/H$. Each agent has the same utility function $u(\rho_q)$, which describes the degree of satisfaction concerning the density of the block it is living in. The collective utility is defined as the total utility of all the agents in the city: $U(C) = \sum_q n_q u(\rho_q)$.

A dynamical rule allows the agents to move from one block to another. At each time step, one picks up at random an agent and a vacant cell, within two different blocks. The agent moves in that empty cell with probability:

$$W(C'|C) = \frac{1}{1 + e^{-\Delta u/T}}, \tag{3.264}$$

where C and C' are the configurations before and after the move respectively, and Δu is the variation of utility associated to the proposed move. The parameter T has the same interpretation as in the standard Schelling model.

It is interesting at this stage to emphasize the difference between the present model and standard physical approaches. It could seem at first sight that the utility is simply the equivalent, up to a sign reversal, of the energy in physics. In the present model however, an economics perspective is adopted, so that the agents are considered as purely selfish. They take decision only according to their own utility change Δu, and do not consider the potential impact of their decision on the other agents. In contrast, in physical models, the probability for a particle to move depends on the energy variation of the whole system, and the effect on the other particles is thus taken into account from the outset. This has important consequences, as we shall see below.

3.1.2 Equilibrium Configurations of the Model

We wish to find the stationary probability distribution *P(C)* of the microscopic configurations *C*. This is not an easy task in general. Yet, if we were able to show that a detailed balance relation holds in this model, we would straightforwardly get the solution. Let us assume that the individual cost Δu can be written as $\Delta u = F(C') - F(C)$, where F is a function on configuration space. From Eq. 3.264, we find that the dynamics satisfies a detailed balance relation:

$$W(C'|C)P(C) = W(C|C')P(C'), \tag{3.265}$$

with a distribution *P(C)* given by

$$P(C) = \frac{1}{Z}\, e^{-F(C)/T}. \tag{3.266}$$

It can be shown that a function F satisfying this condition is given by

$$F(C) = -\sum_{q}\sum_{m=0}^{n_q} u(m/H). \tag{3.267}$$

To characterize the "segregation" phenomenon, the full statistical information on the occupation number of each cell is not necessary. Instead, an aggregated description in terms of densities of the blocks turns out to be more useful. Such a coarse-grained description is obtained by aggregating all configurations with the same number of agents in each block. As there are $H!/(n!(H-n)!)$ ways of ordering n agents in H cells, we obtain the following coarse-grained probability distribution:

$$\tilde{P}(\{n_q\}) = \frac{1}{Z}\, \exp\left(-\frac{H}{T}\sum_{q}\tilde{f}(n_q)\right), \tag{3.268}$$

where we have introduced the function $\tilde{f}$:

$$\tilde{f}(n) = \frac{T}{H}\ln\left(\frac{n!(H-n)!}{H!}\right) - \frac{1}{H}\sum_{m=0}^{n_q} u\left(\frac{m}{H}\right). \tag{3.269}$$

The above expression suggests to consider the limit of large H in order to get a continuous formulation for $\tilde{f}$. Keeping constant the density of each block $\rho_q = n_q/H$ (ρ_q hence becoming a continuous variable) and expanding the factorials using Stirling's formula $\ln n! \approx n\ln n - n$, valid for large n, one obtains for $H \to \infty$

$$\frac{1}{H}\ln\left(\frac{n_q!(H-n_q)!}{H!}\right) \to \rho_q\ln\rho_q + (1-\rho_q)\ln(1-\rho_q). \tag{3.270}$$

Similarly, the last term in the expression of $\tilde{f}$ converges to an integral:

$$\frac{1}{H}\sum_{m=0}^{n_q} u\left(\frac{m}{H}\right) \to \int_{0}^{\rho_q} u(\rho')d\rho'. \tag{3.271}$$

In terms of density ρ_q, the stationary distributions $\tilde{P}(\{n_q\})$ turns into a probability density $\Psi(\{\rho_q\})$ given by (with $\sum_{q=1}^{Q} \rho_q = Q\rho_0$ held fixed):

$$\Psi(\{\rho_q\}) = K \exp\left(-\frac{H}{T}\sum_{q=1}^{Q} f(\rho_q)\right) \tag{3.272}$$

where K is a normalization constant, and where the function $f(\rho)$ is defined as

$$f(\rho) = T\rho \ln \rho + T(1-\rho)\ln(1-\rho) - \int_0^{\rho} u(\rho')d\rho'. \tag{3.273}$$

The function $\Phi(\{\rho_q\}) = \sum_q f(\rho_q)$ may be called a potential, or a large deviation function. It is also the analogue of the free energy functions used in physics. The configurations $\{\rho_q\}$ that minimize the potential $\Phi(\{\rho_q\})$ under the constraint of fixed $\sum_{q=1}^{Q} \rho_q$ are the most probable to come up. In the limit $H \to \infty$, these configurations are the only ones that appear in the stationary state, as the probability of other configurations vanishes exponentially with H.

