Statistical Mechanics
Master de Physique fondamentale et appliquée – Spécialité : Champs, particules et matière

Leticia F. Cugliandolo
Laboratoire de Physique Théorique et Hautes Energies de Jussieu
Laboratoire de Physique Théorique de l'Ecole Normale Supérieure
Membre de l'Institut Universitaire de France
`leticia@lpt.ens.fr`

20 November 2005

Abstract

Ce cours présente une introduction à quelques problèmes d'intérêt actuel en mécanique statistique. Chaque chapitre est focalisé sur un sujet. On expliqera ces sujet en s'appuyant sur les résultats d'expériences. On présentera leur modélisation ainsi qu'une sélection de méthodes théoriques utilisées pour leur étude.

This document includes: (1) A detailed schedule of the lectures, TDs and exam. (2) A **draft** of the Lecture notes on Statistical Mechanics. It presents a summary of the material that will be described during the semester. They are certainly incomplete and may contain errors. Hopefully, we shall improve it with the help of the students. (3) The TDs.

1 Programme

Les TDs sont conçus pour être faits à la maison après une aide donnée en cours. Leurs solutions devront être rendus et cela fera parti de la note finale.

1. Introduction Rappel des notions de probabilités et statistique [1, 2, 3, 4].

   Quelques applications : l'expérience de Luria-Delbrück [5]; statistiques d'extrêmes [6, 7]; un paradigme, le modèle d'Ising et ses applications interdisciplinaires [8, 9, 10, 11, 12].

   TD 1 : exercises simples de rappel.

   Lecture : l'article de Luria-Delbrück [5]; structures dans l'univers [14].

2. Transitions de phase [8, 9, 10, 11, 12]

   Exemples : vapeur-liquide, paramagnétique-ferromagnétique, DNA. Classification des transitions.

   Approche de champ moyen.

   Phénomènes critiques, invariance d'échelle, introduction au groupe de renormalisation [10, 11, 13].

   TD 2 : modèle d'Ising ferromagnétique complètement connecté avec des interactions à p spins, $(p = 2)$ transition de second ordre, $(p > 2)$ transition de premier ordre.

   Lecture : Transitions de phase, un exemple mécanique [15].

3. Désordre [16, 17, 18, 19, 20]

   Définition de désordre recuit et désordre gelé (exemples). Leur traitement statistique.

   Désordre gelé : effets de la frustration, auto-moyennage de l'énergie libre, introduction à la méthode des repliques.

   Exemples : verres de spin, potentiels aléatoires, réseau de neurones.

   TD 3: La chaîne de spins avec des interactions désordonnées.

   Lecture : Le problème de la rupture [21]

4. Processus stochastiques [22, 23]

   Mouvement Brownien. L'équation maîtresse, l'approche de Langevin et Fokker-Planck.

   TD 4 : Ratchets, application à l'electrophorèse.

5. Croissance de surfaces et d'interfaces [24]

   Exemples. Rugosité. Lois d'échelle dynamiques.

   Le 'random deposition model' : solution, limite continue et présentation des équations de Edwards-Wilkinson et de Kardar-Parisi-Zhang.

   TD 5 : Étude de l'équation d'Edwards-Wilkinson.

2 Introduction

2.1 Background

Equilibrium Statistical Mechanics is a very well-established branch of theoretical physics Together with Quantum Mechanics, they form the basis of Modern Physics.

The goal of equilibrium statistical mechanics is to derive the thermodynamic functions of state of a macroscopic system from the microscopic laws that determine the behaviour of its constituents. In particular, it explains the origin of thermodynamic – and intuitive – concepts like presure, temperature, heat, *etc.*

In Table 1 we recall the typical length, time and energy scales appearing in the microscopic and macroscopic World.

	Micro		Macro	
dist (ℓ)	Solid	Gaz	10^{-3}m	
	10^{-10}m	10^{-8}m		
# part (N)	1		Solid	Gaz
			$\left(\frac{10^{-3}}{10^{-10}}\right)^{d=3} = 10^{21}$	$\left(\frac{10^{-3}}{10^{-8}}\right)^{d=3} = 10^{15}$
energy (E)	$1\,eV$		$1J \approx 6\,10^{18}eV$	
time (t)	Solid	Gaz	$1\,s$	
	$\hbar/1eV \approx 6\ 10^{-14}\,s$	$10^{-9}\,s$		

Table 1: Typical length, energy and time scales in the microscopic and macroscopic World.

A reference number is the number of Avogadro, $N_A = 6.02\ 10^{23}$; it counts the number of atoms in a mol, *i.e.* 12gr of ^{12}C, and it yields the order of magnitude of the number of molecules at a macroscopic level. The ionization energy of the Hydrogen atom is 13.6 eV and sets the energy scale in Table 1.

It is clear from the figures in Table 1 that, from a practical point of view, it would be impossible to solve the equations of motion for each one of the $N \approx N_A$ particles – let us keep the discussion *classical*, including quantum mechanical effects would not change the main conclusions to be drawn henceforth – and derive from their solution the macroscopic behaviour of the system. Moreover, the deterministic equations of motion may present a very high sensitivity to the choice of the initial conditions – deterministic chaos – and thus the precise prediction of the evolution of the ensemble of microscopic constituents becomes unfeasible even from a more fundamental point of view.

The passage from the microscopic to the macroscopic is then done with the help of *Statistical methods*, *Probability Theory* and, in particular, the *Law of Large Numbers.* It assumes – and it has been quite well confirmed – that there are no big changes in the fundamental Laws of Nature when going through all these orders of magntide. However, a number of new and interesting phenomena arise due to the unexpected *collective* behaviour of these many degrees of freedom. For example, phase transitions when varying an external parameter occur; these are not due to a change in the form of the microscopic interactions but, rather, to the locking of the full system in special configurations.

Equilibrium statistical mechanics also makes another very important assumption that we shall explain in more detail below: that of the *equilibration* of the macroscopic system. Some very interesting systems do not match this hypothesis. Still, one would like to use Probabilistic arguments to characterize their macroscopic behavior. This is possible in a number of cases and we shall discuss some of them. Indeed, deriving a theoretical framework to describe the behavior of macroscopic systems out of equilibrium is one the present major challenges in theoretical physics.

2.2 This course

In this set of lecture we shall discuss same problems in *equilibrium* and *dynamic* statistical mechanics that either are not fully understood or receive the attention of researchers at present due to their application to problems of interest in physics and other areas of science. The plan of the set of lectures is the following:

In the first Chapter we recall some aspects of Probability Theory and Statistical Mechanics. In Appendix A we recall the definition of a probability, its main properties, and the main probability functions encountered in hard Science. In the main part of the text we show the utility of Probability arguments by discussing the experiment of Luria and Delbrück – considered the founder of Molecular Biology, the authors received the Nobel Prize in 1969 – and some aspects of the laws of large numbers and extreme value statistics. Basic features of the foundations of Statistical Mechanics are recalled next. Finally, we define the Ising model and list some of its more important properties that render it such a standard model.

In the second Chapter we describe the theory of phase transitions: first we explain the mean-field approach, and then we introduce fluctuations and show how the importance of these led to the development of the renormalization group.

The third Chapter is devoted to the discussion of disorder and its effects in phase transitions and low-temperature behavior. We briefly describe the three main routes to describe these systems analytically: scaling, the replica method and functional renormalization group ideas.

In the last two Chapters we introduce *time* into the discussion. First, in Chapter 4 we define stochastic processes, the Langevin and Fokker-Planck formalism and we briefly discuss the dynamics of macroscopic systems close to thermal equilibrium. Next, in Chapter 5 we treat the problem of the random growth of a surface. We introduce some physical examples and two simple models, we discuss the behavior in terms of scaling laws and we also solve one of these models analytically while describing what fails in the other and the need to use improved renormalization group ideas out of equilibrium.

It is clear that the correct explanation of all of these problems and alytical method would need many more hours of teaching. We shall only give the main ingredients of each of them and provide the interested students with references to deepen their knowledge of these subjects. The cours is intimately related to the one of J-M di Meglio (first semester) and O. Martin (second semester).

3 Basic notions

3.1 Applications of probability theory

3.1.1 The Luria-Delbrück experiment

This is a very cute example of the use of probability concepts – and more generality, of a training in Physics! – in varied problems in Science.

This experiment provided the first proof of *genetic mutation* in bacteria. It works as follows. Colonies of bacteria, typically including 10^9 members, are exposed to a virus that kills most of them. However, after some time, some new bacteria appear showing that either
(i) some bacteria acquire resistance to the virus (adaptation, Lamarckian hypothesis), or
(ii) their ancestors already possessed resistance *via* mutation,
when they were exposed to the virus.

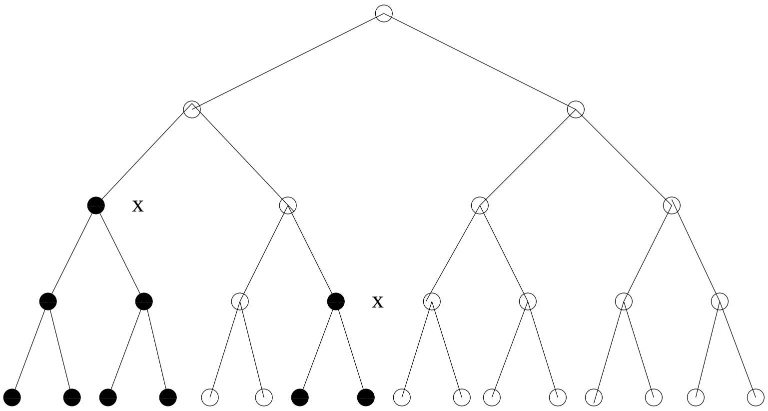


Figure 1: A tree representing the growth of a colony of bacteria under the mutation hypothesis. Each level in the tree corresponds to a generation. Time grows going down along the vertical direction. The nodes represent individuals in the colonies. The individuals shown with a cross are the mutant ones. The individuals painted white are sensitive to the virus while the ones painted black are resistant.

How can one distinguish between the two scenarii? It is clear that if the adaptive hypothesis held true, the spatial distribution of the resistant bacteria in the sample would be uniform and, moreover, the number of resistant bacteria would not increase with the age of the sample. If, in contrast, the mutation hypothesis held true, then, the number of resistant bacteria should be larger for an older sample and, they should appear in groups of individuals related by inheritage. Luria developed an experimental analysis, based on the smart use of Statistical analysis (see Appendix A) that allowed him to determine that the mutation hypothesis is correct.

Luria's idea was to study finite-size sample-to-sample fluctuations to distinguish between adaptation and mutation. Take a colony of bacteria of a given age, say it contains N individuals, and divide it in M samples with $\approx N/M$ individuals each. Expose now each sample to virus and count how many resistant bacteria, n_k with $k = 1, \dots, M$, are in each of them. If the adaptation hypothesis is correct, all sample react in roughly the same way, and n should be described by a Poisson distribution characterized by an average that coincides with the mean-square deviation, $\langle n \rangle = \langle (n - \langle n \rangle)^2 \rangle = \mu$, with μ the parameter of the Poisson distribution. If, instead, the mutation hypothesis is correct some samples should have many more resistant bacteria than others. Thus, the mean-square-displacement should be much larger than the average.

Let us discuss the theory behind Luria's argument in a bit more detail. Remember the usual calculation of the density fluctuations in a gas. Let N and V be the number of particles and volume of a gas and n be the number of particles in a very small volume v within the bulk, $v \ll V$. Since the gas is uniform, the probability that any given particle is in v is just v/V and the probability that n given particles are in v is just $(v/V)^n$. Similarly, the probability that a particle is not in v is $1 - v/V$ and the probability that $N - n$ particles are not in v is $(1 - v/V)^{N-n}$. The probability of having *any* n particles in v is given by the binomial formula:

$$P(n, v; N, V) = \frac{N!}{n!(N-n)!} \left(\frac{v}{V}\right)^n \left(1 - \frac{v}{V}\right)^{N-n} . \tag{3.1}$$

The prefactor gives the number of ways of choosing n particles from the total number N. Now, this expression can be simplified, and becomes Poisson's law, in the limits $v \ll V$, $n \ll N$:

$$P(n) = \frac{\langle n \rangle^n e^{-\langle n \rangle}}{n!} . \tag{3.2}$$

(In the third factor one replaces the exponent $N - n \sim N$; the first factor is approximated by using Stirling formula for the factorials, see Appendix [?]; and then $\langle n \rangle \equiv Nv/V$ as one can also verify by computing the average directly from this expression.) Note that n can be very different from its mean value, $\langle n \rangle$ but it is still much smaller than N that has been taken to infinity in this calculation. From this expression one can compute the average $\langle n^2 \rangle$ and the mean-square fluctuations of the number of particles

$$\sigma^2 \equiv \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle . \tag{3.3}$$

This is the main characteristics of the Poisson law exploited by Luria.

Luria carried out these experiments and he found the desired result, the sample-to-sample fluctuations were much larger than predicted from a Poisson distribution and hence he concluded that the mutation hypothesis was correct. In collaboration with Delbrück they refined the analysis and even found a way to estimate the mutation rate from their experimental observations. They got the Nobel Prize in Medicine in 1969.

3.1.2 Statistics of extremes

The statistics of extreme values is currently appearing in a number of interesting problems in hard and applied Sciences. In all sorts of applications, accurate risk assessment relies on the effective evaluation of the extremal behavior of the process under study. Unlike most of Statistics which tries to say something about typical behavior, Extreme Value Statistics attempts to characterize unlikely behavior, or at least to say how unlikely the behavior is. Applications include: flood risk assessment; financial risk management; insurance assessment; setting industrial safety standards; the prediction of extreme weather conditions. In the context of our lectures, we shall see Statistics of extremes appearing in Phase transitions and diffusion processes.

The mathematical question is the following. Let us study a sequence $s_N \equiv x_1, \dots, x_N$ of realizations of a random variable x with a probability distribution function $p(x)$. Each entry x_i is an independent identically distributed random number. A natural question to ask is what is the maximum value taken by x, *i.e.* what is $x_{max} \equiv \max(x_1, x_2, \dots, x_N)$? it is clear that x_{max} is itself a random variable (since if we drew different sequences s_N we would obtain different values of x_{max}). We then need to characterize x_{max} in probability and the quantity we need to determine is its probability distribution function, $q(x_{max})$.

This kind of question was raised in the context of studies funded by insurance companies in the Netherlands... it was useful to know how high could it be a raise in the sea level... Emil Gumbel developed part of the theory of extreme value statistics. Quoting him *"It seems that the rivers know the theory. It only remains to convince the engineers of the validity of this analysis."* or also *"Il est impossible que l'improbable n'arrive jamais."* With the computational advances and software developed in recent years, the application of the statistical theory of extreme values to weather and climate has become relatively straightforward. Annual and diurnal cycles, trends (*e.g.*, reflecting climate change), and physically-based covariates (*e.g.*, El Niño events) all can be incorporated in a straightforward manner.

Let us call $f(\Lambda)$ the cumulative probability,

$$f(\Lambda) = \int_{-\infty}^{\Lambda} dx\, p(x) \;, \tag{3.4}$$

i.e. the probability of x being smaller than Λ and $g(\Lambda)$ the cumulative probability of x_{max}:

$$g(\Lambda) = \int_{-\infty}^{\Lambda} dx_{max}\, q(x_{max}) \;. \tag{3.5}$$

Now, in order to have $x_{max} < \Lambda$ one needs to have all x's smaller than Λ. Thus,

$$g(\Lambda) = [f(\Lambda)]^N \tag{3.6}$$

Let us call $h(\Lambda)$ the probability that x is larger than Λ,

$$h(\Lambda) = 1 - f(\Lambda) \tag{3.7}$$

If Λ is a large number, $h(\Lambda)$ is expected to be small, and

$$g(\Lambda) = [1 - h(\Lambda)]^N \sim 1 - Nh(\Lambda) \sim e^{-Nh(\Lambda)} \tag{3.8}$$

One needs to evaluate this expression for different functions $h(\Lambda)$. It turns out that for all $p(x)$ that are not bounded and that fall off to zero faster than exponentially one has the *Gumbel* distribution of the maximum

$$q(x_{max}) = b e^{-b(x_{max}-s) - e^{-b(x_{max}-s)}} \;. \tag{3.9}$$

The two parameters b and s fix the mean and the variance of q and depend on the ones of x. Note that a better representation is obtained using the reduced variable $b(q_{max} - s)$ and a logarithmic scale.

Actually one can also derive the pdf of the a-th value in the sequence $a = 1$ being the maximum, $a = 2$ the next one, and so on and so forth, to find

$$q(x_a) = \mathcal{N} e^{-ab(x_a - s) - e^{-b(x_a - s))}} \tag{3.10}$$

with $\mathcal{N}$ the normalization constant that depends on a, b and s.

Other $p(x)$ with different decaying forms at infinity (slower than exponential) fall into different classes of extreme value statistics (Fréchet, Weibull).

We shall see extreme value statistics play a role in the context of critical phenomena (Chapter 2) and surface growth (Chapter 5).

3.2 Elements in statistical mechanics

Let us here recall some important features of Statistical Mechanics [1, 2, 3, 4].

The state of a classical system made of $i = 1, \ldots, N$ particles living in d-dimensional space is fully characterized by a point in the $2dN$ dimensional phase space Γ made of the coordinates, $\vec{q}$, and momenta, $\vec{p}$, of the particles $(\vec{q}, \vec{p}) \equiv (q_1^1, q_1^2, q_1^3, q_2^1, q_2^2, q_2^3, \ldots, q_N^1, q_N^2, q_N^3, p_1^1, p_1^2, p_1^3, p_2^1, p_2^2, p_2^3, \ldots, p_N^1, p_N^2, p_N^3)$. The Hamiltonian of the system is $H(\vec{q}, \vec{p})$ and the time evolution is determined by Hamilton's equation of motion (equivalent to Newtonian dynamics, of course). As time passes the representative point $(\vec{q}(t), \vec{p}(t))$ traces a path in Γ. Energy, E, is conserved if the Hamiltonian does not depend on time explicitly and thus all points in any trajectory lie on a constant energy surface, $H = E$.

Liouville's theorem states that a volume element in phase space does not change in the course of time if each point in it follows the microscopic Hamilton laws of motion. It is pretty easy to show just by computing the Jacobian of a change of variables corresponding to the infinitesimal evolution of a little volume $d\Gamma$ dictated by the equations of motion.