3.1.3 Condition for Phase Separation

Focusing on the large H case, the problem gets back to finding the set $\{\rho_q\}$ which minimizes the potential $\Phi(\{\rho_q\})$ under the constraint that $\sum_q \rho_q$ is fixed. We are interested in knowing whether the stationary state is statistically homogeneous or inhomogeneous. Following standard physics textbooks methods [4], the homogeneous state at density ρ_0 is unstable against a phase separation if there exists two densities ρ_1 and ρ_2 such that

$$\gamma f(\rho_1) + (1-\gamma) f(\rho_2) < f(\rho_0). \tag{3.274}$$

The parameter γ $(0 < \gamma < 1)$ corresponds to the fraction of blocks that would have a density ρ_1 in the diphasic state. This condition simply means that the value of the potential Φ is lower for the diphasic state than for the homogeneous state, so that the diphasic state has a much larger probability to occur. Geometrically, the inequality (3.274) corresponds to requiring that $f(\rho)$ is a non-convex function of ρ. The values of ρ_1 and ρ_2 are obtained by minimizing $\gamma f(\rho_1') + (1-\gamma) f(\rho_2')$ over all possible values of ρ_1' and ρ_2', with γ determined by the mass conservation $\gamma\rho_1' + (1-\gamma)\rho_2' = \rho_0$ (see Fig. 3.1).

We now try to translate the convexity condition (3.274) into a condition on the utility function $u(\rho)$. Phase separation occurs if there is a range of density for which $f(\rho)$ is concave, namely $f''(\rho) < 0$. We thus compute the second derivative of f, yielding

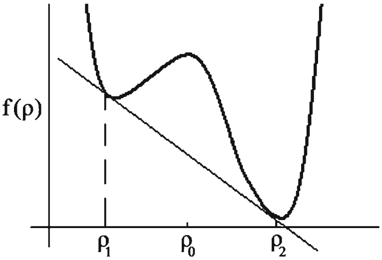

Fig. 3.1 Phase separation: the system of density ρ_0 splits into two phases of densities ρ_1 and ρ_2 to lower its potential

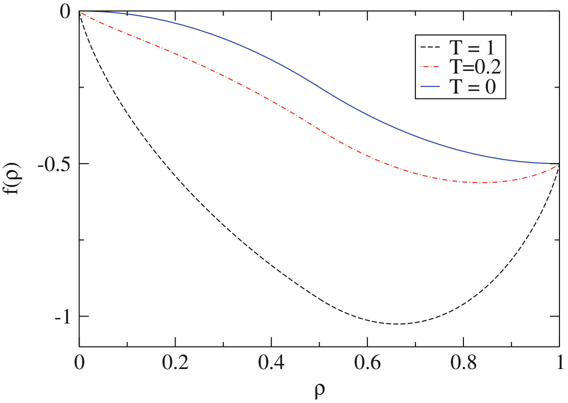

Fig. 3.2 Plot of the function $f(\rho)$ for different values of temperature, T = 1, T = 0.2 and T = 0, illustrating that $f(\rho)$ becomes non-convex for T < 0.5, leading to a phase separation

$$f''(\rho) = \frac{T}{\rho(1-\rho)} - u'(\rho). \tag{3.275}$$

For a given utility function, this condition can be checked explicitly. We note that in the limit $T \to 0$, $f''(\rho) = -u'(\rho)$, so that the homogeneous state is stable if $u(\rho)$ is a monotonously decreasing function of ρ.

The specific form of the utility function is an input of the model, and it can be postulated on a phenomenological basis, or rely on a theory of the interactions among agents. In order to analyze an explicit example of a non-linear utility function, we consider the asymmetrically peaked utility function defined as:

$$u(\rho) = \begin{cases} 2\rho & \text{if } \rho \le \frac{1}{2} \\ 2(1-\rho) & \text{if } \rho > \frac{1}{2}. \end{cases} \tag{3.276}$$

The expression of $f(\rho)$ can be easily deduced from $u(\rho)$, and is illustrated on Fig. 3.2 for different values of T. To study the stability of the homogeneous phase, we look at the sign of $f''(\rho)$. One has for $\rho < 1/2$

$$f''(\rho) = \frac{T}{\rho(1-\rho)} - 2, \tag{3.277}$$

and for $\rho > 1/2$:

$$f''(\rho) = \frac{T}{\rho(1-\rho)} + 2. \tag{3.278}$$

It is easy to check that $f''(\rho)$ is minimum for $\rho \to \frac{1}{2}^-$, the corresponding value being

$$\lim_{\rho \to \frac{1}{2}^-} f''(\rho) = 4T - 2. \tag{3.279}$$

Thus, for $T > 1/2$, the function $f(\rho)$ is convex on the whole interval $0 < \rho < 1$ as $f''(\rho) > 0$ on this interval, and the homogeneous phase is stable. On the contrary, for $T < 1/2$, there exists an interval $\rho_1 < \rho < \rho_2$ where $f(\rho)$ is concave ($f''(\rho) < 0$), so that in the stationary state, the system is split into two phases of respective densities ρ_1 and ρ_2.

The surprising phenomenon here is that a phase separation occurs even in the case $\rho_0 = 1/2$, although all agents have a significant preference for a half-filled neighborhood. This can be understood intuitively as follows. At a small, but non-zero temperature *T*, small fluctuations of density in the blocks are possible. Let us assume that we start from the homogeneous state of density $\rho_0 = 1/2$, with some small fluctuations of this density around the mean value 1/2. If a block has a density smaller that 1/2, then this block becomes less attracting for the agents living in it. So some agents will start to move to the most attractive blocks which have exactly the density 1/2. In doing so, the initial block becomes less and less attractive, thus making more and more agents leave it. This avalanche process, which is related to the selfish behavior of the agents, qualitatively explains the instability of the homogeneous state with density 1/2.

3.2 Condensation Transition

Let us now turn to a different type of situation, involving alternative techniques. While the above variant of the Schelling model could be dealt with by using a mapping to an equilibrium system, in many cases such equilibrium methods are not sufficient to solve the model, due for instance to the presence of fluxes in the system. One must then resort to other kinds of approaches. Among possible approaches, one can consider simple enough stochastic models for which an exact solution of the master equation can be found in the steady state, although detailed balance is not satisfied. A prominent example of such type of models is the so-called zero-range process (ZRP) [5], that we describe below. Another well-known example of exactly solvable non-equilibrium model is the asymmetric simple exclusion process (ASEP), for which the derivation of the solution is however much more technical [6].

3.2.1 Zero Range Process: Definition and Exact Solution

In the ZRP, N particles are randomly placed on the L sites of a one-dimensional lattice with periodic boundary conditions,[1] and can jump from site i to the neighboring site $i+1$ (with the convention $L + 1 \equiv 1$). Motion is thus biased, which generates a current of particles along the ring. The interaction between particles is taken into account through the fact that the probability per unit time to jump from site i to site $i+1$ depends on the current number n_i of particles on site i; this probability is denoted as $u(n_i)$.

A microscopic configuration of the ZRP is given by the set $C = \{n_i\}$ of the occupation numbers of all sites. The transition rate $W(C'|C)$ can be written formally as

$$W(\{n_i'\}|\{n_i\}) = \sum_{i=1}^{L} u(n_i)\, \delta_{n_i',n_i-1}\, \delta_{n_{i+1}',n_{i+1}+1} \prod_{j\neq i,i+1} \delta_{n_j',n_j} \tag{3.280}$$

where $\delta_{n',n}$ is the Kronecker symbol, equal to 1 if $n' = n$, and to 0 otherwise. Using this form of the transition rate, one can write the corresponding master equation, which we do not display here to lighten the presentation. It can be shown [5] that the steady-state distribution can be looked for under a factorized form

$$P(\{n_i\}) = \frac{1}{Z}\left(\prod_{i=1}^{L} f(n_i)\right)\delta_{\sum_j n_j,N} \tag{3.281}$$

where the Kronecker delta symbol accounts for the conservation of the total number of particles. Inserting this form in the master equation, one obtains the expression of $f(n)$:

$$f(n) = \begin{cases} \prod_{k=1}^{n} \dfrac{1}{u(k)} & \text{if } n \geq 1 \\ 1 & \text{if } n = 0. \end{cases} \tag{3.282}$$

Note that the model can also be defined in such a way as to obtain any desired function $f(n)$ in the steady-state distribution: one simply needs to choose $u(n) = f(n-1)/f(n)$, for $n \geq 1$.

3.2.2 Maximal Density and Condensation Phenomenon

One of the interesting properties of the ZRP is the presence of a condensation transition, where a finite fraction of the total number of particles gather on a single site.

---

[1] We consider here for simplicity the ring geometry, but the ZRP can actually be defined on an arbitrary graph [7].

Such a phenomenon appears in the case of a function $f(n)$ decaying as a power-law, $f(n) \sim A/n^{\alpha}$, or equivalently $u(n) = 1+\alpha/n+o(1/n)$. The single-site distribution can be obtained by considering the rest of the system as a reservoir of particles, a situation similar to the canonical ensemble at equilibrium. Assuming the system to be homogeneous, the single-site distribution is then given by

$$p(n) = cf(n)e^{-\mu n} \tag{3.283}$$

where μ is the effective chemical potential of the reservoir. The normalization constant c is determined by

$$\frac{1}{c} = \sum_{n=0}^{\infty} f(n)e^{-\mu n}. \tag{3.284}$$

The convergence of this sum requires that $\mu > 0$ (or $\mu \geq 0$ if $\alpha > 1$). The average density

$$\rho = \langle n \rangle = c\sum_{n=1}^{\infty} nf(n)e^{-\mu n} \tag{3.285}$$

is a decreasing function of μ, which thus reaches its maximum value $\rho_{\rm c}$ for $\mu \to 0$:

$$\rho_{\rm c} = c\sum_{n=1}^{\infty} n\, f(n) \sim C\sum_{n=1}^{\infty} \frac{A}{n^{\alpha-1}}. \tag{3.286}$$

Hence $\rho_{\rm c}$ is infinite if $\alpha \leq 2$, and finite if $\alpha > 2$. As a result, if $\alpha > 2$, a homogeneous density of particles cannot exceed a finite value $\rho_{\rm c}$. If, on the contrary, one imposes a density $\rho_0 > \rho_{\rm c}$, by including in the system a number of particles $N > L\rho_{\rm c}$, the dynamics will necessarily evolve toward a non-homogeneous state. It can be shown [5] that the resulting state is composed of a "fluid phase", homogeneous at density $\rho_{\rm c}$, and a "condensate", that is a single site containing the remaining number of particles in the system, namely $L(\rho_0 - \rho_{\rm c})$.

Applications of this model range from vibrated granular matter (each site corresponding to a vibrated urn containing grains), to road traffic (sites being sections of roads), or network dynamics (that is, the dynamics of attachment and detachment of links on the network) [5]. Note that the ZRP is a very simplified model, so that mapping it to more realistic situations often implies approximations.