If the macroscopic state of the system is characterized by the number of particles N, the volume V and the energy E that, say, lies between E and $E + dE$, all *microstates*, *i.e.* all configurations on the constant energy surface, are equivalent. We can think about all these microstates as being (many) independent *copies* of the original system. This is Gibbs' point of view, he introduced the notion of *ensemble* as the collection of mental copies of a system in identical macroscopic conditions. Different ensembles, correspond to different choices of the parameters characterizing the macroscopic state, (N, V, E) in the microcanonical, (N, V, T) in the canonical and (μ, V, T) in the macrocanonical, with T temperature and μ the chemical potential.

But, how can one describe the evolution of the system in phase space? In practice, given a macroscopic system with $N \gg 1$, one cannot determine the position and momenta of all particles with great precision – uncertainty in the initial conditions, deterministic chaos, *etc.* A probabilistic element enters into play since what one can do is estimate the *probability* that the representative point of the system is in a given region of Γ. Indeed, one introduces a time-dependent probability density $\rho(\vec{q}, \vec{p}; t)$ such that $\rho(\vec{q}, \vec{p}; t) d\Gamma$ is the probability that the representative point is in a region of volume $d\Gamma$ around the point $(\vec{q}, \vec{p})$ at time t. Probability behaves like an incompressible fluid in phase space. The determinist equations of motion for $(\vec{q}, \vec{p})$ allow us to derive the Liouville deterministic equation for the evolution of ρ:

$$\frac{\partial \rho}{\partial t} = -i \left(\frac{\partial H}{\partial p_j^a} \frac{\partial}{\partial q_j^a} - \frac{\partial H}{\partial q_j^a} \frac{\partial}{\partial p_j^a} \right) , \tag{3.11}$$

with the summation convention over repeated indices (i labels particles and a labels coordinates).

Note that if initially one knows the state of the system with great precision, the initial ρ will be concentrated in some region of phase space. At later times, ρ can still be localized – perhaps in a different region of phase – or it can spread. This depends on the system. Following Gibbs, the probability density ρ is interpreted as the one obtained within the (microcanonical) ensemble.

It is important to note that Liouville's equation remains invariant under time-reversal, $t \to -t$ and $\vec{p} \to -\vec{p}$. Thus, for generic initial conditions its solutions oscillate in time and do not approach a single asymptotic stationary solution that could be identified with equilibrium. The problem of how to obtain irreversible decay from Liouville's equation is a fundamental one in Statistical Mechanics. We shall come back to this problem in Chapter 4. Now, let us mention an attempt to understand the origin of irreversibility in terms of flows in phase space, namely, *ergodic theory*, founded by Boltzmann by the end of the XIXth century [3].

In the absence of a good way to determine the evolution of ρ and its approach to a stationary state, we can simply look for stationary solutions of eq. (3.11), *i.e.* ρ such that the right-hand-side vanishes. The simplest such solution is given by a ρ that depends on the energy E only, $\rho(E)$. Even if it is very difficult to show, this solution is very smooth as a function of $(\vec{q}, \vec{p})$ and it is then the best candidate to describe the *equilibrium* state – understood as the one that corresponds to the intuitive knowledge of equilibrium in thermodynamics. In short, one *postulates* that all points in the energy surface E are equally likely – there is *a priori* no reason why some should be more probable than others! – and one proposes the microcanonical measure:

$$\rho(E) \;=\; \begin{cases} \rho_0 & \text{if } H \in (E, E+dE) \;, \\ 0 & \text{otherwise} \;, \end{cases} \tag{3.12}$$

and then constructs all the Statistical Mechanics machinery on it, constructing the other ensembles, showing that thermodynamics is recovered and so on and so forth.

Finally, let us discuss Boltzmann's and Gibb's interpretation of averages and the *ergodic hypothesis*. Boltzmann interpreted macroscopic observations as time averages of the form

$$\overline{A} \equiv \lim_{\tau \to \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} dt \; A(\vec{q}(t), \vec{p}(t)) \; . \tag{3.13}$$

However, in practice, these averages are impossible to calculate. With the introduction of the concept of ensembles Gibbs gave a different interpretation (and an actual way of computing) macroscopic observations. For Gibbs, these are averages are statistical ones over all elements of the statistical ensemble,

$$\langle\, A \,\rangle = c \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; \rho(\vec{q}, \vec{p}) A(\vec{q}, \vec{p}) \; , \tag{3.14}$$

with ρ the measure. In the microcanonical ensemble this is an average over microstates on the constant energy surface taken with the microcanonical distribution (3.12):

$$\langle\, A \,\rangle = c \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; \delta(H(\vec{q}, \vec{p}) - E) A(\vec{q}, \vec{p}) \; , \tag{3.15}$$

and the normalization constant $c^{-1} = \int \prod_{i=1}^{N} \prod_{a=1}^{d} \delta(H(\vec{q},\vec{p}) - E)$. In the canonical ensemble the is computed with the Gibbs-Boltzmann weight:

$$\langle\, A \,\rangle = Z^{-1} \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; e^{-\beta H(\vec{q},\vec{p})} A(\vec{q},\vec{p}) \; . \tag{3.16}$$

Z is the partition function $Z = \int \prod_{i=1}^{N} \prod_{a=1}^{d} dq_i^a dp_i^a \; e^{-\beta H(\vec{q},\vec{p})}$.

The *(weak) ergodic hypothesis* states that under the dynamic evolution the representative point in phase space of a classical system governed by Newton laws can get as close as desired to any point on the constant energy surface.

The *ergodic hypothesis* states that time and ensemble averages, (3.13) and (3.14) coincide in equilibrium for all reasonable observables. This hypothesis cannot be proven in general but it has been verified in a large number of cases. In general, the great success of Statistical Mechanics in predicting quantitative results has given enough evidence to access this hypothesis.

An important activity in modern Statistical Mechanics is devoted to the study of macroscopic systems that do not satisfy the ergodic hypothesis. A well-understood case is the one of phase transitions and we shall discuss it in the next section. Other cases are related to the breakdown of the equilibration. This can occur either because they are externally driven or because they start from an initial condition that is far from equilibrium and their interactions are such that they do not manage to equilibrate. One may wonder whether certain concepts of thermodynamics and equilibrium statistical mechanics can still be applied to the latter problems. At least for cases in which the macroscopic dynamics is slow one can hope to derive an extension of equilibrium statistical mechanics concepts to describe their behavior.

Finally, let us remark that it is usually much easier to work in the canonical ensemble both experimentally and analytically. Thus, in all our future applications we assume that the system is in contact with a heat reservoir with which it can exchange energy and that keeps temperature fixed.

3.3 The Ising model

The Ising model is a mathematical representation of a magnetic system. It describes magnetic moments as classical *spins*, s_i, taking value ± 1, lying on the vertices of a cubic lattice in d dimensional space, and interacting via nearest-neighbor couplings, $J > 0$. The energy is then

$$E = -J \sum_{\langle ij \rangle} s_i s_j - h \sum_i s_i \tag{3.17}$$

where h is an external magnetic field.

The Ising model is specially attractive for a number of reasons:
(i) It is probably the simple example of modeling to which a student is confronted.
(ii) It can be solved in some cases: $d = 1$ (Chapter 2), $d = 2$, $d \to \infty$ (Chapter 2). The solutions have been the source of new and powerful techniques later applied to a variety of different problems in physics and interdisciplinary fields.
(iii) It has not been solved analytically in the most natural case, $d = 3$!
(iv) It has a phase transition, a interesting collective phenomenon, separating two phases

that are well-understood and behave, at least qualitatively, as real magnets with a paramagnetic and a ferromagnetic phase (Chapter 2).
(v) There is an upper, d_u, and lower, d_l, critical dimension. Above d_u mean-field theory correctly describes the critical phenomenon. Below d_l there is no finite T phase transition. Below d_u mean-field theory fails (Chapter 2).
(vi) One can see at work generic tools to describe the critical phenomenon like *scaling* (Chapter 2, 5) and the *renormalization group* (Chapter 2).
(vii) Generalizations in which the interactions and/or the fields are random variables taken from a probability distribution are typical examples of problems with *quenched disorder* (Chapter 3).
(viii) Generalizations in which spins are not just Ising variables but vectors with n components are also interesting: $n = 1$ (Ising), $n = 2$ (XY), $n = 3$ (Heisenberg), ... , $n \to \infty$ ($O(n)$).
(ix) One can add a dynamic rule to update the spins and we are confronted to the new World of stochastic processes (Chapter 4).
(x) Last but not least, it has been a paradigmatic model extended to describe many problems going beyond physics like neural networks, social ensembles, *etc.*

In the rest of this set of Lectures we shall discuss the physics of this model and we shall study its statics and dynamics with a number of analytic techniques.

4 Phase transitions

Take a piece of material in contact with an external reservoir. The material will be characterized by certain observables, energy, magnetization, *etc.*. To characterize macroscopic systems it is convenient to consider densities of energy, magnetization, *etc*, by diving the macroscopic value by the number of particles (or the volume) of the system. The external environment will be characterized by some parameters, like the temperature, magnetic field, pressure, *etc.* In principle, one is able to tune the latter and the former will be a function of them.

Sharp changes in the behavior of macroscopic systems at critical point (lines) in parameter space have been observed experimentally. These correspond to phase transitions, a non-trivial collective phenomenon appearing in the thermodynamic limit. In this Section we shall review the main features of, and analytic approaches used to study, phase transitions.

4.1 Order and disorder

When one cools down a magnetic sample it undergoes a sharp change in structure, as shown by a sharp change in its macroscopic properties, at a well-defined value of the temperature which is called the *critical temperature* or the Curie temperature. Assuming that this *annealing* process is done in equilibrium, that is to say, that at each temperature step the system manages to equilibrate with its environment after a relatively short transient – an assumption that is far from being true in *glassy* systems but that can be safely assumed in this context – the two states above and below T_c are equilibrium states that can be studied with the standard Statistical Mechanics tools.

More precisely, at T_c the equilibrium magnetization density changes from 0 above T_c to a finite value below T_c, see Fig. 2. The high temperature state is a *disordered* paramagnet while the low temperature state is an *ordered* ferromagnet.

One identifies the magnetization density as the *order parameter* of the phase transition. It is a macroscopic observable that vanishes above the transition and takes a continuously varying value below T_c. The transition is said to be continuous since the order parameter grows continuously from zero at T_c.

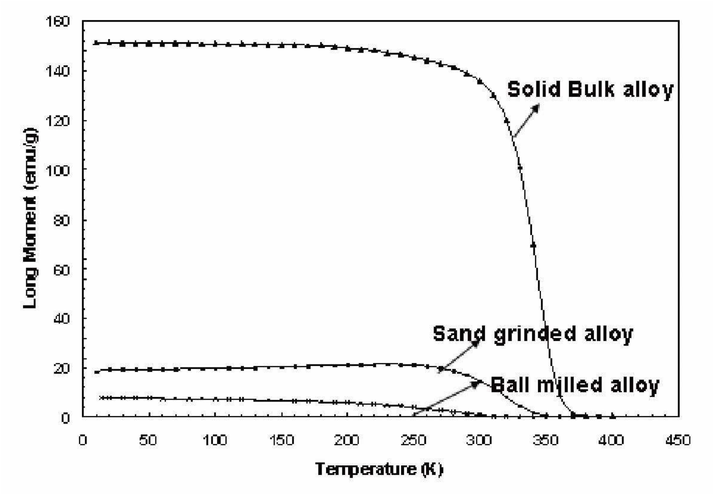
If one looks in more detail into the behavior of the variation of the magnetization density close T_c one would realize that the magnetic *susceptibility*,

$$\left.\frac{\partial m_h}{\partial h}\right|_{h=0} = \left.\frac{\partial}{\partial h}\left(-\frac{\partial}{\partial h} f_h\right)\right|_{h=0} \tag{4.1}$$

i.e. the linear variation of the magnetization density with respect to its conjugate magnetic field h diverges when approaching the transition from both sides. As the second identity shows, the susceptibility is just a second derivative of the free-energy density. Thus, a divergence of the susceptibility indicates a *non-analyticity* of the free-energy density. This can occur only in the infinite volume or *thermodynamic limit*, $N \to \infty$. Otherwise the free-energy density is just the logarithm of the partition function, a finite number of terms that are exponentials of analytic functions of the parameters, and thus an analytic function of the external parameters itself.

Figure 1 http://www.biomagres.com/content/2/1/4/figure/F1?highres=y

Figure 1 Resolution: standard / high



Magnetization and magnetic phase transition (Curie temperature) of the as-produced, sand grinded and ball-milled copper-nickel alloy (29% wt. copper, 71% wt nickel), from top to bottom, respectively.

Figure 2: The magnetization as a function of temperature for three magnetic compounds.

What is observed near such a critical temperature are called *critical phenomena*. Since the pioneering work of Curie, Langevin and others, the two phases, paramagnetic and ferromagnetic are well-understood. Qualitative arguments (see Sect. 4.3) as well as the mean-field approach (see Sect. 4.6) captures the two phases and their main characteristics. However, what happens close to the critical point has remained difficult to describe quantitatively until the development of scaling and the renormalization group (Sect. 4.8).

4.2 Discussion

1 of 1 10/18/05 14:30

Let su discuss some important concepts, pinning fields, broken ergodicity and broken symmetry, with the help of a concrete example, the Ising model (3.17). The dicuss ion is however much more general and introduces the concepts mentioned above.

4.2.1 Pinning field

In the absence of a magnetic field for pair interactions the energy is an even function of the spins, $E(\vec{s}) = E(-\vec{s})$ and, consequently, the equilibrium magnetization density computed as an average over *all* spin configurations with their canonical weight, $e^{-\beta H}$, vanishes at all temperatures.

At high temperatures, $m = 0$ characterizes completely the equilibrium properties of the system since there is a unique paramagnetic state with vanishing magnetization density. At low temperatures instead if we perform an experiment we *do observe* a net magnetization density. In practice, what happens is that when the experimenter takes the system through the transition one cannot avoid the application of tiny external fields – the experimental set-up, the Earth... – and there is always a small *pinning field* that actually selects one of the two possible equilibrium states, with positive of negative magnetization density, allowed by symmetry. In the course of time, the experimentalist should see the full magnetization density reverse, however, this is not see in practice since astronomical time-scales would be needed. We shall see this phenomenon at work when solving mean-field models exactly below.

4.2.2 Broken ergodicity

Introducing dynamics into the problem [1], ergodicity breaking can be stated as the fact that the temporal average over a long (but finite) time window is different from the statical one, with the sum running over all configurations with their associated Gibbs-Boltzmann weight:

$$\overline{A} \neq \langle A \rangle \; . \tag{4.2}$$

In practice the temporal average is done in a long but finite interval $\tau < \infty$. During this time, the system is positively or negatively magnetized depending on whether it is in "one or the other degenerate equilibrium states". Thus, the temporal average of the orientation of the spins, for instance, yields a non-vanishing result $\overline{A} = m \neq 0$. If, instead, one computes the statistical average summing over *all* configurations of the spins, the result is zero, as one can see using just symmetry arguments. The reason for the discrepancy is that with the time average we are actually summing over half of the available configurations of the system. If time τ is not as large as a function of N, the trajectory does not have enough time to visit all configurations in phase space. One can reconcile the two results by, in the statistical average, summing only over the configurations with positive (or negative) magnetization density. We shall see this at work in a concrete calculation below.

4.2.3 Spontaneous broken symmetry

In the absence of an external field the Hamiltonian is symmetric with respect to the simultaneous reveral of all spins, $s_i \to -s_i$ for all i. The phase transition corresponds to a *spontaneous symmetry breaking* between the states of positive and negative magnetization. One can determine the one that is chosen when going through T_c either by applying a small *pinning field* that is taken to zero only after the thermodynamic limit, or by imposing adequate boundary conditions like, for instance, all spins pointing up on the borders of the sample. Once a system sets into one of the equilibrium states this is completely stable in the $N \to \infty$ limit.

Ergodicity breaking necessarily accompanies spontaneous symmetry breaking but the reverse is not true; see [9] for an example and the discussion in Sect. 5 on disordered systems. Indeed, spontaneous symmetry breaking generates disjoint ergodic regions in phase space, related by the broken symmetry, but one cannot prove that these are the only ergodic components in total generality. Mean-field spin-glass models provide a counterexample of this implication.

4.3 Energy *vs* entropy

Let us first use a thermodynamic argument to describe the high and low temperature phases.

---

[1] Note that Ising model does not have a natural dynamics associated to it. We shall see in Section 6 how a dynamic rule is attributed to the evolution of the spins.

The free energy of a system is given by $F = U - k_BTS$ where U is the internal energy, $U = \langle H\rangle$, and S is the entropy. In the following we measure temperature in units of k_B, $k_BT \to T$. The equilibrium state may depend on temperature and it is such that it minimizes its free-energy F. A competition between the energetic contribution and the entropic one may then lead to a change in phase at a definite temperature, *i.e.* a different group of microconfigurations, constituting a state, with different macroscopic properties dominate the thermodynamics at one side and another of the transition.

At zero temperature the free-energy is identical to the internal energy U. In a system with ferromagnetic couplings between magnetic moments, the magnetic interaction is such that the energy is minimized when neighboring moments are parallel. Thus the preferred configuration is such that all moments are parallel and the system is fully ordered.

Switching on temperature thermal agitation provokes the reorientation of the moments and, consequently, misalignments. Let us then investigate the opposite, infinite temperature case, in which the entropic term dominates and the chosen configurations are such that entropy is maximized. This is achieved by the magnetic moments pointing in random independent directions. For example, for a model with N Ising spins, the entropy at infinite temperature is $S \sim N \ln 2$.

Decreasing temperature disorder becomes less favorable. The existence or not of a finite temperature phase transitions depends on whether long-range order, as the one observed in the low-temperature phase, can remain stable with respect to fluctuations, or the reversal of some moments, induced by temperature. Up to this point, the discussion has been general and independent of the dimension d.