3.3 Synchronization Transition

In the two previous examples, the basic "entities" were effectively modeled as simple particles, with no internal dynamics. There are however cases where the internal dynamics plays a key role. This is the case for instance in several types of biological processes, from the cellular scale to that of living organisms, which have a periodic internal activity. In some cases, a population of such dynamical “entities” may enter a synchronized state, in which all oscillators have essentially the same phase, leading to global oscillations at the level of the population. A nice example is provided by a specific species of fireflies living in South Asian forests. During the night, swarms of such fireflies emit flashes of light in a synchronous way [8]. Many other examples can also be found in biological, chemical, physical or social systems (like the synchronization of applause in the latter case). One of the simplest models describing the synchronization phenomenon is the Kuramoto model, that we describe below.

3.3.1 The Kuramoto Model of Coupled Oscillators

The Kuramoto model [9] consists in a set of N oscillators of phase θ_j, evolving according to the coupled equations

$$\frac{d\theta_j}{dt} = \omega_j + \sum_{k=1}^{N} K_{jk} \sin(\theta_k - \theta_j), \qquad j = 1, \dots, N, \tag{3.287}$$

where ω_j is the natural frequency of oscillator j, and K_{jk} is the coupling constant between oscillators j and k. Applications of the Kuramoto model range from chemical oscillators to neural networks, laser arrays or Josephson junctions [8]. We shall here mostly follow the presentation of this model given in Ref. [8], and refer the reader to this specialized review for further details.

For reasons that will appear clear below, it is convenient to work in the rotating frame having the angular velocity $\Omega = \frac{1}{N}\sum_{j=1}^{N}\omega_j$. In this frame, the new phase is given by $\theta'_j(t) = \theta_j(t) - \Omega t$ and the frequencies are $\omega'_j = \omega_j - \Omega$. As the phase difference $\theta'_k - \theta'_j$ in the rotating frame is equal to the phase difference $\theta_k - \theta_j$ in the original frame, it turns out that the phases θ'_j also obey Eq. 3.287, simply replacing the phases and frequencies by their “primed” counterparts. A synchronized state is then a state such that $\theta'_j(t)$ is constant in time. In the following, we work in the rotating frame, and drop the primes to lighten the notations.

The most simple version of the Kuramoto model is obtained by choosing uniform (mean-field type) couplings $K_{ik} = K/N$, such that any pair of oscillators has the same coupling. The $1/N$ scaling is included so that the sum of all coupling terms does not trivially dominate the natural frequency in Eq. 3.287. The evolution of θ_j is then given by

$$\frac{d\theta_j}{dt} = \omega_j + \frac{K}{N}\sum_{k=1}^{N} \sin(\theta_k - \theta_j), \qquad j = 1, \dots, N. \tag{3.288}$$

In order to characterize the possible synchronization of the oscillators resulting from the coupling terms, it is convenient to introduce the complex order parameter $r\,e^{i\psi}$ defined as

$$r\,e^{i\psi} = \frac{1}{N}\sum_{k=1}^{N} e^{i\theta_k}. \tag{3.289}$$

In the absence of synchronization, the (mean) value of this order parameter is equal to zero, while the presence of synchronization is indicated by a value $r > 0$, the phase ψ corresponding to the "average" phase of the oscillators. It is convenient to reformulate Eq. 3.288 as

$$\frac{d\theta_j}{dt} = \omega_j + Kr\sin(\psi - \theta_j), \qquad j = 1, \ldots, N, \tag{3.290}$$

using the fact that from Eq. 3.289,

$$r\,e^{i(\psi-\theta_j)} = \frac{1}{N}\sum_{k=1}^{N} e^{i(\theta_k-\theta_j)} \tag{3.291}$$

for any j, and taking the imaginary part of Eq. 3.291.

We shall now focus on the limit of a very large number of coupled oscillators, $N \to \infty$. In this case, the natural frequencies are described by the density $g(\omega)$, which means that the fraction of oscillators having a natural frequency ω_j in the infinitesimal range $[\omega, \omega + d\omega]$ is $g(\omega)d\omega$. The density $g(\omega)$ is normalized as $\int_{-\infty}^{\infty} g(\omega)d\omega = 1$. In addition, the statistics of the phases of oscillators having a given frequency ω is encoded into the time-dependent probability distribution $\rho(\theta|\omega, t)$. This distribution, normalized according to $\int_{-\infty}^{\infty} \rho(\theta|\omega, t)d\theta = 1$, describes the statistics of a set of identical oscillators having different initial conditions. Taking into account Eq. 3.290, the evolution of the distribution $\rho(\theta|\omega, t)$ is governed by the equation[2]

$$\frac{\partial\rho}{\partial t}(\theta|\omega, t) + \frac{\partial}{\partial\theta}\Big[\Big(\omega + Kr\sin(\psi - \theta)\Big)\rho(\theta|\omega, t)\Big] = 0. \tag{3.292}$$

In the infinite N limit considered here, the expression (3.289) of the order parameter reduces to

$$r\,e^{i\psi} = \langle e^{i\theta}\rangle \equiv \int_{-\pi}^{\pi} d\theta \int_{-\infty}^{\infty} d\omega\, e^{i\theta}\rho(\theta|\omega, t)g(\omega). \tag{3.293}$$

In the following, we look for steady-state solutions and study whether the oscillators get synchronized or not in this regime, depending on the coupling strength K.

---

[2] This equation can be thought of as a Fokker–Planck equation (see Sect. 2.2.3) in the zero noise limit.

3.3.2 Synchronized Steady State

In order to find the steady-state solution of the model, we need to find for all frequency ω the time-independent distribution $\rho(\theta|\omega)$ solution of Eq. 3.292, in which r and ψ are self-consistently determined from Eq. 3.293. It can easily be checked that the uniform distribution $\rho(\theta|\omega) = (2\pi)^{-1}$, which leads to r=0, is a solution of Eq. 3.292 for all coupling strength K. This solution corresponds to a complete lack of synchronization between oscillators. While such a situation is likely to be relevant at low coupling, it is however possible that other solutions exist if the coupling strength K is strong enough.

To look for such possible solutions, we start from a given value of the order parameter $re^{i\psi}$ with $r > 0$, determine the solution of Eq. 3.292 for these values of r and ψ, and then check whether a self-consistent solution of Eq. 3.293 can be found. We first note that if a stationary solution with global phase ψ exists, then another steady-state solution of phase $\psi + \alpha$ can be obtained by shifting all the phases θ_j by the same amount α. Hence we can restrict our study to the case $\psi = 0$, the other cases being deduced by a simple phase shift.

Under this assumption, the steady-state solution of Eq. 3.292 satisfies

$$(\omega - Kr \sin\theta)\,\rho(\theta|\omega) = C \tag{3.294}$$

where C is a constant. The condition $\rho(\theta|\omega) \geq 0$ implies that such a solution exists only if $|\omega| \geq Kr$. The case $|\omega| = Kr$ is further excluded as it would lead to a non-normalizable distribution $\rho(\theta|\omega)$. As a result, one finds

$$\rho(\theta|\omega) = \frac{1}{2\pi}\,\frac{\sqrt{\omega^2 - (Kr)^2}}{|\omega - Kr\sin\theta|}, \qquad |\omega| > Kr. \tag{3.295}$$

If $|\omega| \leq Kr$, the distribution (3.295) is no longer valid. We leave aside the discussion of the marginal case $|\omega| = Kr$, which plays no role in the following, and focus on the situation $|\omega| < Kr$. In this case, the evolution equation (3.290) has two fixed points, solutions of

$$\omega - Kr\sin\theta = 0. \tag{3.296}$$

To check the linear stability of a fixed point θ_0, we set $\theta = \theta_0 + \epsilon$, with $\epsilon \ll 1$. Expanding Eq. 3.290 to linear order in ϵ, we get

$$\frac{d\epsilon}{dt} = -(Kr\cos\theta_0)\,\epsilon, \tag{3.297}$$

so that the fixed point θ_0 is stable if $\cos\theta_0 > 0$ and unstable if $\cos\theta_0 < 0$. Taking into account Eq. 3.296, the stable fixed point is thus given by

$$\theta_0 = \sin^{-1}\left(\frac{\omega}{Kr}\right). \tag{3.298}$$

The distribution $\rho(\theta|\omega)$ associated to this fixed point solution is a Dirac delta function, that is an infinitely peaked solution around the fixed point:

$$\rho(\theta|\omega) = \delta\left(\theta - \sin^{-1}(\omega/Kr)\right), \qquad |\omega| < Kr. \tag{3.299}$$

Now that we have determined $\rho(\theta|\omega)$ for both $|\omega| < Kr$ and $|\omega| > Kr$, we can self-consistently determine r from Eq. 3.293, setting $\psi = 0$:

$$\begin{aligned} r = &\int_{-\pi}^{\pi} d\theta \int_{-Kr}^{Kr} d\omega\, e^{i\theta} \delta\left(\theta - \sin^{-1}(\omega/Kr)\right)\, g(\omega) \\ &+ \frac{\sqrt{\omega^2 - (Kr)^2}}{2\pi} \int_{-\pi}^{\pi} d\theta \int_{|\omega|>Kr} d\omega\, \frac{e^{i\theta}\, g(\omega)}{|\omega - Kr\sin\theta|}. \end{aligned} \tag{3.300}$$

Let us now further assume that $g(\omega)$ is an even function, that is for all $\omega,\ g(-\omega) = g(\omega)$. Using the symmetries of the sine function, it can be shown that the second integral in Eq. 3.300 is equal to zero. The first integral can be computed thanks to the properties of the δ function, namely

$$\int_a^b dx f(x)\, \delta(x - x_0) = f(x_0) \tag{3.301}$$

for any function *f*, provided that $a < x_0 < b$. One thus finds, exchanging the order of integration between θ and ω:

$$r = \int_{-Kr}^{Kr} d\omega\, g(\omega)\, e^{i \sin^{-1}(\omega/Kr)}. \tag{3.302}$$

Using the parity of $g(\omega)$, the imaginary part of the integral vanishes, and Eq. 3.302 reduces to

$$r = 2\int_0^{Kr} d\omega\, g(\omega)\, \cos\left(\sin^{-1}(\omega/Kr)\right). \tag{3.303}$$