The competition argument made more precise allows one to conclude that there is no finite temperature phase transition in $d = 1$ while it suggests there is one in $d > 1$. Take a one dimensional ferromagnetic Ising model with closed boundary conditions (the case of open boundary conditions can be treated in a similar way), $H = -J \sum_{i=1}^{N} s_i s_{i+1}$, $s_{N+1} = s_1$. At zero temperature it is ordered and its internal energy is just

$$U_o = -JN \tag{4.3}$$

with N the number of links and spins. Since there are two degenerate ordered configurations the entropy is

$$S_o = \ln 2 \tag{4.4}$$

The internal energy is extensive while the entropy is just a finite number. At temperature T the free-energy of the ordered state is then

$$F_o = U_o - TS_o = -JN - T \ln 2 \; . \tag{4.5}$$

Adding a *domain* of the opposite order in the system, *i.e.* reversing n spins, two bonds are unsatisfied and the internal energy becomes

$$U_2 = -J(N-2) + 2J = -J(N-4) \; , \tag{4.6}$$

for all n. Since one can place the misaligned spins anywhere in the lattice, there are N equivalent configurations with this internal energy. The entropy of this state is then

$$S_2 = \ln(2N) \; . \tag{4.7}$$

The factor of 2 inside the logarithm arises due to the fact that we consider a reversed domain in each one of the two ordered states. At temperature T the free-energy of a state with one reversed spin and two *domain walls* is

$$F_2 = U_2 - TS_2 = -J(N-4) - T\ln(2N) \ . \tag{4.8}$$

The variation in free-energy between the ordered state and the one with one domain is

$$\Delta F = F_2 - F_o = 4J - T\ln N \ . \tag{4.9}$$

Thus, even if the internal energy increases due to the presence of the domain wall, the increase in entropy is such that the free-energy of the state with a droplet in it is much more favorable at any finite temperature T. We conclude that spin flips are favorable and order is destroyed at any finite temperature. The ferromagnetic Ising chain does not have a finite temperature phase transition.

A similar argument in $d > 1$ suggests that one can have, as indeed happens, a finite temperature transition in these cases (see, *e.g.* [9]).

At low temperatures, the structure of *droplets*, meaning patches in which the spins point in the direction of the opposite state, have been studied in detail with numerical simulations. Their knowledge has been used to derive phenomenological theories, the *droplet model* for systems with quenched disorder. At criticality one observes ordered domains of the two equilibrium states at all length scales – with *fractal* properties – and these have also been studied in detail. Right above T_c finite patches of the system are indeed ordered but these do not include a finite fraction of the spins in the sample and the magnetic density vanishes. However, these patches are enough to generate non-trivial thermodynamic properties very close to T_c and the richness of the critical phenomena.

4.4 Stiffness

A fundamental difference between an ordered and a disordered phase is their stiffness (or rigidity). In an ordered phase the free-energy cost of changing one part of the system with respect to the other part far away is of the order k_BT and usually diverges as a power law of the system size. In a disordered phase the information about the reversed part propagates only a finite distance (of the order of the correlation length, see below) and the stiffness vanishes.

The calculation of the stiffness is usually done as follows. Antiparallel configurations (or more generally the two ground states) are imposed at the opposite boundaries of the sample. A domain wall is then generated somewhere in the bulk. Its energy (or free-energy) cost, *i.e.* the difference between the modified configuration and the equilibrium one, is then measured.

4.5 Classification

Phase transitions are commonly classified by their *order*. The more common ones are those of first and second order.

In first order phase transition the order parameter jumps at the critical point from a vanishing value in the disordered side to a finite value right on the ordered side of the critical point. This is accompanied by discontinuities in various thermodynamic quantities and it is related to the fact that a first derivative of the free-energy density diverges. In such a transition the high and low temperature phases coexist at the critical point. Well-known examples are the melting of three dimensional solids and the condensation of a gas into a liquid. These transitions often exhibit hysteresis or memory effects since the system can remain in the metastable phase when the external parameters go beyond the critical point.

In second order phase transition the order parameter is continuous at the transition, *i.e.* it smoothly departs from zero at the critical point, but its variation with respect to the conjugate field in the zero field limit, or linear susceptibility, diverges. This is a second derivative of the free-energy density. At the critical point there is no phase coexitence, the system is in one critical state; the two phases on either side of the transition become identical at the critical point.

Higher order phase transitions appear when higher derivatives of the free-energy density diverge, *i.e.* when $\partial^n f/\partial y^n$ is discontinuous the transition is of n-th order, where y is any argument of f.

In disordered systems (see Sect. 5) a mixed case occurs in which the order parameter is discontinuous at the transition but all first derivatives of the free-energy density are finite. This is called a *random first order* transition and it provides a scenario for the glassy arrest [26].

4.6 Mean-field theory

In spite of their apparent simplicity, the statics of ferromagnetic Ising models has been solved analytically only in one and two dimensions. The mean-field approximation allows one to solve the Ising model in *any* spatial dimensionality. Even if the qualitative results obtained are correct, the quantitative comparison to experimental and numerical data shows that the approximation fails below an *upper critical dimension* d_u. It is however very instructive to see the mean-field approximation at work.

4.6.1 The naive mean-field approximation

The naive mean-field approximation consists in assuming that the probability density of the system's spin configuration is factorizable in independent factors

$$P(\{s_i\}) = \prod_{i=1}^{N} P_i(s_i) \qquad \text{with} \qquad P_i(s_i) = \frac{1+m_i}{2}\delta_{s_i,1} + \frac{1-m_i}{2}\delta_{s_i,-1} \tag{4.10}$$

and $m_i = \langle\, s_i \,\rangle$, where the thermal average has to be interpreted in the restricted sense described in the previous sections, *i.e.* taken over one ergodic component, in a way that $m_i \neq 0$. Note that one introduces an order-parameter dependence in the probabilities. Using this assumption one can compute the total free-energy

$$F = U - TS \tag{4.11}$$

where the average is taken with the factorized probability distribution (4.10) and the entropy S is given by

$$S = -\sum_{\{s_i\}} P(\{s_i\}) \ln P(\{s_i\}) \; . \tag{4.12}$$

One can use this approximation to treat finite dimensional models. Applied to the d-dimensional pure ferromagnetic Ising model with nearest-neighbor interactions on a cubic lattice $J_{ij} = J/2$ for nearest-neighbors and zero otherwise. One finds the internal energy

$$U = -\frac{J}{2} \sum_{\langle ij \rangle} \langle s_i s_j \rangle - h \sum_i \langle s_i \rangle = -\frac{J}{2} \sum_{\langle ij \rangle} m_i m_j - h \sum_i m_i \ , \tag{4.13}$$

and the entropy

$$\begin{aligned} S &= -\sum_{s_i=\pm 1} \prod_{k=1}^{N} P_k(s_k) \ln \prod_{l=1}^{N} P_l(s_l) \\ &= -\sum_{l=1}^{N} \sum_{s_l=\pm 1} P_l(s_l) \ln P_l(s_l) \\ &= -\sum_i \frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \ . \end{aligned} \tag{4.14}$$

For a uniformly applied magnetic field, all local magnetization equal the total density one, $m_i = m$, and one has the 'order-parameter dependent' free-energy density:

$$f(m) = -dJm^2 - hm + T \left[\frac{1+m}{2} \ln \frac{1+m}{2} + \frac{1-m}{2} \ln \frac{1-m}{2} \right] \ . \tag{4.15}$$

The extrema, $df(m)/dm = 0$, are given by

$$m = \tanh\left(\beta 2dJm + \beta h\right) \ , \tag{4.16}$$

with the factor $2d$ coming from the connectivity of the cubic lattice. This equation predicts a second order phase transition at $T_c = 2dJ$ when $h = 0$. This result is qualitatively correct in the sense that T_c increases with increasing d but the actual value is incorrect at all finite dimensions. In particular, this treatment predicts a finite T_c in $d = 1$ which is clearly wrong. The critical behavior is also incorrect in all finite d, with exponents (see Sect. 4.7) that do not depend on dimensionality and take the mean-field values. Still, the nature of the *qualitative* paramagnetic-ferromagnetic transition in $d > 1$ is correctly captured. We postpone the study of the solutions to eq. (4.16) to the next Subsection where we shall treat a similar, and in some respects, more general case. Having an expression for the free-energy density as a function of the order parameter, that is determined by eq. (4.16), one an compute all observables and, in particular, their critical behavior. We shall discuss it below.

4.6.2 The fully-connected Ising ferromagnet

Let us now solve exactly the fully-connected Ising ferromagnet with interactions between all p uplets of spins in an external field:

$$H = -\sum_{i_1 \neq \ldots \neq i_p} J_{i_1 \ldots i_p} s_{i_1} \ldots s_{i_p} - \sum_i h_i s_i \ , \tag{4.17}$$

$s_i = \pm 1$, $i = 1, \dots, N$. For the sake of generality – and to include the disordered models to be discussed in Sect. 5 – we use a generic interaction strength $J_{i_1 \dots i_p}$. The ferromagnetic model corresponds to

$$J_{i_1 \dots i_p} = \frac{J}{p! N^{p-1}} \tag{4.18}$$

with $0 < J = O(1)$, *i.e.* finite, and p is a fixed integer parameter, $p = 2$ or $p = 3$ or ..., that defines the model. The normalization with N^{p-1} of the first term ensures an extensive energy in the ferromagnetic state at low temperatures, and thus a sensible thermodynamic limit. The factor $p!$ is chosen for later convenience. This model is a source of inspiration for more elaborate ones with dilution and/or disorder (see Sect. 5). Using the factorization of the joint probability density that defines the mean-field approximation, one finds

$$\begin{aligned} F(\{m_i\}) &= -\sum_{i_1 \neq \dots \neq i_p} J_{i_1 \dots i_p} m_{i_1} \dots m_{i_p} - \sum_i h_i m_i \\ &\quad + T \sum_{i=1}^{N} \left[\frac{1+m_i}{2} \ln \frac{1+m_i}{2} + \frac{1-m_i}{2} \ln \frac{1-m_i}{2} \right] . \end{aligned} \tag{4.19}$$

The local magnetizations, m_i, are then determined by requiring that they minimize the free-energy density, $\partial f(\{m_j\})/\partial m_i = 0$ and a positive definite Hessian, $\partial^2 f(\{m_j\})/\partial m_i \partial m_j$ (*i.e.* with all eigenvalues being positive at the extremal value). This yields

$$m_i = \tanh \left(p\beta \sum_{i_2 \neq \dots \neq i_p} J_{i i_2 \dots i_p} m_{i_2} \dots m_{i_p} + \beta h_i \right) \tag{4.20}$$

If $J_{i_1 \dots i_p} = J/(p! N^{p-1})$ for all p uplets and the applied field is uniform, $h_i = h$, one can take $m_i = m$ for all i and these expressions become (4.22) and (4.24) below, respectively. The mean-field approximation is exact for the fully-connected pure Ising ferromagnet. [Note that the fully-connected limit of the model with pair interactions ($p = 2$) is correctly attained by taking $J \to J/N$ and $2d \to N$ in (4.16) leading to $T_c = J$.]

Let us solve the ferromagnetic model exactly. The sum over spin configurations in the partition function can be traded for a sum over the a variable, $x = N^{-1} \sum_{i=1}^{N} s_i$, that takes values $x = -1, -1 + 2/N, -1 + 4/N, \dots, 1 - 4/N, 1 - 2/N, 1$. Neglecting subdominant terms in N, one then writes

$$Z = \sum_x e^{-N\beta f(x)} \tag{4.21}$$

with the order-parameter dependent free-energy density

$$f(x) = -\frac{J}{p!} x^p - hx + T \left[\frac{1+x}{2} \ln \frac{1+x}{2} + \frac{1-x}{2} \ln \frac{1-x}{2} \right] . \tag{4.22}$$

The first two terms are the energetic contribution while the third one is of entropic origin since $N!/(N(1+x)/2)!(N(1-x)/2)!$ spin configurations have the same magnetization density.

In the large N limit, the partition function can be evaluated using the saddle-point method (see Appendix 3)

$$Z \approx \sum_\alpha e^{-N\beta f(m^\alpha_{sp})} \quad , \tag{4.23}$$

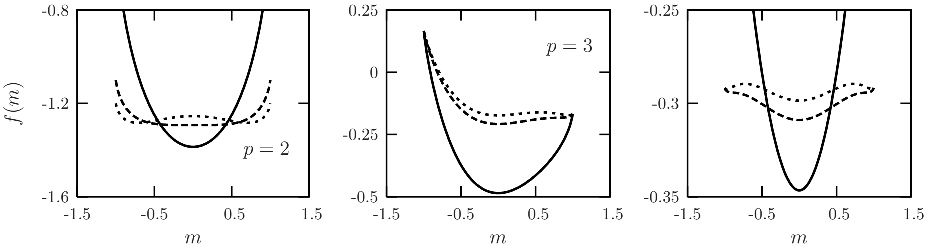


Figure 3: The free-energy density $f(m)$ of the $p=2$ (left), $p=3$ (center) and $p=4$ (right) models at three values of the temperature $T<T_c$ (light dashed line), $T=T_c$ (dark dashed line) and $T>T_c$ (solid line) and with no applied field.

where m_{sp}^{α} are the absolute minima of $f(x)$ given by the solutions to $\partial f(x)/\partial x|_{m_{sp}}=0$,

$$m_{sp} = \tanh\left(\frac{\beta J}{(p-1)!}\, m_{sp}^{p-1} + \beta h\right) \ , \tag{4.24}$$

together with the conditions $d^2 f(x)/dx^2|_{m_{sp}^{\alpha}} > 0$.

High temperature

In a finite magnetic field, eqn. (4.24) has a unique positive – negative – solution for positive – negative – h at all temperatures. The model is ferromagnetic at all temperatures and there is no phase transition in this parameter.

2nd order transition for $p=2$

In the absence of a magnetic field this model has a paramagnetic-ferromagnetic phase transition at a finite T_c. The order of the phase transition depends on the value of p. This can be seen from the temperature dependence of the free-energy density (4.22). Figure 3 displays $f(x)$ in the absence of a magnetic field at three values of T for the $p=2$ (left), $p=3$ (center) and $p=4$ (right) models (we call the independent variable m since the stationary points of $f(x)$ are located at the magnetization density of the equilibrium and metastable states, see below). At high temperature the unique minimum is $m=0$ in all cases. For $p=2$, when one reaches T_c, the $m=0$ minimum splits in two that slowly separate and move towards higher values of $|m|$ when T decreases until reaching $|m|=1$ at $T=0$ (see Fig. 3-left). The transition occurs at $T_c=J$ as can be easily seen from a graphical solution to eqn. (4.24), see Fig. 4-left. Close but below T_c, the magnetization increases as $m \sim (T_c-T)^{\frac{1}{2}}$. The linear magnetic susceptibility has the usual Curie behavior at very high temperature, $\chi \approx \beta$, and it diverges as $\chi \sim |T-T_c|^{-1}$ on both sides of the critical point. The order parameter is continuous at T_c and the transition is of second-order thermodynamically.

1st order transition for $p>2$

For $p>2$ the situation changes. For even values of p, at T^* two minima (and two maxima) at $|m| \neq 0$ appear. These coexist as metastable states with the stable minimum

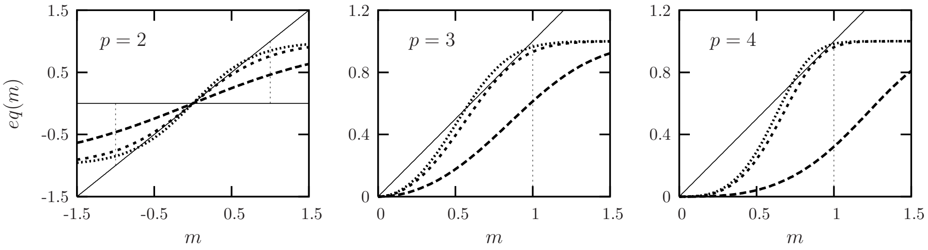


Figure 4: Graphical solution to the equation fixing the order parameter x for $p = 2$ (left), $p = 3$ (center) and $p = 4$ (right) ferromagnetic models at three values of the temperature $T < T^*$, $T = T^*$ and $T > T^*$ and with no applied field. Note that the rhs of this equation is antisymmetric with respect to $m \to -m$ for odd values of p while it is symmetric under the same transformation for even values of p. We show the positive quadrant only to enlarge the figure. T^* is the temperature at which a second minimum appears in the cases $p = 3$ and $p = 4$.

at $m = 0$ until a temperature T_c at which the three free-energy densities coincide, see Fig. 3-right. Below T_c the $m = 0$ minimum continues to exist but the $|m| \neq 0$ ones are favored since they have a lower free-energy density. For odd values of p the free-energy density is not symmetric with respect to $m = 0$. A single minimum at $m^* > 0$ appears at T^* and at T_c it reaches the free-energy density of the paramagnetic one, $f(m^*) = f(0)$, see Fig. 3-center. Below T_c the equilibrium state is the ferromagnetic minimum. For all $p > 2$ the order parameter is discontinuous at T_c, it jumps from zero at T_c^+ to a finite value at T_c^-. The linear magnetic susceptibility also jumps at T_c. While it diverges as $(T - T_c)^{-1}$ on the paramagnetic side, it takes a finite value given by eqn. (4.26) evaluated at m^* on the ferromagnetic one. In consequence, the transition is of first-order.

Pinning field, broken ergodicity and spontaneous broken symmetry

The saddle-point equation (4.24) for $p = 2$ [or the mean-field equation (4.16)] admit two equivalent solutions in no field. What do they correspond to? They are the magnetization density of the equilibrium ferromagnetic states with positive and negative value. Indeed, if we computed the averaged magnetization density with the partition sum restricted to the configurations with positive (or negative) value of x we find $m = m_{sp}$. For $T < T_c$ if one computed this result from $m = N^{-1} \sum_{i=1}^N \langle s_i \rangle = \sum_x e^{-\beta N f(x)} x$ summing over the two minima of the free-energy density one finds $m = 0$ as expected by symmetry.

In practice, the restricted sum is performed by applying a small magnetic field, computing the statistical properties in the $N \to \infty$ limit, and then setting the field to zero. In other words,

$$m_\pm \equiv \frac{1}{N} \sum_{i=1}^N \langle s_i \rangle_\pm = \left(\frac{1}{\beta N} \frac{\partial \ln Z}{\partial h} \right) \Bigg|_{h \to 0^\pm} = \frac{\partial f(m_{sp})}{\partial h} \Bigg|_{h \to 0^\pm} = \pm |m_{sp}| \; . \tag{4.25}$$

The limit $N \to \infty$ taken in a field selects the positive (or negatively) magnetized states.