Performing the change of variable $\omega = Kr\sin x$, one eventually finds the following self-consistent equation, taking into account the assumption $r > 0$

$$\int_0^{\pi/2} dx\, (\cos x)^2\, g(Kr\sin x) = \frac{1}{2K}. \tag{3.304}$$

The solutions of this equation depend on some generic properties of the function $g(\omega)$. In the following, we assume that $g(\omega)$ has its maximum at $\omega = 0$, that is for all $\omega \neq 0$, $g(\omega) < g(0)$. Denoting as $I(r)$ the integral on the left-hand-side of Eq. 3.304, we have for all $r > 0$, $I(r) < I(0)$. Hence if the coupling constant K is such that $(2K)^{-1} > I(0)$, Eq. 3.304 has no solution for r, while a solution $r > 0$ exists for $(2K)^{-1} < I(0)$. This defines the critical coupling $K_c = [2I(0)]^{-1}$, above which a solution $r > 0$ exists. Expanding $g(\omega)$ for small ω as

$$g(\omega) = g(0) - \frac{1}{2}|g''(0)|\omega^2 + \mathcal{O}(\omega^4), \tag{3.305}$$

with $g''(0) < 0$, one finds after some algebra the following relation, for $0 < K - K_c \ll K_c$,

$$r \sim \sqrt{\frac{16(K - K_c)}{\pi K_c^4 |g''(0)|}}. \tag{3.306}$$

The above result is valid for any regular function $g(\omega)$ having its maximum at $\omega = 0$. In the specific case

$$g(\omega) = \frac{1}{\pi} \frac{\omega_0}{\omega_0^2 + \omega^2}, \tag{3.307}$$

where $\omega_0 > 0$ is a constant, the solution of Eq. 3.304 can be given explicitly for all $K > K_c \equiv 2\omega_0$, namely:

$$r = \sqrt{1 - \frac{2\omega_0}{K}}. \tag{3.308}$$

Finally, once the steady state solutions have been determined, their stability should be studied. A detailed stability analysis goes beyond the scope of the present booklet, but the main result is that the synchronized solution $r > 0$ obtained for $K > K_c$ corresponds to the stable state of the system [8].

3.4 Collective Motion of Active Particles

Active "particles" are particles able to sustain a continuous motion thanks to some external energy input. This concept is used by physicists to describe for instance the motion of animals or bacteria. A very schematic model of active particle is a point-like particle with a velocity vector of constant modulus, but arbitrary direction. When two (or more) active particles interact, they tend to align their velocities. Such simplified physicists' models have been extensively studied through numerical

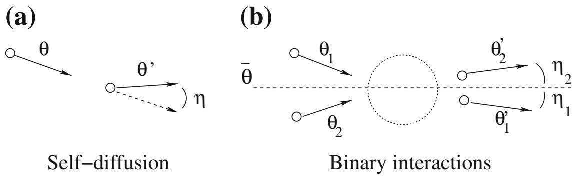

Fig. 3.3 Schematic representation of the dynamics of the model. The *dotted circle* on the right panel represents the interaction area

simulations [10–12]. A transition from disordered motion when the density of active particles is low, to ordered collective motion when the density is high, has been reported. This transition exhibits some properties similar to that of phase transitions observed in physical systems. It is also possible to develop analytical approaches, either by postulating phenomenological equations of motion at the macroscopic scale (hydrodynamic equations) [13], or by using a Boltzmann approach to derive such hydrodynamic equations [14, 15]. We present here a brief summary of the results obtained from the latter approach.

3.4.1 Definition of the Model

We consider self-propelled point-like particles moving on a continuous two-dimensional space, with a velocity vector $\mathbf{v}$ of fixed magnitude v_0 (to be chosen as the velocity unit) in a reference frame. The velocity of the particles is simply defined by the angle θ between $\mathbf{v}$ and a given reference direction. Particles move in straight line, following their velocity vector, until they experience either a self-diffusion event (a random scattering), or a binary collision that tends to align the velocities of the two particles—see Fig. 3.3. Self-diffusion events are defined as follows: the velocity angle θ of any particle is changed with a probability λ per unit time to a value $\theta' = \theta + \eta$, where η is a Gaussian noise with distribution $p_0(\eta)$ and variance σ_0^2. Binary collisions, that are the only interactions between particles, occur when the distance between two particles becomes less than d_0 (in the following, we set $d_0 = \frac{1}{2}$). The velocity angles θ_1 and θ_2 of the two particles are then changed into $\theta_1' = \overline{\theta} + \eta_1$ and $\theta_2' = \overline{\theta} + \eta_2$, as shown on Fig. 3.3b. In the last expression, $\overline{\theta} = \arg(e^{i\theta_1} + e^{i\theta_2})$ is the average angle, and η_1 and η_2 are independent Gaussian noises with the same distribution $p(\eta)$ and variance σ^2. Note that these binary collisions are different from the collisions in usual gases, as in this latter case, collisions are ruled by energy and momentum conservation laws.

3.4.2 Description Through a Boltzmann Equation

A useful mathematical tool to describe statistically the dynamics of the system is the one-particle phase-space distribution $f(\mathbf{r}, \theta, t)$, namely the probability to find a particle at position $\mathbf{r}$ and with a velocity angle θ, at time t. The evolution of this one-particle phase-space distribution is ruled by the Boltzmann equation, which reads

$$\frac{\partial f}{\partial t}(\mathbf{r},\theta,t) + \mathbf{e}(\theta)\cdot\nabla f(\mathbf{r},\theta,t) = I_{\mathrm{dif}}[f] + I_{\mathrm{col}}[f]. \tag{3.309}$$