The magnetic linear susceptibility is given by

$$\chi \equiv \left.\frac{\partial m}{\partial h}\right|_{h\to 0^\pm} = \left.\frac{\partial m_{sp}}{\partial h}\right|_{h\to 0^\pm} = \frac{\beta}{\cosh^2(\frac{\beta J}{(p-1)!}\, m_{sp}^{p-1}) - \beta J} \ . \tag{4.26}$$

For even p the two magnetized states have the same divergent susceptibility at T_c.

For odd values of p there is only one non-degerate minimum of the free-energy density at low temperatures and the application of a pinning field is then superfluous.

The existence of two degenerate minima of the free-energy density, that correspond to the two equilibrium ferromagnetic states at low temperatures, implies that ergodicity is broken in these models. In pure static terms this means that one can separate the sum over all spin configurations into independent sums over different sectors of phase space that correspond to each equilibrium state. In dynamic terms it means that temporal and statistical averages (taken over all configurations) do not coincide.

For any even value of p and at all temperatures the free-energy density in the absence of the field is symmetric with respect to $m \to -m$, see the left and right panels in Fig. 3. The phase transition corresponds to a *spontaneous symmetry breaking* between the states of positive and negative magnetization. One can determine the one that is chosen when going through T_c either by applying a small *pinning field* that is taken to zero only after the thermodynamic limit, or by imposing adequate boundary conditions. Once a system sets into one of the equilibrium states this is completely stable in the $N \to \infty$ limit.

For all odd values of p the phase transition is not associated to symmetry breaking, since there is only one non-degenerate minimum of the free-energy density that corresponds to the equilibrium state at low temperature.

4.6.3 Landau theory

The exercise that we solved in the last subsection corresponds to a fully solvable model for which mean-field theory is exact. Now, can one get an idea of the limit of validity of mean-field theory and when it is expected to fail?

Landau proposed an extension of Weiss mean-field theory for ferromagnets (Sect. 4.6) that has a much wider range of application, includes space, allows to predict when it applies and when it fails. In a few words – we shall not develop Landau theory here – it consists in proposing a field theory for coarse-grained fields that represent the averaged relevant variable. In the case of an Ising spin system, the field in each coarse-graining volume $v = \ell^d$ within the sample is defined as:

$$\phi(\vec{x}) \equiv \frac{1}{\ell^d} \sum_{j\in v_{\vec{x}}} s_j \ , \tag{4.27}$$

and the effective free-energy of the interacting system is proposed to be

$$F(\phi) = \int d^dx \left[c(\nabla\phi(\vec{x}))^2 + \frac{\lambda}{4!}\phi^4(\vec{x}) + \frac{T-T_c}{T_c}\phi^2(\vec{x}) \right] \ . \tag{4.28}$$

The first term mimics an elastic energy related to the ferromagnetic interactions. The second term is an expansion of the entropic contribution in powers of $T - T_c$ that is expected to be valid only close to T_c. One keeps the leading terms in the Taylor expansion.

Note that this 'order-parameter dependent' free-energy is not quadratic due to the term ϕ^4. If one then integrates over all ϕ configurations – to compute the partition function – and then evaluates this integral with a saddle-point approximation one can also include the fluctuations (see Appendix 3) and see when these ones become too important and thus make the saddle-point evaluation invalid. This analysis is called *Ginzburg criterion* and tells us that there is an upper critical dimension,

$$d_u = 4 \tag{4.29}$$

above which mean-field theory is *exact*! and below which it fails. However, it does not fail everywhere in parameter space. It just fails when one gets very close to the critical point. How close, it depends on the system, and this is called the *critical region.* The behavior of is then well-described by the Landau-Ginzburg theory away from the critical region. Inside the critical region the approach fails. In particular, it predicts the mean-field exponents in (4.34) for all d.

Landau and Ginzburg got Nobel Prizes in Physics that were strongly associated to their work along these lines.

4.7 Critical exponents

The rest of the discussion will focus on second order phase transitions for which the order parameter departs from zero smootly when entering the ordered phase.

When studying the observables close to the critical point one realizes that they depend on the distance from the critical point in the form of power laws. For instance, in zero field the order parameter increases as

$$m \sim (T_c - T)^{\beta} \tag{4.30}$$

while at T_c and as a function of the conjugate field it behaves as

$$m \sim h^{1/\delta} \ . \tag{4.31}$$

The divergence of the linear susceptibility at T_c is characterized by two exponents

$$\chi \sim \left\{ \begin{array}{ll} (T-T_c)^{-\gamma} & T > T_c \ , \\ (T_c - T)^{-\gamma'} & T < T_c \ . \end{array} \right. \tag{4.32}$$

The specif heat also diverges at T_c:

$$C_V \sim \left\{ \begin{array}{ll} (T-T_c)^{-\alpha} & T > T_c \ , \\ (T_c - T)^{-\alpha'} & T < T_c \ . \end{array} \right. \tag{4.33}$$

While the values of T_c are material dependent, all ferromagnetic transitions of systems in $d = 3$ that with an order parameter of the same dimensionality can be described by the same – within error bars – critical exponent! This feature indicates the existence of *universality classes*, *i.e.* groups of systems for which the details of the microscopic interactions do not matter and whose macroscopic critical behavior is identical.

It is simple to compute the exponents in the naive mean-field approximation for any d or for the fully connected model with $p = 2$. They read

$$\alpha = 0 \;, \qquad \beta = \frac{1}{2} \;, \qquad \gamma = 1 \;, \qquad \delta = 3 \;, \qquad \eta = 0 \;, \qquad \nu = \frac{1}{2} \;, \tag{4.34}$$

and they are independent of d. These values are to be confronted to the experimental values. In ferromagnetic phase transition of Ising symmetry they are

d	β	α	γ	δ	ν	η	
2	1/8	0	7/4	15	1	1/4	exact
3	0.325	0.11	1.24	4.82	0.63	0.032	approx

Table 2: Critical exponents in the Ising universality class.

4.8 Towards an understanding of critical phenomena

The fact that very different systems share the same critical properties, the mere existence of universality classes, suggested that a very general framework should be able to describe the critical behavior of all these systems at once. The fact that the mean-field critical exponents were slightly different from the observed ones was not very important as a quantitative disagreement but it was from a fundamental point of view. Something important was going on and needed an explanation.

In this Subsection we introduce and discuss the concepts that allowed a qualitative and quantitative understanding of critical phenomena. The ideas and methods introduced actually go beyond this problem and have been exported to other situations like dynamical processes in and out of equilibrium (see Sects. 6 and 7).

4.8.1 Scaling

Scaling concepts are fundamental in describing the behaviour of systems made of large number of constituents, interacting non-linearly and according to laws that are sometimes poorly understood. The idea is to isolate a few relevant variables that characterize the behaviour at a certain length and time scale and to postulate simple *scaling relations* between them. When there is only one independent variable, the scaling relations take the form of power laws with exponents that are not rational numbers.

Systems that may be microscopically very different but share the same scaling relations belong to the same *universality class.*

Scaling arguments apply to many different physical situations (in and out of equilibrium) and they can be explained using *renormalization* ideas. In most cases, the renormalization approach does not have a formal basis yet. It is in the context of critical phenomena in equilibrium that scaling and renormalization can be derived systematically.

In the discussion of critical phenomena we have defined 6 critical exponents (α for the specific heat, β for the order parameter, γ for the susceptibility, δ for the order parameter at the critical point as a function of the conjugate field, η for the correlation function and ν for the correlation length). But, actually, not all these exponents are independent, the scaling hypothesis allows one to show that only two of them are. One example of these relations is the one called *Rushbrooke scaling law*

$$\alpha + 2\beta + \gamma = 2 \; . \tag{4.35}$$

These relations follow from the fact that all observables can be computed from the free-energy density and, if one assumes a *scaling form* for it, one is forced to have relations between the exponents characterizing the thermodynamic observables and correlation functions. We shall not discuss these relations here, the interested student can find them in any book on critical phenomena, see *e.g.* [9].

Let us take a simple viewpoint and discuss a way to collapse data close to a critical point, a property closely related to scaling. The power law expressions (4.30) and (4.31) suggest Widom scaling for the order parameter:

$$m(t,h) \sim |t|^\beta \Phi_\pm \left(\frac{h}{|t|^{\beta\delta}} \right) \qquad t \equiv \frac{|T-T_c|}{T_c} \; , \tag{4.36}$$

with $\Phi_\pm(0) = 1$ and $\Phi_\pm(x \to \infty) \sim x^{1/\delta}$. With these limits, (4.31) is recovered on the critical isotherm and (4.30) follows at strictly zero field and for $|t| \ll 1$. An example of data collapse is given in Fig. 5.

Surprisingly enough, *all* systems undergoing a ferromagnetic transition can be *scaled* in this way using the *same* functions $\Phi_\pm$ above and below the critical temperature, respectively! The way of checking this hypothesis is by plotting $m/|t|^\beta$ against $|h|/|t|^\beta\delta$ for different systems and looking for *data collapse.* Of course, we do not know the values of the universal exponents β and δ and the material dependent critical temperature T_c *a priori*, so we need to manipulate a bit the data before obtaining collapse. Note that the scaling law (4.36) is independent of the dimension d.

One can also write the scaling hypothesis for the free-energy density and derive the one for the magnetization by differentiation.

Scaling relations that involve the dimension are called *hyperscaling.* The correlation function satisfies

$$G(\vec{r}; t, h) = \frac{1}{r^{d-2+\eta}} \, g\left(r|t|^\nu, \frac{h}{|t|^{\beta\delta}} \right) \tag{4.37}$$

Kadanoff proposed that this quite incredible feature could be explained assuming that near a critical point a system looks the same at all length scales. This is called *scale invariance.*

4.8.2 The correlation length

A very important concept in critical phenomena is that of a *correlation length* usually denoted by ξ.

The correlation length is the distnce over which the fluctuations of the microscopic degrees of freedom are significantly correlated. A simple way to understand its meaning is the following. Take a macroscopic sample and measure some macroscopic observable under some external conditions, *i.e.* temperature T and pressure P. Now, repeat the measurement after cutting the sample in two pieces and keeping the external conditions

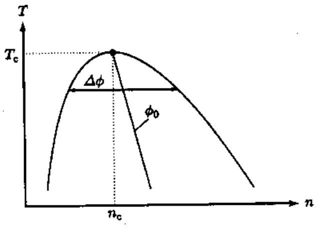


Fig. 4.4.3. Schematic phase boundary for the liquid-gas transition showing the asymmetry of the order parameter about the critical density.

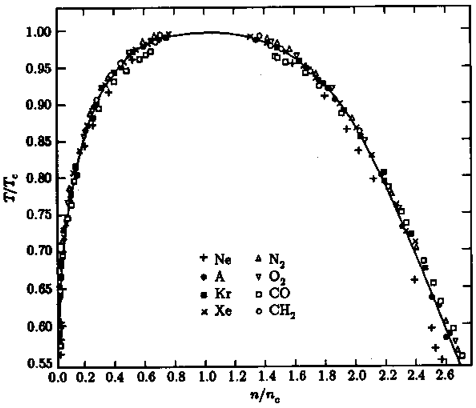


Fig. 4.4.4. Phase boundary in units of reduced temperature and density for eight different molecular fluids near their liquid-gas transitions. Note the universal behavior and the fact that the solid line is $\Delta\phi \propto (T_c - T)^\beta$ with $\beta = 1/3$ rather than the mean-field result $\beta = 1/2$. [E.A. Guggenheim, *J. Chem. Phys.* 13, 253 (1945).]

Chaikin-Lubensky より

Figure 5: Critical scaling in gas-liquid transitions at constant pressure. At very low density and low temperature, at the left of the curve the system is a gas, at very large density and still low temperature, at the right of the curve the system is a liquid. In the region below the curve there is coexistence of gas and liquid. Above the curve the system goes continuously from a gas to a liquid when increasing the density. The critical line behaves as $|\rho_l - \rho_g| \sim |T - T_c|^\beta$ with $\beta \sim 0.327$ close to the maximum. Note that scaling holds as far as $T/T_c \sim 0.55$!

unchanged. The result for the macroscopic observable is the same. Repeating this procedure, one finds the same result until the system size reaches the correlation length of the material.

At finite temperature, one can have *droplets* of the wrong phase within the correct one, due to thermal agitation. The size of these droplets will be a function of temperature and at a given instant, a snapshot of the system reveals the existence of a number of them with different sizes. One expects though that they have a well-defined average (taken, for instance, over different snapshots taken at different times). This average size can be taken as a qualitative indication of the value of the correlation length (we shall give a more precise definition below).

In first order phase transition the correlation length is finite for all values of the parameters. In second order phase transitions, the correlation length is usually very short, of the order of a few lattice spacing, at low temperature. It increases when approaching T_c, it diverges at T_c, and then decreases again in the high temperature phase when getting away from the critical point. See Fig. 6.

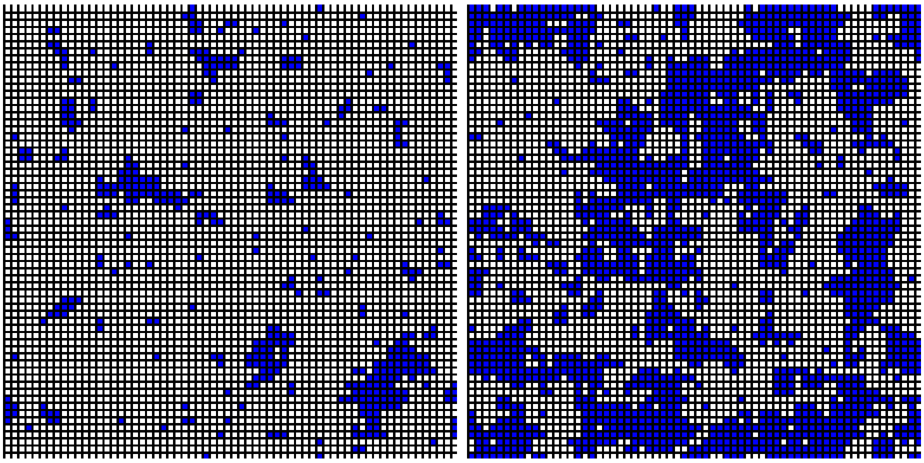


Figure 6: Two snapshots of the spin configuration in a $2d$ Ising model. Left: below T_c; right: at T_c.

The fact that one finds coherent structures at all lengths at the critical point means that there is no spatial scale left in the problem and then all scales participate in the critical behaviour. The system becomes *scale invariant* and, if one looks at it with different microscopes one essentially sees the same.

The actual definition of the correlation length is based on the use of the *static suceptibility sum rule*. Let us define the two-point connected correlation function

$$G(\vec{r}_i, \vec{r}_j) \equiv \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle \tag{4.38}$$

where $\vec{r}_i$ is the spatial position of the spin s_i. A simple calculation allows one to show that the linear susceptibility reads

$$\chi \equiv \frac{\partial m_h}{\partial h} = - \left. \frac{\partial^2 f}{\partial h^2} \right|_{h=0} = \frac{\beta}{N} \sum_{ij} G(\vec{r}_i - \vec{r}_j) \; . \tag{4.39}$$

Let us prove this statement. Nothing indicates that spatial translational invariance should be violated, thus, the correlation should be a function of the distance between the points $\vec{r}_i$ and $\vec{r}_j$ only:

$$G(\vec{r}_i, \vec{r}_j) = G(\vec{r}_i - \vec{r}_j) \; . \tag{4.40}$$

One then has

$$\chi \equiv \frac{\partial m}{\partial h} = \beta \sum_i G(\vec{r}_i) = \frac{\beta}{a^d} \int_V d^d r \; G(\vec{r}) \; . \tag{4.41}$$

This means that the divergence of the susceptibility at the critical point must be accompanied by a special behavior of the correlation function. Indeed, one finds that

$$G(\vec{r}) \sim \frac{e^{-r/\xi}}{r^{d-1} \xi^{(d-3)/2}} \; , \qquad\qquad \text{for } \; r \gg \xi \; . \tag{4.42}$$

This expression is integrable over the full volume unless the exponential factor disappears. This is indeed what happens at T_c where the correlation length diverges, again as a power law of the distance to the critical point

$$\xi \sim |T - T_c|^{-\nu} \; . \tag{4.43}$$

Finally, one observes that the correlation function right at the critical point also diverges as a power law:

$$G(\vec{r}) \sim r^{-(d-2+\eta)} \; , \tag{4.44}$$

with η another critical exponent.

Let us discuss the correlation length in a simple solvable case, the Ising model in $d = 1$ with, say, open boundary conditions. In this case, the finite temperature correlation function is

$$G_{kl} \;\; = \;\; \langle s_k s_l \rangle - \langle s_k \rangle \langle s_l \rangle = \langle s_k s_l \rangle \tag{4.45}$$

since $\langle s_k \rangle = 0$ at any $T > 0$. Introducing, for convenience, different coupling constants $K_i = \beta J_i$ on the links, G_{kl} reads

$$G_{kl} = Z^{-1} \sum_{\{s_i = \pm 1\}} e^{\sum_i K_i s_i s_{i+1}} s_k s_l = Z^{-1} \frac{\partial}{\partial K_k} \frac{\partial}{\partial K_{k+1}} \dots \frac{\partial}{\partial K_{l-1}} Z \; . \tag{4.46}$$

At the end of the calculation one takes $K_i = K = \beta J$ for all i. Thus, at finite temperature the connected correlation between any two spins can be computed as a number of derivatives (depending on the distance between the spins) of the partition function conveniently normalized. Using the change of variables $\eta_i = s_i s_{i+1}$, one finds

$$Z = \sum_{\{\eta_i = 1\}} e^{\sum_i K_i \eta_i} = 2 \prod_{i=1}^{N-1} 2\cosh(K_i) \to 2(2\cosh \beta J)^{N-1} \; . \tag{4.47}$$

Taking the distance between the chosen spins s_k and s_l to be $k - l = r$ the correlation function is then given by

$$G(r) \;\; = \;\; [\tanh(\beta J)]^r = e^{r \ln[\tanh(\beta J)]} = e^{-r/\xi} \tag{4.48}$$

with

$$\xi = \frac{1}{\ln \coth(\beta J)} \sim e^{4J/(T-T_c)} \; , \qquad T \sim 0 \; . \tag{4.49}$$

In this one dimensional example we found a essential singularity, an exponential divergence, of the correlation length when approaching $T_c = 0$. In general, in higher d, one has a power law divergence of the form (4.44).

4.8.3 Finite size effects

A real system is large but finite, $1 \ll N_A \sim 10^{23} < \infty$. Finite size effects will then play a role in the phase transition that is rounded by the fact that $N_A < \infty$. Finite size effects become important when $\xi \sim L$, the linear size of the system, say $L \sim 1$cm for an actual sample. A rough estimation of how close to T_c one needs to get to see deviations from critical scaling shows that finite size effects are quite negligible in experiments but are certainly not in numerical simulations and have to be taken into account very carefully when trying to compare numerical data to analytical predictions.

Finite size effects are taken into account by introducing correcting functions in the scaling laws, for example,

$$\chi_L \sim |t|^{-\gamma} g\left(\frac{L}{\xi}\right) \tag{4.50}$$

with $g(x \to \infty) \to 1$ and $g(x \to 0) \to x^{\gamma/\nu}$, and

$$m \sim L^{\beta/\nu} \; . \tag{4.51}$$

4.8.4 The renormalization group

The development of the renormalization group by K. Wilson in the early 70s gave a totally new way of understanding condensed-matter and particle physics phenomena. He transformed the picture of phase transitions that developed in the 60s – with the understanding of concepts like scaling, universality and correlations – into a calculational tool and got the Nobel Prize in Physics in 1982.

Let us describe how the RG works on one example, again the one dimensional Ising model with N (even) spins in the absence of an applied field. The partition function reads

$$Z(N,J) = \sum_{s_i=\pm 1} e^{\beta J \sum_{i=1}^{N-1} s_i s_{i+1}} \tag{4.52}$$

We shall call $K \equiv \beta J$. The sum in the exponential is

$$s_1 s_2 + s_2 s_3 + s_3 s_4 + s_4 s_5 + \dots \tag{4.53}$$

It is clear that s_2 enters only in the first two terms, s_4 enters only on the third and forth term and so on and so forth. One can then sum over all configurations of the spins with an even label. For example, the sum over $s_2 = \pm 1$ yields

$$e^{K(s_1+s_3)} + e^{-K(s_1+s_3)} \; . \tag{4.54}$$

One then obtains:

$$Z(N,K) = \sum_{s_i=\pm 1; i \; odd} \prod_{i=1}^{N-1} [e^{K(s_i+s_{i+2})} + e^{-K(s_i+s_{i+2})}] \tag{4.55}$$

If we find a function $\kappa(K)$ and a new coupling constant K' such that each of these terms can be written as

$$e^{K(s_i+s_{i+2})} + e^{-K(s_i+s_{i+2})} = \kappa(K) e^{K' s_i s_{i+2}} \; , \tag{4.56}$$

the right-hand-side in eq. (4.55) would be proportional to the partition function of another one dimensional Ising model with $N/2$ spins and a different coupling constant K':

$$Z(N,K) = \kappa(K)^{N/2} Z(N/2, K') \; . \tag{4.57}$$

Indeed, the solution to (4.56) can be easily found; it is enough to consider all the cases $s_i = \pm 1$ and $s_{i+2} = \pm 1$ to obtain:

$$\begin{aligned} K' &= \frac{1}{2} \ln \cosh(2K) \;, && (4.58) \\ \kappa(K) &= 2 \cosh^{1/2}(2K) \;. && (4.59) \end{aligned}$$

Now, we know that the free-energy, and in particular $\ln Z$, should be linear in N;

$$-\beta F(N, K) = \ln Z(N, K) = N\zeta(K) \;. \quad (4.60)$$

Thus, taking the ln of eq. (4.62) and using (4.60)

$$\begin{aligned} \ln Z(N, K) &= \frac{N}{2} \ln \kappa(K) + \ln Z(N/2, K') \;, && (4.61) \\ \zeta(K) &= \frac{1}{2} \ln \kappa(K) + \frac{1}{2} \zeta(K') \;. && (4.62) \end{aligned}$$

or equivalently

$$\zeta(K') = -\ln[2 \cosh^{1/2}(2K) + 2\zeta(K)] \quad (4.63)$$

Inserting a value of K on the right-hand-side one obtains the new coupling constant K' and the new free-energy $\zeta(K')$.

Equations (4.58) and (4.63) provide recursion relations for the coupling constant and the partition function. Note that K' is always smaller than K. One can also solve for K as a function of K':

$$\begin{aligned} K &= \frac{1}{2} \cosh^{-1}(e^{2K'}) && (4.64) \\ \zeta(K) &= \frac{1}{2} \ln 2 + \frac{1}{2} K' + \frac{1}{2} \zeta(K') \;, && (4.65) \end{aligned}$$

obtaining now an increasing flow $K(K')$.

One can use these results to compute the value of the partition function for any K. The argument goes as follows. For very small K', *i.e.* very high temperature, the spins are basically independent and $Z(K') \sim 2^N$ and $\zeta(K') = \ln 2$. Using then (4.64) and (4.65) one computes K and $\zeta(K)$. One then iterates using these values as starting points K' and $\zeta(K')$. The agreement between the values generated this way and the results of the exact calculation are quite amazing (see, *e.g* [13]).

The process described in the previous paragraph is a *flow* in the space of parameters. Starting from any non-zero value of the coupling constant the iteration converges to $K \to \infty$. There are then two *fixed points*, a *stable fixed point* at $K \to \infty$ (to which trajectories are attracted) and an *unstable fixed point* $K = 0$ from which trajectories depart. The stable fixed point is the critical point one is looking for. The critical behavior can then be obtained from the dependence of $\zeta(K)$ on the parameter $K - K_c$.

A similar procedure can be applied to the $d = 2$ problem. In this case, however, the *decimation* of spins cannot be done exactly, as in $d = 1$, and one is forced to use some approximation. There a number of successful recipes in the literature. Basically, one finds that the flow in parameter space approaches a stable fixed point at finite value of K which corresponds to the phase transition.

4.8.5 Fluctuations of macroscopic observables

A direct consequence of having a diverging correlations length is that the critical measure-to-measure fluctuations global observables like, for instance, the magnetization density, are not Gaussian. The reason is simple, if $L < \xi$, a global measurement is not the result of an average over many uncorrelated regions and, thus, one can not use the central limit theorem to argue for a normal distribution of fluctuations. Recently, the study of critical fluctuations of macroscopic observables received much attention. The best adapted model for this analysis is the $2d$ XY model, that is critical on a finite interval of temperatures (and not only at a single precise value of T_c) [27]. This model, in an approximation, is mapped onto an interface model, the Edwards-Wilkinson one, that we shall discuss in Sect. 7.

In the limit we are now interested in L is finite with respect to ξ and finite size effects are important. *Finite size scaling* implies

$$p_L(M) = L^{\beta/\nu}\, \Pi\left(ML^{\beta/\nu}, \frac{\xi}{L}\right) . \tag{4.66}$$

The calculation of this probability function in the so-called spin-wave approximation of the $2d$ XY model showed that it has the form of Gumbel distributions found in the staudy of extreme value statistics that we discussed in Sect. 3.1.2. The role payed by extreme We values in this context remains, however, not clear. shall give more details on the analysis of these non-Gaussian fluctuations in the language of interfaces.

5 Disordered systems

No material is perfectly homogeneous: impurities of different kinds are distributed randomly throughout the samples.

A natural effect of disorder should be to lower the critical temperature. Much attention has been payed to the effect of *weak* disorder on phase transitions, that is to say, situations in which the nature of the ordered and disordered phases is not modified by the impurities but the critical phenomenon is. On the one hand, the critical exponents of second order phase transitions might be modified by disorder, on the other hand, disorder may smooth out the discontinuities of first order phase transitions rendering them of second order. *Strong* disorder instead changes the nature of the low-temperature phase and before discussing the critical phenomenon one needs to understand how to characterize the new ordered 'glassy' phase.

In this Chapter we shall discuss several types of *quenched* disorder and models that account for it. We shall also overview some of the theoretical methods used to deal with the static properties of models with quenched disorder.

5.1 Quenched and annealed disorder

First, one has to distinguish between *quenched* and *annealed* disorder. Imagine that one mixes some random impurities in a melt and then very slowly cools it down in such a way that the impurities and the host remain in thermal equilibrium. If one wants to study the statistical properties of the full system one then has to compute the full partition function in which one sums over all configurations of original components and impurities. This is called annealed disorder. In the opposite case in which upon cooling the host and impurities do not equilibrate but the impurities remain blocked in random fixed positions, one talks about quenched disorder. Basically, the relaxation time associated with the diffusion of the impurities in the sample is so long that these remain trapped. The former case is easier to treat analytically but is less physically relevant. The latter is the one that leads to new phenomena and ideas that we shall discuss next.

5.2 Bond disorder: the case of spin-glasses

Spin-glasses are alloys in which magnetic impurities substitute the original atoms in positions randomly selected during the chemical preparation of the sample. The interactions between the impurities are of RKKY type:

$$V_{rkky} = -J \, \frac{\cos(2k_F r_{ij})}{r_{ij}^3} \, s_i s_j \tag{5.1}$$

with $r_{ij} = |\vec{r}_i - \vec{r}_j|$ the distance between them and s_i a spin variable that represents their magnetic moment. Clearly, the initial location of the impurities varies from sample to sample. The time-scale for diffusion of the magnetic impurities is much longer than the time-scale for spin flips. Thus, for all practical purposes the positions $\vec{r}_i$ can be associated to quenched random variables distributed according to a uniform probability distribution that in turn implies a probability distribution of the exchanges. This is called *quenched disorder*.

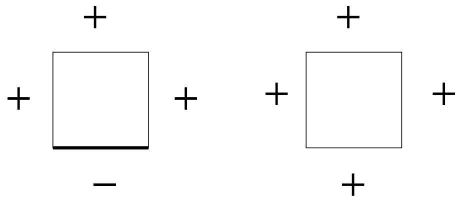


Figure 7: A frustrated (left) and an unfrustrated (center) square plaquette. A frustrated triangular plaquette (right).

5.2.1 Frustration

Depending on the value of the distance r_{ij} the numerator in eqn. (5.1) can be positive or negative implying that both ferromagnetic and antiferromagnetic interactions exist. This leads to *frustration*, which means that some two-body interactions cannot be satisfied by any spin configuration. An example with four sites and four links is shown in Fig. 7-left, where we took three positive exchanges and one negative one all, for simplicity, with the same absolute value, J. Four configurations minimize the energy, $E_f = -2J$, but none of them satisfies the lower link. One can easily check that any closed loop such that the product of the interactions takes a negative sign is frustrated. Frustration naturally leads to a higher energy and a large degeneracy of the number of ground states. This is again easy to grasp by comparing the number of ground states of the frustrated plaquette in Fig. 7-left to its unfrustrated counterpart shown on the central panel. Indeed, the energy and degeracy of the ground state of the unfrustrated plaquette is $E_u = -4J$ and $n_u = 2$, respectively.

Frustration may be due to pure geometrical constraints. The canonical example is an antiferromagnet on a triangular lattice in which each plaquette is frustrated, see Fig. 7.

5.2.2 Self-averageness

If each sample is characterized by its own realization of the exchanges, should one expect a totally different behavior from sample to sample? Fortunately, many generic static and dynamic properties of spin-glasses (and other systems with quenched disorder) do not depend on the specific realization of the random couplings and are *self-averaging.* This means that the typical value is equal to the average over the disorder:

$$A_J^{typ} = [\, A_J \,]_{av} \; . \tag{5.2}$$

In particular, the free-energy of models with short-ranged interactions is expected to be self-averaging.

The meaning of this property can grasped from the solution of the random bond Ising chain defined by the energy function $E = -\sum_i J_i s_i s_{i+1}$ with spin variables $s_i = \pm$, for $i = 1, \ldots, N$ and random bonds J_i independently taken from a probability distribution $P(J_i)$. For simplicity, we consider periodic boundary conditions. The disorder-dependent partition function reads

$$Z_J = \sum_{\{s_i = \pm 1\}} e^{\beta \sum_i J_i s_i s_{i+1}} \tag{5.3}$$

and this can be readily computed introducing the change of variables $\sigma_i \equiv s_i s_{i+1}$. One finds.

$$Z_J = \prod_i 2\cosh(\beta J_i) \qquad \Rightarrow \qquad -\beta F_J = \sum_i \ln\cosh(\beta J_i) + N \ln 2 \;. \tag{5.4}$$

The partition function is a *product* of *i.i.d.* random variables and it is itself a random variable with a log-normal distribution. The free-energy instead is *sum* of *i.i.d.* and, using the central limit theorem, it becomes a Gaussian random variable that is narrowly peaked about its maximum in the large N limit. The typical value, determined by the maximum in the distribution coincides with the average.

A simple argument that justifies the self-averageness of the free-energy in generic finite dimensional systems with short-range interactions is the following. Let us divide the system of volume V in n subsystems of volume v with $V = nv$. If the interactions are short-ranged, the total energy is the sum of two terms, a contribution from the bulk of the subsystems and a contribution from the interfaces between the subsystems. In the thermodynamic limit, the latter is negligible with respect to the former. The disorder dependent free-energy density is then a of random numbers, each one being the disorder dependent free-energy of the bulk of each subsystem. In the limit of a very large number of subsystems ($V \gg v$) the central limit theorem implies that the total free-energy is Gaussian distributed with the maximum reached at a value F_J^* that coincides with the average of all realizations of the randomness $[F_J]$. Thus, the typical F_J is identical to the averaged $[F_J]_{av}$. One can then compute the latter to understand the static properties of the typical system.

Once one has $[F_J]$, one derives all disordered average thermal averages by taking derivatives of the disordered averaged free-energy with respect to sources introduced in the partition function. For example,

$$[\,\langle\, s_i \,\rangle\,]_{av} = -\left.\frac{\partial [F_J]_{av}}{\partial h_i}\right|_{h_i=0} \;, \tag{5.5}$$

$$[\,\langle\, s_i s_j \,\rangle - \langle\, s_i \,\rangle\langle\, s_j \,\rangle\,]_{av} = T \left.\frac{\partial [F_J]_{av}}{\partial h_i h_j}\right|_{h_i=0} \;, \tag{5.6}$$

with $E \to E - \sum_i h_i s_i$.

5.3 Models

In early 70s *Edwards and Anderson* proposed a rather simple model that should capture the main features of spin-glasses. The interactions (5.1) decay with a cubic power of the distance and hence they are relatively short-ranged. This suggests to put the spins on a regular cubic lattice model and to trade the randomness in the positions into random nearest neighbor exchanges taken from a Gaussian (or bimodal) probability distribution:

$$E_{ea} = -\sum_{\langle ij \rangle} J_{ij} s_i s_j \qquad \text{with} \qquad P(J_{ij}) = (2\pi\sigma^2)^{-\frac{1}{2}}\, e^{-\frac{J_{ij}^2}{2\sigma^2}} \;. \tag{5.7}$$

A natural extension of the EA model in which all spins interact has been proposed by Sherrington and Kirkpatrick

$$E = -\sum_{i_1 \neq j} J_{ij} s_i s_j - \sum_i h_i s_i \tag{5.8}$$

and it is called the *SK model*. The interaction strengths J_{ij} are taken from a Gaussian pdf and they scale with N in such a way that the thermodynamic is non-trivial:

$$P(J_{ij}) = (2\pi\sigma_N^2)^{-\frac{1}{2}} \, e^{-\frac{J_{ij}^2}{2\sigma_N^2}} \qquad\qquad \sigma_N^2 = \sigma^2 N \; . \tag{5.9}$$

This is a case for which a mean-field approximation is expected to be exact.

A further extension of the EA model is called the *p spin model*

$$E = -\sum_{i_1 \neq \ldots \neq i_p} J_{i_1 \ldots i_p} s_{i_1} \ldots s_{i_p} - \sum_i h_i s_i \tag{5.10}$$

with $p \geq 3$. The exchanges are now taken from a Gaussian probability distribution with

$$[J_{i_1 \ldots i_p}]_{av} = 0 \; , \qquad [J^2_{i_1 \ldots i_p}]_{av} = \frac{J^2}{2N^{p-1}} \; . \tag{5.11}$$

Indeed, an interesting thermodynamic limit is achieved by scaling $J_{i_1 \ldots i_p}$ with $N^{-(p-1)/2}$. This scaling is justified as follows. The "force" $F_i \equiv \sum_{ii_2 \neq i_p} J_{ii_2 \ldots i_p} m_{i_2} \ldots m_{i_p}$ should be of order one. At low temperatures the m_i's take plus and minus sign. In particular, we can estimate the order of magnitude of this term working at $T = 0$ and taking $m_i = \pm$ with probability $\frac{1}{2}$. In order to make the discussion simple, let us take $p = 2$. In this case, if the strengths are of order one, F_i is a sum of N independent, identically distributed random variables, with zero mean and variance equal to 1. Thus, F_i has zero mean and variance equal to N. Thus, one can argue that F_i is of order $\sqrt{N}$. To make it finite we then chose J_{ij} to be of order $1/\sqrt{N}$ or, in other words, we impose $[J_{ij}^2]_{av} = J^2/(2N)$. The generalization to $p \geq 2$ is straightforward.

Cases that find an application in computer science are defined on random graphs with fixed or fluctuating finite connectivity. In the latter case one places the spins on the vertices of a graph with links between couples or groups of p spins chosen with a probability c. These are called *dilute spin-glasses*.

Let us now discuss some, *a priori* simpler cases. An example is the Mattis random magnet in which the interaction strengths are given by

$$J_{i_1 \ldots i_p} = \xi_{i_1} \ldots \xi_{i_p} \qquad \text{with} \qquad \xi_j = \pm \quad \text{with} \quad p = 1/2 \; . \tag{5.12}$$

In this case a simple *gauge transformation*, $\eta_i \equiv \xi_i s_i$, allows one to transform the disordered model in a ferromagnet, showing that there was no true frustration in the system.

Random ferromagnets are systems in which the strengths of the intearctions are not all identical but their sign is. Thus, there is no frustration in this systems either.