The functionals $I_{\mathrm{dif}}[f]$ and $I_{\mathrm{col}}[f]$ respectively account for the self-diffusion and collision phenomena. The vector $\mathbf{e}(\theta)$ is the unit vector in the direction θ. The diffusion functional $I_{\mathrm{dif}}[f]$ is given by

$$\begin{aligned} I_{\mathrm{dif}}[f] = -\lambda f(\mathbf{r},\theta,t) + \lambda \int_{-\pi}^{\pi} d\theta' \int_{-\infty}^{\infty} d\eta\, p_0(\eta) \\ \times \sum_{m=-\infty}^{\infty} \delta(\theta' + \eta - \theta + 2m\pi) f(\mathbf{r},\theta',t), \end{aligned} \tag{3.310}$$

where the sum of Dirac delta functions takes into account the periodicity of the angles. The evaluation of the collision term $I_{\mathrm{col}}[f]$ is more subtle. We know that two particles collide if their distance becomes less than the interaction range d_0. In the frame of particle 1, particle 2 has a velocity $\mathbf{v}_2' = \mathbf{e}(\theta_2) - \mathbf{e}(\theta_1)$. Hence, particles that collide with particle 1 between t and $t+dt$ are those that lie, at time t, in a rectangle of length $|\mathbf{v}_2'|\,dt$ and of width $2d_0$, yielding for the collision functional

$$\begin{aligned} I_{\mathrm{col}}[f] = -f(\mathbf{r},\theta,t) \int_{-\pi}^{\pi} d\theta'\, |\mathbf{e}(\theta') - \mathbf{e}(\theta)| f(\mathbf{r},\theta',t) \\ + \int_{-\pi}^{\pi} d\theta_1 \int_{-\pi}^{\pi} d\theta_2 \int_{-\infty}^{\infty} d\eta\, p(\eta)\, |\mathbf{e}(\theta_2) - \mathbf{e}(\theta_1)| \\ \times f(\mathbf{r},\theta_1,t) f(\mathbf{r},\theta_2,t) \sum_{m=-\infty}^{\infty} \delta(\overline{\theta} + \eta - \theta + 2m\pi), \end{aligned} \tag{3.311}$$

with $\overline{\theta} = \arg(e^{i\theta_1} + e^{i\theta_2})$. One can check that the uniform angular distribution $f(\mathbf{r},\theta,t) = \rho/2\pi$ is a solution of Eq. 3.309 for an arbitrary constant density ρ, and for any noise distributions $p_0(\eta)$ and $p(\eta)$.

3.4.3 Hydrodynamic Equations and Phase Diagram

In order to deal with more convenient physical quantities, we introduce the hydrodynamic density and velocity fields $\rho(\mathbf{r},t)$ and $\mathbf{u}(\mathbf{r},t)$:

$$\rho(\mathbf{r},t) = \int_{-\pi}^{\pi} d\theta\, f(\mathbf{r},\theta,t) \tag{3.312}$$

$$\mathbf{u}(\mathbf{r},t) = \frac{1}{\rho(\mathbf{r},t)} \int_{-\pi}^{\pi} d\theta\, f(\mathbf{r},\theta,t)\, \mathbf{e}(\theta). \tag{3.313}$$

Integrating the Boltzmann equation (3.309) over θ, one directly obtains the continuity equation for $\rho(\mathbf{r},t)$:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0. \tag{3.314}$$

The operator ∇ is the vectorial differential operator[3] of components $(\partial/\partial x, \partial/\partial y)$. The derivation of a hydrodynamic equation for the velocity field is less straightforward, and involves an approximation scheme. The reader is referred to Refs. [14, 15] for more details on the derivation. Introducing for convenience the momentum field $\mathbf{w}(\mathbf{r},t) = \rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t)$, we find the following hydrodynamic equations:

$$\frac{\partial\mathbf{w}}{\partial t} + \gamma(\mathbf{w}\cdot\nabla)\mathbf{w} = -\frac{1}{2}\nabla(\rho-\kappa\mathbf{w}^2) + (\mu-\xi\mathbf{w}^2)\mathbf{w} + \nu\nabla^2\mathbf{w} - \kappa(\nabla\cdot\mathbf{w})\mathbf{w}. \tag{3.316}$$

It is interesting to give a physical interpretation of the different terms appearing in this hydrodynamic equation. The first term in the r.h.s. of Eq. 3.316 can be interpreted as a pressure gradient, considering $p = \frac{1}{2}(\rho - \kappa\mathbf{w}^2)$ as an effective pressure. The second term accounts for the local relaxation of $\mathbf{w}$, while the third term is analogous to the standard viscous term appearing in the Navier-Stokes equation describing usual fluids. Finally, the last term corresponds to a feedback on the flow from the compressibility effects.

The different coefficients appearing in Eq. 3.316 can be computed explicitly as a function of the microscopic parameters of the model. They are given by

$$\nu = \frac{1}{4}\left[\lambda\left(1-e^{-2\sigma_0^2}\right) + \frac{4}{\pi}\rho\left(\frac{14}{15}+\frac{2}{3}e^{-2\sigma^2}\right)\right]^{-1}, \tag{3.317}$$

$$\gamma = \frac{8\nu}{\pi}\left(\frac{16}{15} + 2e^{-2\sigma^2} - e^{-\sigma^2/2}\right), \tag{3.318}$$

---

[3] More explicitly, Eq. 3.314 reads

$$\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial x}(\rho u_x) + \frac{\partial}{\partial y}(\rho u_y) = 0, \tag{3.315}$$

where (u_x, u_y) are the components of the vector $\mathbf{u}$.