Link randomness is not the only type of disorder encountered experimentally. Random fields, that couple linearly to the magnetic moments, are also quite common; the classical model is the ferromagnetic random field Ising model:

$$E_{rfim} = -J \sum_{\langle ij \rangle} s_i s_j - \sum_i s_i h_i \qquad \text{with} \qquad P(h_i) = (2\pi\sigma^2)^{-\frac{1}{2}} \, e^{-\frac{h_i^2}{2\sigma^2}} \; . \tag{5.13}$$

Models with site or link dilution are also interesting:

$$E_{site\ dil} = -J \textstyle\sum_{\langle ij \rangle} s_i s_j \epsilon_i \epsilon_j \; , \qquad E_{link\ dil} = -J \textstyle\sum_{\langle ij \rangle} s_i s_j \epsilon_{ij} \; , \qquad . \tag{5.14}$$

with $P(\epsilon_i = 0, 1) = p, 1-p$ in the first case and $P(\epsilon_{ij} = 0, 1) = p, 1-p$ in the second. The dilute antiferromagnet in a uniform magnetic field is believed to behave similarly to the ferromagnetic random field Ising model. Experimental realizations of the former are common and measurements have been performed in samples like $\mathrm{Rb_2Co_{0.7}Mg_{0.3}F_4}$.

Once again, disorder is not only present in magnetic systems. An example that has received much attention is the case of a d dimensional directed manifold moving in an embedding $N+d$ dimensional space under the effect of a quenched random potential. The simplest case with $d = 0$ corresponds to a particle moving in an embedding space with N dimensions. If, for instance $N = 1$, the particle moves on a line, if $N = 2$ it moves on a plane and so on and so forth. If $d = 1$ one has a line that can represent a polymer, a vortex line, etc. The fact that the line is directed means that it does not have overhangs. If the line moves in a plane, the embedding space has $(N = 1) + (d = 1)$ dimensions. The quenched random potential is, *say*, $V(\vec{x})$, with $\vec{x}$ the coordinate on the embedding space and one characterises it with its probability distribution function $P(V)$.

5.4 The Griffiths phase

The effects of quenched disorder shows up already in the paramagnetic phase of a finite dimensional model, as the ones introduced above. Below the critical point of the pure case (no disorder) one sees finite regions of the system that order due to fluctuations in the couplings that can make certain parts of the sample more ferromagnetic than others. These properties manifest in non-analyticities of the free-energy that appear in a full interval of temperatures above the critical temperature of the disordered model, as shown by Griffiths. We shall not discuss these features in detail here.

5.5 The random field Ising model

The random fields give rise to many metastable states that modify the equilibrium and non-equilibrium behaviour of the RFIM. In one dimension the RFIM does not order at all, in $d = 2$ there is strong evidence that the model is disordered even at zero temperature, in $d = 3$ it there is a finite temperature transition towards a ferromagnetic state. Whether there is a glassy phase near zero temperture and close to the critical point is still and open problem.

The RFIM at zero temperature has been proposed to yield a generic description of material cracking through a series of avalaches. In this problem one cracking domain triggers others, of which size, depends on the quenched disorder in the samples. In a random magnetic systems this phenomenon corresponds to the variation of the magnetization in discrete steps as the external field is adiabatically increased (the time scale for an avalanche to take place is much shorter than the time-scale to modify the field) and it is accessed using Barkhausen noise experiments. Disorder is responsible for the jerky motion of the domain walls. The distribution of sizes and duration of the avalanches is found to decay with a power law tail cut-off at a given size. The value of the cut-off size depends on the strength of the random field and it moves to infinity at the critical point.

5.6 The spin-glass transition

Let us now discuss a problem in which disorder is so strong as to modify the nature of the low temperature phase. If this is so, one needs to define a new order parameter, capable of identifying order in this phase.

5.6.1 The order parameter

The spin-glass equilibrium phase is one in which spins "freeze" in randomly-looking configurations. In finite dimensions these configurations are spatially irregular. A snapshot looks statistical identical to a high temperature paramagnetic configuration in which spins point in both directions. However, while at high temperatures the spins flip rapidly and another snapshot taken immediately after would look completely different from the previous one, at low temperatures two snapshots taken at close times are highly correlated.

In a spin-glass state the local magnetization is expected to take a non-zero value, $m_i = \langle\, s_i \,\rangle \neq 0$, where the average is interpreted in the restricted sense discussed above, *i.e.* within a pure state. Instead, the global magnetization density, $m = N^{-1}\sum_{i=1}^N m_i$, vanishes since one expects to have as many spins pointing up as spins pointing down with each possible value of $|m_i|$. Thus, the *total magnetization*, $m = N^{-1}\sum_{i=1}^N m_i$, of a spin-glass vanishes at all temperatures and it is not a good order parameter.

The spin-glass transition is characterized by a finite peak in the linear magnetic susceptibility and a diverging non-linear magnetic susceptibility. Let us discuss the former first and show how it yields evidence for the freezing of the local magnetic moments. For a generic magnetic model such that the magnetic field couples linearly to the Ising spin, $E \to E - \sum_i h_i s_i$, the linear susceptibility is related, via the static *fluctuation-dissipation theorem* to the correlations of the fluctuations of the magnetization:

$$\chi_{ij} \equiv \left.\frac{\partial \langle s_i \rangle_h}{\partial h_j}\right|_{h=0} = \beta \,\langle\, (s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle)\,\rangle \;. \tag{5.15}$$

This relation is proven by using the definition of $\langle s_i \rangle$ and simply computing the derivative with respect to h_j. In particular,

$$\chi_{ii} = \beta \left(1 - m_i^2\right) \geq 0 \;, \tag{5.16}$$

with $m_i = \langle s_i \rangle$. The total susceptibility measured experimentally is $\chi \equiv \frac{1}{N}\sum_{ij}\chi_{ij}$. On the experimental side we do not expect to see order one sample-to-sample fluctuations in this global quantity. On the analytical side one can use a similar argument to the one presented in Sect. 5.2.2 to argue that χ should be self-averaging. Thus, the experimentally observed susceptibility should be given by

$$[\,\chi\,]_{av} = N^{-1}\sum_{ij}[\,\chi_{ij}\,]_{av} \approx N^{-1}\sum_i [\,\chi_{ii}\,]_{av} = N^{-1}\sum_i \beta\left(1 - [\,m_i^2\,]_{av}\right) \;, \tag{5.17}$$

since we can expect that cross-terms cancel under the disorder average. The fall of χ at low temperatures with respect to its value at T_c signals the freezing of the *local magnetizations*, m_i, in non-zero values. Note that this argument is based on the assumption that the measurement is done in equilibrium.

Thus, the natural *global* order parameter that characterizes the spin-glass transition is

$$q_{ea} \equiv N^{-1} \sum_i [\, m_i^2 \,]_{av} \tag{5.18}$$

as proposed in the seminal Edwards-Anderson paper. q_{ea} vanishes in the high temperature phase since all m_i are zero but it does not in the low temperature phase since the square power takes care of the different signs. Averaging over disorder eliminates the site dependence. Thus, q_{ea} is also given by

$$q_{ea} = [\, m_i^2 \,]_{av} \; . \tag{5.19}$$

In all this discussion the statistical average has to be interpreted in the same restricted sense as the one discussed in the paramagnetic - ferromagnetic transition of the usual Ising model (see Sect. 4). Whether there are two equilibrium states (as in the usual ferromagnetic phase) or many more (as we shall see appear in fully-connected spin-glass models) the average is taken in one of them. Since in the latter case it is found that even if the individual m_i take different values in different states, q_{ea} is 'state-independent', this definition does not allow to distinguish between the two scenarios.

We shall see below a more pertinent definition of an order parameter that allows one to distinguish between the simple, two-state, and the many state scenarios. In practice it is related to the definition of the *overlap* – or correlation – between two spin configurations, say $\{s_i\}$ and $\{\sigma_i\}$, defined as

$$q_{s\sigma} = N^{-1} \sum_i s_i \sigma_i \; . \tag{5.20}$$

$q_{s\sigma}$ takes values between O and 1. It equals one if $\{s_i\}$ and $\{\sigma_i\}$ differ in a number of spins that is smaller than O(N), is equals -1 when the two configurations are totally anticorrelated – with the same proviso concerning a number of spins that is not O(N) – and it equals zero when $\{s_i\}$ and $\{\sigma_i\}$ are completely uncorrelated. Note that the *self-overlap* of a configuration with itself is identically one for Ising spins.

5.6.2 Divergent susceptibility

In a pure magnetic system with a second-order phase transition, see the discussion in Sect. 4, the susceptibility of the order parameter to a field that couples linearly to it diverges at the transition. In a spin system, in particular a ferromagnet, one induces a local magnetization with a local field

$$\langle s_i \rangle = \sum_{j=1}^{N} \chi_{ij} h_j \tag{5.21}$$

with χ_{ij} the linear susceptibilities. Using this expression, the Edwards-Anderson parameter becomes

$$q_{ea} = \frac{1}{N} \sum_{i=1}^{N} [\langle s_i \rangle^2]_{av} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} [\chi_{ij} \chi_{ik} h_j h_k]_{av} \tag{5.22}$$

If the applied fields are random and taken from a probability distribution such that $\overline{h_j h_k} = \sigma^2 \delta_{jk}$ one obtains

$$q_{ea} = \frac{1}{N} \sum_{i=1}^{N} [\langle s_i \rangle^2]_{av} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} [\chi_{ij}^2]_{av} \sigma^2 = \chi_{SG} \sigma^2 \tag{5.23}$$

σ^2 acts then as a field conjugated to the order parameter q_{ea}. (One can also argue that a uniform field looks random to a spin-glass sample and thus the same result holds.) The *spin-glass susceptibility* is then defined as

$$\chi_{SG} \equiv \frac{1}{N}\sum_{ij}[\chi_{ij}^2]_{av} = \frac{1}{N}\sum_{ij}[\,(\langle s_i s_j\rangle - \langle s_i\rangle\langle s_j\rangle)^2\,]_{av} \tag{5.24}$$

and one can check whether it diverges at T_c as expected in a second-order phase transition. One can show that the divergence of χ_{SG} is related to the divergence of the non-linear usual magnetic susceptibility

$$\chi^{(3)} \equiv \frac{1}{N}\sum_{i=1} \left.\frac{\partial^3\langle s_i\rangle}{\partial h^3}\right|_{h=0} \tag{5.25}$$

that can be accessed experimentally. A careful experimental measurement of $\chi^{(3)}$, $\chi^{(5)}$ and $\chi^{(7)}$ demonstrated that all these susceptibilities diverge at T_c.

5.7 Statics of models with strong disorder

Let us discuss in this part the statics of spin-glass models and the two main competing descriptions of it. Similar arguments can be used to study other models with strong disorder, as a manifold in a random potential.

5.7.1 Scaling arguments and the droplet theory

The droplet theory is a phenomenological model that assumes that the low temperature phase of a spin-glass model has only two equilibrium states related by an overall spin flip. It is then rather similar to a ferromagnet, only that the nature of the order in the two equilibrium states is not easy to see, it is not just most spins pointing up or most spins pointing down with some thermal fluctuations within. At a glance, one sees a disordered paramagnetic like configuration and a more elaborate order parameter has to be measured to observe the order. The spin-glass phase is then called a *disguised ferromagnet* and a usual spontaneous symmetry breaking (between the two equilibrium states related spin reversal symmetry) leading to usual ergodicity breaking is supposed to take place at T_c.

Once this assumption has been done, renormalization group arguments are used to describe the scaling behavior of several thermodynamic quantities. The results found are then quantitatively different from the ones for a ferromagnet but no *novelties* appear.

Let us now just discuss one simple argument that is at the basis of what is needed to derive the results of the droplet theory without entering into the complications of the calculations.

The Imry-Ma argument

Take a ferromagnetic Ising model in a random field, defined in eq. (5.13). In zero applied field and low enough temperature, if $d > 1$ there is phase transition between a paramagnetic and a ferromagnetic phase. If one considers the effect of a random field with very strong typical strength, the the spins will align with the local external fields and the system will be paramagnetic. However, It is, however, non-trivial to determine the effect of a relatively weak random field on the ferromagnetic phase at sufficiently low temperature. The long-range ferromagnetic order could be preserved or else the field could be enough to break up the system into large but finite domains of the two ferromagnetic phases.

A qualitative argument to estimate whether the ferromagnetic phase survives or not in presence of the external random field is called the Imry-Ma argument in this context and it is a simple extension of the discussion in Sect. 4.3. Let us fix $T = 0$ and switch on a random field. If a domain $\mathcal{D}$ of a the opposite order (say down) is created within the bulk of the ordered state (say up) the system pays an energy due to the unsatisfied links lying on the boundary that is

$$\Delta E_{border} \sim 2JL^{d-1} \tag{5.26}$$

where L is the linear length of the border and $d-1$ is the dimension of the border of a domain embedded in d a dimensional volume, assuming it is compact. Typically the magnetic energy in the interior of the domain due to the external field is

$$\Delta E_{rf} \sim -hL^{d/2} \tag{5.27}$$

since there are $N = L^d$ spins inside the domain of linear length L and, using the central limit theorem, $-h\sum_{j\in\mathcal{D}} s_i \sim -h\sqrt{N} = -hL^{d/2}$. The comparison between these two energy scales yields

$$JL^{d-1} \sim -hL^{d/2} \qquad \left(\frac{J}{h}\right)^{\frac{2}{2-d}} \sim L \tag{5.28}$$

This indicates that the creation of domains at zero temperature is favorable in $d < 2$ and it is not favorable in $d > 2$. (The marginal case $d = 2$ is more subtle and we do not discuss it here.) The size of the domain is given by the expression above.

A very similar reasoning is used to argue that there cannot be spin-glass order in an Edwards-Anderson model in an external field. The only difference is that the domain wall energy is here assumed to be proportional to L^y with an *a priori* unknown d-dependent exponent y that is related to the geometry of the domains.

5.7.2 The naive mean-field approximation

Disordered models have quenched random interactions. Due to the fluctuating values of the exchanges, one expects that the equilibrium configurations be such that the spins freeze in different directions. The local averaged magnetizations need not be identical. At low-temperatures, and quite surprisingly *a priori* the naive mean-field equations (4.20) have an *exponential in N number of minima*. These can be identified as *different* equilibrium states that could be accessed by applying the corresponding site-dependent pinning fields.

Even if it has been shown by Thouless-Anderson-Palmer [16, 17] that these equations are not correct in the fully-connected disordered case, a term which is called the Onsager reaction term is missing, the fact that the number of equilibrium states diverges as an exponential of N in the thermodynamic limit remains true for these models. We shall not derive the correct TAP equations here but we shall describe the main features they predict.

The free-energy density as a function of the local magnetizations m_i defines what is usually called *the free-energy landscape*. Note that this function depends on $N \gg 1$ variables, m_i, that are not all identical in the disordered case in which the interactions between different groups of spins are different.

For all models, at high temperatures $f(m_i)$ it is characterized by a single minimum in which all local magnetizations vanish, as expected; this is the paramagnetic state.

At very low temperature there are many equilibrium states (and not just two as in an Ising ferromagnetic model) and they are not related by a symmetry (as spin reversal in the Ising ferromagnet or a rotational symmetry in the Heisenberg ferromagnet). These are characterized by non-zero values of the local magnetizations m_i that are different in different states.

On can also access the saddle-points of the free-energy landscape by looking for stationary points of $f(m_i)$ that are not the absolute minima or have unstable directions. The number and structure of saddle-points is particularly interesting in the $p \geq 3$ cases and it is indeed the reason why these models, with a random first order transition, have been proposed to mimic the structural glass arrest.

The derivation and understanding of these results is quite subtle and goes beyond the scope of these Lectures. Still, we shall briefly present the phenomenology of these models here, jus to give a flavor of their complexity. Below a temperature T_c, an exponential (in N) number of metastable states contribute to the thermodynamics in such a non-trivial way that their combined contribution to the observables makes them those of a paramagnet. Even if each of these states is non-trivial (the m_i's are different from zero) the statistical average over all of them yields results that are identical to those of a paramagnet. At a lower temperature T_s ($T_s < T_c$) there is an entropy crisis, less than an exponential number of metastable states survive, and there is a static phase transition to a glassy state.

5.7.3 The replica method

A picture that is consistent with the one arising from the naive mean-field approximation but contradicts the initial assumption of the droplet model arises from the *exact* solution of fully-connected spin-glass models. These results are obtained using a method which is called the *replica trick* and that we shall briefly present below.

In Sect. 5.2.2 we argued that the typical properties of a disordered system can be computed from the disorder averaged free-energy

$$[\, F_J \,]_{av} \equiv \int dJ P(J) F_J \; . \tag{5.29}$$

The replica method allows one to compute $[\, F_J \,]_{av}$ for fully-connected models. It is based on the smart use of the identity

$$\ln Z = \lim_{n\to 0} \frac{Z^n - 1}{n} \; . \tag{5.30}$$

Thus,

$$-\beta[\, F_J \,]_{av} = -\int dJ P(J) \ln Z_J = -\lim_{n\to 0} \frac{1}{n} \left(\int dJ P(J) Z_J^n - 1 \right) \; , \tag{5.31}$$

where we have exchanged the limit $n \to 0$ with the integration over the exchanges. For *integer* n the *replicated* partition function, Z_J^n, reads

$$Z_J^n = \sum_{\{s_i^a\}} e^{-\beta [E_J(\{s_i^1\}) + \ldots + E_J(\{s_i^n\}]} \; . \tag{5.32}$$

Here $\sum_{\{s^a\}} \equiv \sum_{\{s^1=\pm 1\}} \ldots \sum_{\{s^n=\pm 1\}}$. Z_J^n corresponds to n identical copies of the original system, that is to say, all of them with the same realization of the disorder. Each copy is characterized by an ensemble of n spins, $\{s_i^a\}$, that we label with a replica index $a = 1, \ldots, n$.

With this method one can study p spin disordered spin models. for which Z_J^n takes the form

$$Z_J^n = \sum_{\{s_i^a\}} e^{\beta \sum_{a=1}^n \left[\sum_{i_1 \neq \ldots \neq i_p} J_{i_1 \ldots i_p} s_{i_1}^a \ldots s_{i_p}^a + \sum_i h_i s_i^a \right]} \; . \tag{5.33}$$

The average over disorder amounts to computing a Gaussian integral for each set of spin indices $i_1, \ldots i_p$. One finds

$$\begin{aligned}
[\, Z_J^n \,]_{av} &= \sum_{\{s_i^a\}} e^{\frac{\beta^2 J^2}{2N^{p-1}} \sum_{i_1 \neq \ldots \neq i_p} (\sum_a s_{i_1}^a \ldots s_{i_p}^a)^2 + \beta \sum_a \sum_i h_i s_i^a} \\
&\equiv \sum_{\{s_i^a\}} e^{-\beta F(\{s_i^a\})} \; .
\end{aligned} \tag{5.34}$$

The function $\beta F(\{s_i^a\})$ is not random. It depends on the spin variables only but it includes terms that couple different replica indices:

$$\begin{aligned}
F(\{s_i^a\}) &= -\frac{\beta^2 J^2}{2N^{p-1}} \sum_{i_1 \neq \ldots \neq i_p} \sum_{ab} s_{i_1}^a \ldots s_{i_p}^a s_{i_1}^b \ldots s_{i_p}^b - \beta \sum_a \sum_i h_i s_i^a \\
&\approx -\frac{N\beta^2 J^2}{2} \left[\sum_{a \neq b} \left(\frac{1}{N} \sum_i s_i^a s_i^b \right)^p + n \right] - \beta \sum_a \sum_i h_i s_i^a \; .
\end{aligned} \tag{5.35}$$

In writing the last expression we have dropped terms that are subleading in N (in complete analogy with what we have done for the pure p spin ferromagnet). The constant term $-Nn\beta^2 J^2/2$ originates in the terms with $a = b$, for which $(s_i^a)^2 = 1$.