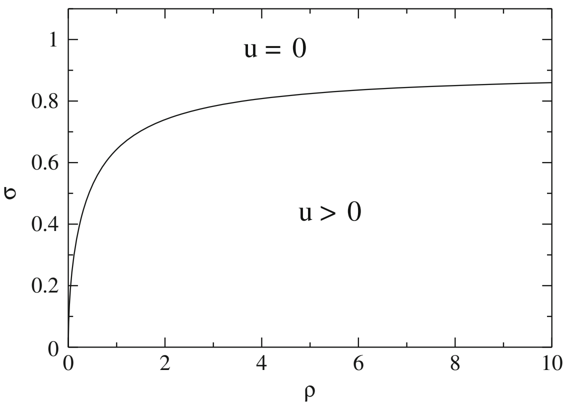

Fig. 3.4 Phase diagram of the model in the noise-density plane. A transition line separates the domains with zero hydrodynamic velocity, $u = |\mathbf{u}| = 0$, from the domain of collective motion, $u > 0$. Parameters: $\lambda = 1,\ d_0 = 0.5,\ v_0 = 1$.

$$\kappa = \frac{8\nu}{\pi}\left(\frac{4}{15} + 2e^{-2\sigma^2} + e^{-\sigma^2/2}\right), \tag{3.319}$$

$$\mu = \frac{4}{\pi}\rho\left(e^{-\sigma^2/2} - \frac{2}{3}\right) - \lambda\left(1 - e^{-\sigma_0^2/2}\right), \tag{3.320}$$

$$\xi = \frac{64\nu}{\pi^2}\left(e^{-\sigma^2/2} - \frac{2}{5}\right)\left(\frac{1}{3} + e^{-2\sigma^2}\right). \tag{3.321}$$

Note that ν, γ and κ are always positive; μ can change sign, and $\xi > 0$ whenever $\mu > 0$.

Turning to the study of the spontaneous onset of collective motion in the present model, we look for possible instabilities of the spatially homogeneous flow, that is the appearance of a uniform, nonzero, velocity field $\mathbf{u}$ (or momentum field $\mathbf{w}$). Considering a time-dependent, but spatially homogeneous flow, we get

$$\frac{\partial \mathbf{w}}{\partial t} = (\mu - \xi \mathbf{w}^2)\mathbf{w}. \tag{3.322}$$

Obviously, $\mathbf{w} = 0$ is a solution for arbitrary values of the coefficients. However, this solution becomes unstable for $\mu > 0$, when a nonzero solution $\mathbf{w}_0 = \sqrt{\mu/\xi}\,\mathbf{e}$ appears ($\mathbf{e}$ is a unit vector pointing in a arbitrary direction). From the expression (3.320) of μ, it turns out that $\mu = 0$ corresponds to a threshold value ρ_t given by

$$\rho_t = \frac{\pi\lambda(1 - e^{-\sigma_0^2/2})}{4(e^{-\sigma^2/2} - \frac{2}{3})}. \tag{3.323}$$

The transition line defined by ρ_t in the plane (ρ, σ) is plotted on Fig.3.4, in the case $\sigma_0 = \sigma$. The instability is seen to occur at any density, provided the noise is low enough. The transition line saturates at a value $\sigma_t = (2\ln\frac{3}{2})^{1/2} \approx 0.90$. Note that if one would instead fix the noise σ_0 to a given value, the transition to collective motion observed when decreasing σ would disappear below a finite density $\rho_t^0 = \frac{3}{4}\pi\lambda(1 - e^{-\sigma_0^2/2})$.

As a conclusion, we have seen that the numerically observed transition to collective motion in simple models of self-propelled particles can be obtained from theoretically-derived hydrodynamic equations. We note however that further instabilities leading to more complicated patterns are also observed [11, 14, 15], and analytical approaches to these instabilities are still an open issue.

References

1. S.F. Edwards, R.B.S. Oakeshott, Physica A **157**, 1080 (1989)
2. T.C. Schelling, J. Math. Sociol. **1**, 143 (1971)
3. S. Grauwin, E. Bertin, R. Lemoy, P. Jensen, Proc. Natl. Acad. Sci. **106**, 20622 (2009)
4. H. Callen, *Thermodynamics and An Introduction to Thermostatistics* (Wiley, New York, 1985)
5. M.R. Evans, T. Hanney, J. Phys. A **38**, R195 (2005)
6. B. Derrida, M.R. Evans, V. Hakim, V. Pasquier, J. Phys. A **26**, 1493 (1993)
7. M.R. Evans, S.N. Majumdar, R.K.P. Zia, J. Phys. A **39**, 4859 (2006)
8. J.A. Acebrón, L.L. Bonilla, C.J. Pérez Vicente, F. Ritort, R. Spigler, Rev. Mod. Phys. **77**, 137 (2005)
9. Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer, New York, 1984)
10. T. Vicsek et al., Phys. Rev. Lett. **75**, 1226 (1995)
11. G. Grégoire, H. Chaté, Phys. Rev. Lett. **92**, 025702 (2004)
12. J. Toner, Y. Tu, S. Ramaswamy, Ann. Phys. **318**, 170 (2005)
13. J. Toner, Y. Tu, Phys. Rev. Lett. **75**, 4326 (1995)
14. E. Bertin, M. Droz, G. Grégoire, Phys. Rev. E **74**, 022101 (2006)
15. E. Bertin, M. Droz, G. Grégoire, J. Phys. A: Math. Theor. **42**, 445001 (2009)