To summarize, we started with an interacting spin model. Next, we enlarged the number of variables from N spins to $N \times n$ replicated spins by introducing n non-interacting copies of the system. By integrating out the disorder we decoupled the sites but we payed the price of coupling the replicas. Hitherto the replica indices act as a formal tool introduced to compute the average over the bond distribution. Nothing distinguishes one replica from another and, in consequence, the "free-energy" $F(\{s_i^a\})$ is invariant under permutations of the replica indices.

The next step to follow is to identify the order parameters and transform the free-energy into an order-parameter dependent expression to be rendered extremal at their equilibrium values. In a spin-glass problem we already know that the order parameter is not the global magnetization as in a pure magnetic system but the parameter q – or more generally the overlap between states. Within the replica calculation an overlap between replicas

$$q_{ab} \equiv N^{-1} \sum_i s_i^a s_i^b \tag{5.36}$$

naturally appeared in eq. (5.35). It can be shown that this expression plays the role of an overlap between states labeled, say, by α and β

$$q_{\alpha\beta} \equiv N^{-1} \sum_i \langle s_i \rangle^\alpha \langle s_i \rangle^\beta \tag{5.37}$$

The statistics of possible overlaps is then characterized by a probability function

$$P(q) \equiv \sum_{\alpha\beta} \delta(q - q_{\alpha\beta}) \tag{5.38}$$

and it is the function that extends the concept of the order parameter and allows one to distinguish between the droplet picture prediction that corresponds to

$$P(q) = \frac{1}{2} \left(\delta(q - q_{ea}) + \delta(q + q_{ea}) \right) \tag{5.39}$$

and a more complicated situation in which $P(q)$ has the two delta functions at $\pm q_{eq}$ plus a non-zero value on a continuous support in between.

The idea is to write a free-energy density as a function of the order parameter q_{ab} and then look for the extreme that minimizes it, in complete analogy with what has been done for the fully-connected ferromagnet. This is, of course, a tricky business, since the order parameter is here a matrix with number of elements n going to zero! A recipe for identifying the form of order parameter (or the correct saddle-point solution) has been proposed by G. Parisi in the late 70s and early 80s. This solution has been recently proven to be exact for mean-field models by two mathematical physics, Guerra and Talagrand. Whether the very rich physical structure that derives from this rather formal solution survives in finite dimensional systems remains a subject of debate.

Symmetry and ergodicity breaking

Let us summarize some of the main features of the solution to the SK model obtained with replicas. There is a phase transition at a finite T_c at which the *replica symmetry* is broken. This is a rather abstract symmetry. This symmetry breaking is accompanied by ergodicity breaking as in the usual case.

Right below T_c an exponential in N number of equilibrium states appear. The transition is continuous, the order parameter approaches zero right below T_c. Lowering further the temperature each ergodic component breaks in many other ones. In this sense, the full spin-glass phase, $T < T_c$, is 'critical' and not only the single point T_c.

6 Stochastic processes

Simultaneously to the development of equilibrium statistical mechanics, which assumes that systems can be described in terms of Boltzmann-Gibbs probability distributions, the question of how this equilibrium is reached was raised (in fact by Boltzmann himself), and led to the well known debate about irreversibility and the arrow of time. Although the conceptual issues are formidable, one can take a phenomenological point of view and postulate the nature of the coupling between a given system and a thermal bath such that equilibrium is, at least in principle, reached at long times. This is known as the Langevin approach, which provides a consistent description of the *dynamics* of systems subject to thermal noise, such that the stationary solution reproduces the Boltzmann-Gibbs distribution.

On can be interested in situations such that the dynamics should lead at long times to equilibrium, but the time needed to do so is either infinite or very large compared to experimental time scales.[2] These systems are usually referred to as ‘glassy’. In these cases, equilibrium concepts are *a priori* useless; the description of the system is inherently of dynamical nature. The theoretical framework available to describe the dynamics of a system subject to thermal noise is good for both cases.

Almost any physical system is subject to fluctuations that have an unknown origin and/or can only be characterized only statistically. This noise is one of the manifestations of the exchange of energy between the system and its environment; the other accompanying feature is, as we will see in more detail below, *dissipation*. The time evolution of a system coupled to its environment can be described in two equivalent ways. One is the Langevin approach that consists in studying Newton’s equations with the addition of two terms representing friction and thermal noise. The other description is probabilistic and is concerned with the evolution of the probability distribution of the relevant degrees of freedom of the system. Both approaches turn out to be very useful to understand the dynamics of model systems.

The environment is often assumed to be a heat bath in equilibrium at a given temperature. In these conditions, many systems do equilibrate with such an environment after a short transient. The dynamics of equilibrated systems has several special features that we also review in this Chapter. How these properties are modified in systems that never reach equilibrium is a problem of current interest in research but we shall not develop it here.

6.1 The Langevin equation

The Langevin equation is a stochastic differential equation that describes phenomenologically a large variety of problems. It describes the time evolution of a set of slow variables coupled to a much larger set of fast variables that are usually (but not necessarily) assumed to be in thermal equilibrium at a given temperature. We first introduce it in the context of Brownian motion and later derive it in more generality.

---

[2] There might of course be systems that never reach thermal equilibrium, such as driven dissipative systems for example.

6.1.1 Brownian motion

The Brownian motion is the erratic motion of a mesoscopic colloidal particle (in the micrometer range) immersed in an equilibrated fluid made up of much lighter molecules. The jerky motion of the colloidal particle is due to the collisions with the molecules in the liquid. Knowing that each individual collision deflects the trajectory of the particle by a tiny amount, Langevin proposed to collect the effect of all molecules in the fluid in a time-dependent random force. If the particle moves in d dimensions, has mass m, its center of mass position at time t is represented by $\vec{x} = (x_1, \dots, x_d)$ and its velocity is $\vec{v} = \dot{\vec{x}}$, Newton's equation reads

$$m\dot{\vec{v}}(t) = \vec{F}(\vec{x}, t) + \vec{f}(t) \ . \tag{6.1}$$

The force $\vec{F}$ designates all external deterministic forces, and depends in general also on the position of the particle $\vec{x}$, while $\vec{f}$ represents the net force exerted by all the individual molecules in the fluid on the massive particle. The latter depends on the time-dependent positions of all the molecules and is a rapidly fluctuating function of time. The characteristic time for the variation of this force is related to the time interval between successive collisions, that we call τ_c, and can be estimated to be of the order of the pico-second or even shorter for a typical liquid. [3]

Due to its rapid fluctuations, the time-dependence of the force $\vec{f}$ cannot be specified. One can, instead, make reasonable assumptions about its average over a large number of identical macroscopic situations and characterize it in statistical terms. More precisely, one considers an ensemble of n systems made of one tracer particle moving in a fluid and that are prepared in identical conditions, and defines ensemble-average quantities

$$\langle \, \mathcal{O}(t) \, \rangle = \frac{1}{n} \sum_{k=1}^{n} \mathcal{O}^{(k)}_{\vec{f}}(t) \ , \tag{6.2}$$

with the label k identifying the copy in the ensemble and $\mathcal{O}$ is an observable that depends on the force $\vec{f}$. (Equivalently, one can consider a single experiment in which one uses n identical non-interacting tracer particles moving in the medium.) Equation (6.1) implies that the position and velocity of the particle are both fluctuating quantities that depend on $\vec{f}$. The aim is then to predict the average result (and the fluctuations) over a large number of experiments performed in identical conditions, or equivalently over a large number of particles provided these are non interacting.

In the absence of external forces, the problem is fully isotropic and the ensemble-averaged velocity can only vanish. If however the particle acquires a non-vanishing velocity $\vec{v}$, then the environment will react and induce a non zero force which, by symmetry, must be in the direction of $\vec{v}$. Thus, one can propose the following decomposition:

$$\vec{f} = f_1(v)\vec{v} + \xi \ , \tag{6.3}$$

where f_1 is some function of the modulus of $\vec{v}$, and ξ is an isotropic random force, the *thermal noise*, that keeps the agitation of the particle (this name has its origin in the

---

[3] This does not mean, however, that the correlation function of this random force is short range in time. Because of momentum conservation in the surrounding fluid, there appears a now well known 'long time tail' in the force correlation function, only decaying as $t^{-3/2}$ in three dimensions.

random noise that one can actually hear in electric circuits). ξ has vanishing ensemble-average at each instant t: $\langle\, \vec{\xi}_i(t)\, \rangle = 0$, for all i and all times. The average over different realizations of the history of the system introduced in eqn (6.2) corresponds now to an average over histories of the time-dependent random force, $\vec{\xi}$. Thus, henceforth the angular brackets represent:

$$\langle\, \mathcal{O}(t)\, \rangle = \int_{\infty}^{\infty} \prod_{j=0}^{\mathcal{N}_t} \prod_{i=1}^{d} d\xi_i(t_j)\, P[\xi_i(t_j)]\; \mathcal{O}_{\vec{\xi}}(t_j) \equiv \int \mathcal{D}\xi\; P(\vec{\xi})\; \mathcal{O}_{\vec{\xi}}(t) \; ,$$

where $P[\xi_i(t_j)]$ is the probability of occurrence of ξ_i at time $t_j = j\delta$, with $j = 1, \ldots, \mathcal{N}_t$, and δ an infinitesimal time-interval. In the second term above we wrote it explicitly, and in the third term we used a short-hand notation that we adopt hereafter. The subindex $\vec{\xi}$ in $\mathcal{O}$ indicates that it is evaluated in the solution to the Langevin equation and hence depends on the thermal noise realization.

The simplest assumption for $f_1(v)$ is that it tends to a constant for small velocities, leading to the familiar *friction force*:

$$f_1(v \to 0)\vec{v} \;\; = \;\; -\gamma\vec{v} \qquad \gamma > 0 \; , \tag{6.4}$$

that opposes the motion of the particle. The *friction coefficient*, γ, is proportional to the *shear viscosity* η of the medium[4], $\gamma = c\eta > 0$, with c a constant of geometric origin that depends on the size and shape of the colloidal particle. When the medium is a normal fluid, and the particle is a sphere of radius a that is much larger than the mean free-path of the molecules one has $c = 6\pi a$ and one recovers the *Stokes law* for a spherical particle in a viscous fluid. The friction coefficient γ is of the order of $n_p \ell \tau_c T$ with n_p the density of particles in the fluid, ℓ the mean free path, τ_c the average time between collisions, the temperature T is measured in units of the Boltzmann's constant k_B. For a typical liquid, like water in normal conditions, the viscosity is of the order of 0.01 Poise.[5] With the choice (6.4) eqn (6.1) becomes

$$m\dot{\vec{v}}(t) = -\gamma\vec{v}(t) + \vec{F}(\vec{x}, t) + \vec{\xi}(t) \; . \tag{6.5}$$

So far we have characterized the random force exerted by the fluid by giving its average value, the friction force. In order characterize more completely the motion of the particle, one also needs to know how the random force fluctuates in time. This information is (in part) contained in its *correlation function*, defined by comparing the thermal noise at two subsequent times t and t', with $t = t' + \tau$. Since collisions are very irregular, one can assume that the forces at two different times are statistically independent for long enough time-separations period (but see footnote 3), i.e.

$$\langle\, \xi_i(t)\xi_j(t')\, \rangle = \langle\, \xi_i(t)\, \rangle \langle\, \xi_j(t')\, \rangle = 0 \quad \text{if} \;\; \tau \equiv t - t' \gg \tau_c \quad \forall\, i, j, = 1, \ldots, d \; .$$

---

[4] The shear viscosity of a system measures its resistance to flow. A flow field can be established by placing the system between two plates and pulling them apart in opposite directions creating a *shear force*. The rate at which the plates are pulled apart is called the *shear rate*. Other geometries are also possible.

[5] The friction coefficient and the viscosity are measured in units of [mass]/[time] and Poise≡ [energy×time]/[volume] in the CGS system, respectively.

The correlation between the same component of the random force evaluated at different times is, by definition, a symmetric function of times, $\langle\, \xi_i(t)\xi_i(t')\,\rangle = \langle\, \xi_i(t')\xi_i(t)\,\rangle$. In addition, one assumes that the correlation between the same component of the force at two different times is *stationary*, that is to say, that it only depends on the time-difference τ. This is a property of the reservoir (the fluid in our case) in thermal equilibrium. Finally, since we assume that all directions of space are equivalent, the components of the random force in different directions are uncorrelated (even at time-differences that are shorter than τ_c):

$$\langle\, \xi_i(t)\xi_j(t')\,\rangle \;=\; \delta_{ij}\, g(|\tau|) \;, \qquad \forall\, i,j = 1,\dots,d \;. \tag{6.6}$$

This comes from the fact that δ_{ij} is the only rotationally invariant tensor. In the above equation, $\langle \xi_i^2(t')\rangle = g(0) > 0$ and $g(|\tau|)$ sharply peaked around $\tau = 0$ with a support of the order of τ_c. At macroscopic time-scales, τ is actually much longer than τ_c, and one can approximate $g(|\tau|)$ by a delta function of weight $2A$

$$\langle\, \xi_i(t)\xi_j(t')\,\rangle = 2A\delta_{ij}\,\delta(\tau) \;, \qquad \forall\, i,j = 1,\dots,d \;. \tag{6.7}$$

The Fourier transform of the stationary correlation of the noise defines its *spectral density*, $S(\omega)$. In this case, $S(\omega)$ is independent of the frequency ω, defining a *white noise.*

Two-time correlations do not characterize time-dependent random variables completely. This is achieved either by giving all higher-order correlations or, equivalently, by specifying its full multivariate probability distribution function. The simplest choice is a multivariate Gaussian distribution compatible with the above two-time correlation. In discrete time, we write:

$$P(\vec{\xi}) = \frac{1}{\mathcal{N}} \exp\left(-\frac{1}{4A} \sum_{j=0}^{\mathcal{N}_t} \sum_{i=1}^{d} \xi_i^2(t_j) \right) \;, \tag{6.8}$$

with $\mathcal{N}$ the normalization constant. The Gaussian hypothesis is based on the central limit theorem. Indeed, if one observes the thermal force acting on a sufficiently large particle, with a time-grid that is much larger that τ_c, $\vec{\xi}$ is the result of a large number of random forces with finite average and variance, all sharing the same distribution law. However, one could imagine 'sporadic' thermal baths that would lead to deviations from a Gaussian distribution of random forces, in particular in the tails.

The Gaussian hypothesis implies that all higher-order correlations can be expressed as functions of the two-time correlation (6.6). This is the content of Wick's theorem, which states that:

$$\langle\, \xi_{i_{2n}}(t_{2n}) \dots \xi_{i_1}(t_1)\,\rangle = \sum_{pairs} \langle\, \xi_{i_{2n}}(t_{2n})\xi_{i_{2n-1}}(t_{2n-1})\,\rangle \dots \langle\, \xi_{i_2}(t_2)\xi_{i_1}(t_1)\,\rangle \;,$$

with the sum running over all the different ways of separating the product of $2n$ variables into n pairs. The average of a product of an odd number of noise factors vanishes identically since the thermal noise has zero-average (remember that the average of the random force, namely the friction force, has been subtracted off). So far the variance, $2A$, is a free parameter. We shall see later on that if the environment is assumed to be in equilibrium at a temperature T, then A must in fact relate to γ and T.

Irreversibility and dissipation.

The friction force $-\gamma\vec{v}$ in the Langevin equation (6.5) explicitly breaks the time-reversal ($t \to -t$) invariance, a property that has to be respected by any set of microscopic dynamic equations. This is the well known paradox raised by the irreversibility of thermodynamics. However, the Langevin equation is an effective equation that only describes the particle and *not* the individual motion of the molecules of the surrounding fluid. Of course, Newton's equation describing the whole system, the particle and all the molecules of the fluid, must be time reversal invariant. However, time-reversal can be broken in the *reduced* equation, where the ability of the thermal bath to reach equilibrium is assumed from the start. The Langevin approach is a clever way to hide the irreversibility problem under the rug (by transferring the conceptual difficulties to the reservoir) and allows to investigate in a phenomelogical way the dynamics of the particle alone. In Sect. **??** we shall study a simple model where the assumptions behind Langevin's equation can be made more transparent.

Note that the energy of the particle is not conserved and, in general, flows to the bath leading to *dissipation*. At very long times however, the particle may reach a stationary regime in which the exchange of energy becomes symmetric on average: the particle gives and receives energy from the bath at equal rate.

Generation of memory.

The Langevin equation (6.5) is a first order differential equation on the velocity. The full dynamics of the particle is determined by this equation together with $\vec{v}(t) = \dot{\vec{x}}(t)$ which is also a first-order differential equation.

These features imply that the pair velocity-position of the particle at time $t+\delta$, with δ an infinitesimal time-step, depends on the pair velocity-position at time t and the value of the noise at time t. Thus, the full set of equations defines a *Markov process*, that is a stochastic process that depends on its history only through its very last step. Notice, however, that the pair of first order differential equations could also be described by a single second-order differential equation:

$$m\ddot{\vec{x}}(t) + \gamma\dot{\vec{x}}(t) = \vec{F}(\vec{x},t) + \vec{\xi}(t) \; . \tag{6.9}$$

Having replaced the velocity by its definition in terms of the position the Markov character of the process is lost. This is a very general feature: by integrating away some degrees of freedom one generates memory in the evolution. Generalizations of the Langevin equation, such as the one that we present in Sect. 6.1, and the ones that will be generated to describe the slow evolution of super-cooled liquids and glasses, do have memory.

Fluctuation – dissipation relation of the second kind.

In the heuristic derivation of the Langevin equation that we presented above the constant A is not fixed. The simplest way of fixing the value of this parameter is to study the velocity fluctuations of a Brownian particle in a constant external force. The time-dependent velocity follows from the integration over time of eqn (6.5)

$$\vec{v}(t) = \vec{v}_0 e^{-\frac{\gamma}{m}t} + \frac{1}{m}\int_0^t dt' \; e^{-\frac{\gamma}{m}(t-t')} \, [\, \vec{F} + \vec{\xi}(t')\,] \; ,$$

with $\vec{v}_0$ the initial velocity at $t=0$. Using the fact that the noise has zero average one finds

$$\langle \vec{v}(t) \rangle = \vec{v}_0 e^{-\frac{\gamma}{m}t} + \frac{\vec{F}}{\gamma}\left(1 - e^{-\frac{\gamma}{m}t}\right) \;\to\; \frac{\vec{F}}{\gamma} \quad \text{when } \; t \gg t_c^v \equiv \frac{m}{\gamma} \; .$$

Using the noise-noise auto-correlation in eqn (6.7), and setting $\vec{v}_0 = \vec{0}$ for simplicity, one readily calculates the mean-square displacement of the velocity in each direction of space, $\sigma^2_{v_i}(t) \equiv \langle\, (v_i(t) - \langle v_i(t)\rangle)^2 \,\rangle$,

$$\sigma^2_{v_i}(t) = \frac{1}{m^2}\int_0^t dt' \int_0^t dt'' \; e^{-\frac{\gamma}{m}(2t-t'-t'')} \; \langle\, \xi_i(t')\xi_i(t'') \,\rangle = \frac{A}{\gamma m}\left(1 - e^{-\frac{2\gamma}{m}t}\right) \; .$$

Since the Langevin equation is a phenomenological description of the approach to thermal equilibrium, we must impose for consistency that the above quantity saturates to the expected value calculated with the canonical distribution at temperature T. Thus,

$$\lim_{t \gg t_c^v} \sigma^2_{v_i}(t) = \frac{A}{\gamma m} = \langle\, (v_i - \langle v_i \,\rangle)^2 \rangle_{eq} = \frac{T}{m} \; ,$$

where $\langle \quad \rangle_{eq}$ denotes an average taken with Maxwell's velocity distribution. For this equality to hold one enforces that:

$$A = \gamma T \; . \tag{6.10}$$

This relation is known under the name of fluctuation–dissipation theorem (FDT) of the 'second kind' in Kubo's nomenclature. (The 'first kind' will be discussed below; these names are here a little unfortunate).

It is important to note that this FDT characterizes the surrounding fluid and not the particle, since it relates the noise-noise correlation to the friction coefficient. In the case of the Brownian particle this relation ensures that after a transient of the order of t_c, the bath maintains the mean kinetic energy of the particle constant and equal to its equilibrium value. The Gaussian distribution of the noise and the linear relation linking it to the velocity imply that the velocity of the particle is indeed distributed according to Maxwell's distribution. We shall see later that even when the environment satisfies a fluctuation – dissipation relation (FDR) the system in contact with it does not necessarily follow and satisfy an FDR itself. This is one of the main characteristics of non-equilibrium systems in contact with equilibrated environments.

Diffusion in velocity space.

For the sake of simplicity let us focus in this section on a one dimensional problem. The two-time velocity-velocity 'variogram', defined as: $\Delta_{vv}(t,t') \equiv \langle\, (v(t) - v(t'))^2 \,\rangle = C_{vv}(t,t) + C_{vv}(t',t') - 2C_{vv}(t,t')$ is a simple function of the two-time correlation of the velocity $C_{vv}(t,t') = \langle\, v(t)v(t') \,\rangle$ that itself is easily calculated to be:

$$C_{vv}(t,t') = v_0^2\, e^{-\frac{\gamma}{m}(t+t')} + \frac{T}{m}\left(e^{-\frac{\gamma}{m}|t-t'|} - e^{-\frac{\gamma}{m}(t+t')}\right) \; . \tag{6.11}$$

The first term comes from the initial condition and the second term sometimes goes under the name of *Dirichlet correlator*. When $t + t' \gg t_c$ the initial condition is forgotten, the last term vanishes, and the correlation and displacement become functions of $|t - t'|$ only, as expected in equilibrium (see Sect. 6.3.1). The variogram crosses over from a diffusive regime to saturation at the characteristic time t_c^v:

$$\Delta_{vv}(t,t') \quad \sim \quad \begin{cases} 2D_v|t-t'| & \text{when } |t-t'| \ll t_c^v \text{ with } D_v \equiv \dfrac{T\gamma}{m^2} \;, \\ \dfrac{4T}{m} & \text{when } |t-t'| \gg t_c^v \;. \end{cases}$$

The coefficient D_v of the linear regime is called the *velocity diffusion coefficient.*

Diffusion in position space.

The position-position variogram is similarly defined as $\Delta_{xx}(t,t') \equiv \langle\, (x(t) - x(t'))^2 \,\rangle$ and measures the square of the typical displacement of the particle between t and t'. Choosing $x_0 = v_0 = 0$ one finds:

$$\Delta_{xx}(t,t') = \frac{2D_v m^2}{\gamma^2} \left[|t-t'| + \frac{2m}{\gamma}\left(e^{-\frac{\gamma}{m}|t-t'|} - 1\right) - \frac{m}{2\gamma}\left(e^{-\frac{2\gamma}{m}|t-t'|} - 1\right) \right] \;,$$

and there is also a change in behavior at time-differences of the order of t_c^v:

$$\Delta_{xx}(t,t') \quad \sim \quad \begin{cases} \dfrac{2D_v}{3}|t-t'|^3 & \text{when } |t-t'| \ll t_c^v \;, \\ 2D_x|t-t'| & \text{when } |t-t'| \gg t_c^v \text{ with } D_x \equiv \dfrac{T}{\gamma} \;. \end{cases}$$

At small times, the particle is subject to a random acceleration; its velocity is thus of the order of $\sqrt{|t-t'|}$, leading to a typical displacement of the order of $|t-t'|^{3/2}$. For times larger than t_c^v, the velocity saturates under the effect of friction, and normal diffusion sets in, with a typical displacement given by $\sqrt{D_x|t-t'|}$.

Einstein relation.

The *mobility* of the particle is defined as

$$\mu \equiv \lim_{F\to 0} \lim_{t\to\infty} \frac{\langle\, v(t) \,\rangle}{F}. \tag{6.12}$$

Using the above result on D_x, we find:

$$\mu = \frac{1}{\gamma} \equiv \frac{D_x}{T} \quad \text{when } |t-t'| \gg t_c^v \;, \tag{6.13}$$

This identity between μ and D_x is known as the *Einstein relation* between the transport properties of the particle, its diffusion in real space, and the temperature of the surroundings. It expresses the *fluctuation–dissipation theorem of the first kind*, to be distinguished from the one of the second kind by the fact that it now describes a dynamic property of the particle induced by those of the bath.

Stokes-Einstein relation.

We have mentioned that in a normal fluid the viscosity and friction coefficient of the frictional force exerted on an spherical particle of radious a are related by the Stokes law $\gamma = 6\pi a\eta$. This relation, combined with (6.13) implies the *Stokes-Einstein* relation

$$6\pi a\eta = \frac{T}{D_x}$$

linking the viscosity to the temperature and spatial diffusion constant. This prediction first obtained by Einstein and then confirmed by Langevin was experimentally verified by Perrin. Much more recently, though, it has been noticed that in liquids that are super-cooled this relation ceases to be valid.

Smoluchowski (strong overdamped) limit.

In many situations in which friction is very strong the inertial term $m\dot{\vec{v}}$ can be dropped from eqn (6.5):

$$\gamma\dot{\vec{x}}(t) \;=\; \vec{F}(\vec{x},t)+\vec{\xi}(t)\;. \tag{6.14}$$

This limit is acceptable when the observation times are much longer than the characteristic time for relaxation $t_c^v = m/\gamma$, such that the mean-squared displacement of the velocity saturates and the position diffuses. The range of validity of this approximation can be easily put to test in the example of a particle moving in a harmonic potential.

In the white-noise case the friction coefficient can be eliminated from the Langevin equation in the Smoluchowski limit and the noise-noise correlation with the rescaling of time: $\tau \equiv t\gamma^{-1}$, $\tilde{x}(\tau) \equiv x(t\gamma^{-1})$.

6.1.2 Generalized Langevin equations

Langevin-like equations are used to describe the dynamics of a much more general microscopic or macroscopic systems coupled to environments. The applications in physics, chemistry and engineering are numerous. One can cite, for instance, the description of the dynamics of macromolecules in solution, with the analysis of the electrophoresis technique as a particular case. The time required to dissociate molecules or the transition rate between molecular configurations are subjects of great interest in chemistry that have been attacked with the Langevin approach. Stochastic equations with damping and white noise are also used to describe noisy electric circuits. To treat glassy problems that are typically macroscopic systems constituted by particles (colloids, atoms, molecules, spins...) in contact with an environment (the solution, phonons...) that is described statistically we need to justify them beyond the simple Brownian motion problem.

Langevin equations with multiplicative noise.

So far the random force appeared as a separate term in the Langevin equation. Extensions of the Langevin approach with *multiplicative noise* lead to the so-called *non-linear Langevin equations* [?] in which the usually delta-correlated Gaussian noise $\vec{\xi}$ multiplies a certain function $e(\vec{x},t)$ of the stochastic variable $\vec{x}$ itself:

$$m\ddot{\vec{x}}+\gamma\dot{\vec{x}} = -\vec{F}(\vec{x},t)+e(\vec{x},t)\vec{\xi}(t)\;. \tag{6.15}$$

However, as written, this equation is at best ambiguous. A proper prescription in a discrete time setting must be specified. This is a very important point that we shall not discuss here. It leads to what is called Ito and Stratonovich calculus, basically different prescriptions as to how to interpret the continuous differential equation and how to discretize the time-derivative.

Langevin equations with memory.

A derivation of a generalized linear Langevin equation with memory is very simple. In general, one studies the coupled (closed) system made by the actual system of interest in interaction with an environment. The description of the bath and of its interaction with the system depends on the problem at hand. The simplest choice is that of an independent ensemble of harmonic oscillators that couple linearly to each coordinate-like degree of freedom in the system. This choice allows one to solve the dynamic equations for the bath variables analytically. After introducing their solution in the dynamic equations for the system, one thus obtains the dynamics of the *reduced* system. Until this point the dynamics of the system remains deterministic and is completely determined by its initial conditions as well as those of the reservoir variables. The statistical element comes in when one proposes that the initial coordinates and momenta of the oscillators in the bath are distributed according to an equilibrium measure. This induces randomness and friction in the dynamics of the reduced system. After performing explicitly these calculations one ends up with the generalized Langevin equation:

$$m\dot{\vec{v}}(t) = \vec{F}(\vec{x},t) + \vec{\xi}(t) - \int_0^t dt' \, \gamma(t-t')\dot{\vec{x}}(t') \; . \tag{6.16}$$

with the Gaussian thermal noise characterized by

$$\begin{aligned} \langle \xi_i(t) \rangle &= 0 \, , \qquad \forall \, i \text{ and } \forall \, t \, , & (6.17) \\ \langle \xi_i(t)\xi_j(t') \rangle &= T \, \delta_{ij}\gamma(t-t') \, , & (6.18) \end{aligned}$$

and $\gamma(t-t')$ a *retarded friction*. A multiplicative retarded noise arises from a model in which one couples the coordinates of the oscillators to a generic function of the coordinates of the system.

Different oscillator reservoirs are characterized by different kernels $\gamma(t-t')$. The spectral function of the colored bath is usually assumed to have the form

$$S(\omega) = T\tilde{\gamma}(\omega) = 2T\gamma \left(\frac{|\omega|}{\tilde{\omega}} \right)^{\delta-1} f_c\left(\frac{|\omega|}{\Lambda} \right) \; . \tag{6.19}$$

The function $f_c(x)$ is a high-frequency cut-off of typical width Λ and is usually chosen to be an exponential. The frequency $\tilde{\omega} \ll \Lambda$ is a reference frequency that allows one to have a coupling strength γ with the dimensions of a viscosity. If $\delta = 1$, the friction is said to be *Ohmic*, $S(\omega)$ is constant when $|\omega| \ll \Lambda$ and one recovers a white noise. When $\delta > 1$ ($\delta < 1$) the bath is *superOhmic* (*subOhmic*). The exponent δ is taken to vary in the interval $[0,2]$ to avoid divergencies.

Time-dependent, $\vec{f}(t)$, and constant non-potential forces, $\vec{f}^{np}$, as the ones applied to granular matter and in rheological measurements, respectively, are simply included in the right-hand-side (RHS) as part of the deterministic force. When the force derives from a potential, $F_i(t) = -\partial_{x_i} V(\vec{x}(t))$.

In so far we have discussed systems with position and momentum degrees of freedom. Other variables might be of interest to describe the dynamics of different kind of systems. In particular, a continuous Langevin equation for classical spins can also be used if one replaces the hard Ising constraint, $s_i = \pm 1$, by a soft one implemented with a potential term of the form $V(s_i) = u(s_i^2 - 1)^2$ with u a coupling strength (that one eventually takes to infinity to recover a hard constraint). The soft spins are continuous unbounded variables, $s_i \in (-\infty, \infty)$, but the potential energy favors the configurations with s_i close to ± 1. Even simpler models are constructed with spherical spins, that are also continuous unbounded variables globally constrained to satisfy $\sum_{i=1}^N s_i^2 = N$.

6.2 Averages, correlations and responses

In Sect. 6.1 we analyzed several averaged properties of Brownian motion. In this Section we introduce the averaged observables on which we shall focus in the rest of the book. We use a vector $\vec{x}$ to encode all dynamic degrees of freedom in the system. These can be position and momenta for a particle system, spins for a magnetic sample, or even a field such as the local and time-dependent density.

6.2.1 One-time functions

The simplest dynamic observables depend only on one time and are defined by

$$\langle A(\vec{x}, t) \rangle = \int \mathcal{D}\vec{\xi}\, P(\vec{\xi})\, A(\vec{x}_\xi, t) \;, \tag{6.20}$$

where $\vec{\xi}$ symbolically denotes the noise at all times. The integral indicates a sum over all possible noise history realizations, each counted with its weight $P(\xi)$. This is a *path-integral* and we shall not discuss it here. Even if A can in principle be any complicated functional of the system's degrees of freedom, in most cases of interest it is a simple scalar function of these variables. Typical examples we shall examine later are the energy density, the magnetization density in a magnetic system, the particle density in a liquid or a glass, etc.

6.2.2 The correlation functions

Given any two functionals of $\vec{x}$, say $A(\vec{x}, t)$ and $B(\vec{x}, t)$, one defines the *correlation function* between A and B at two subsequent times t and t' as

$$C_{AB}(t, t') \equiv \langle A(\vec{x}, t) B(\vec{x}, t') \rangle = \int \mathcal{D}\vec{\xi}\, P(\vec{\xi})\, A(\vec{x}_\xi, t) B(\vec{x}_\xi, t') \;.$$

Note that the *auto-correlations* are, by definition, symmetric under exchanges of t and t', $C_{AA}(t, t') = C_{AA}(t', t)$.

In a generic situation $C_{AB}(t, t')$ is a function of *both* times t and t'. We shall see that when the system reaches a steady state, and in particular the equilibrium measure, $C_{AB}(t, t')$ becomes stationary and a function of time-differences only $C_{AB}(t, t') = C^{st}_{AB}(t - t') = C^{st}_{AB}(\tau)$ with $\tau \equiv t - t'$. In complete generality we write the two-time correlator as $C_{AB}(t, t') = \mathcal{C}_{AB}(\tau, t')$ and we define Fourier transforms with respect to the time difference τ:

$$\tilde{\mathcal{C}}_{AB}(\omega, t') \equiv \int_{-\infty}^{\infty} d\tau\, \mathcal{C}_{AB}(\tau, t')\, e^{i\omega\tau}$$

with the inverse Fourier transform given by

$$\mathcal{C}_{AB}(\tau, t') = \int \frac{d\omega}{2\pi}\, \tilde{\mathcal{C}}_{AB}(\omega, t')\, e^{-i\omega\tau} \;.$$

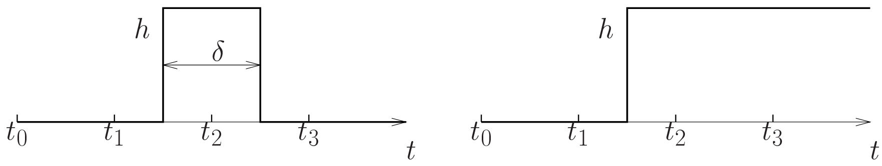


Figure 8: Sketch of perturbations of strength h; a kick (left) and a step (right).

In the stationary limit these become the usual expressions

$$\tilde{\mathcal{C}}_{AB}(\omega) \equiv \int_{-\infty}^{\infty} d\tau \; C^{st}_{AB}(\tau)\, e^{i\omega\tau} \;, \qquad C^{st}_{AB}(\tau) \equiv \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \, \tilde{\mathcal{C}}_{AB}(\omega)\, e^{-i\omega\tau} \;,$$

and $\tilde{\mathcal{C}}_{AA}(\omega)$ is the spectral density of A.

Sometimes it is useful to calculate the *connected correlation* functions in which one correlates the fluctuations of A and B about their averages:

$$C^{c}_{AB}(t,t') \equiv \langle \, (A(\vec{x},t) - \langle A(\vec{x},t)\rangle)\, (B(\vec{x},t') - \langle B(\vec{x},t')\rangle) \, \rangle \;.$$

Other important *two-time quantities* in the analysis of dynamical problems with diffusion are the *displacement between observables*:

$$\Delta^{(1)}_{AB}(t,t') \equiv \left\langle \, (A(\vec{x},t) - B(\vec{x},t'))^2 \, \right\rangle = C_{AB}(t,t) + C_{AB}(t',t') - 2\, C_{AB}(t,t') \;,$$

and the *displacement between the fluctuations* of these observables,

$$\begin{aligned} \Delta^{(2)}_{AB}(t,t') &\equiv \left\langle \, (A(\vec{x},t) - \langle A(\vec{x},t)\rangle - B(\vec{x},t') + \langle B(\vec{x},t')\rangle)^2 \, \right\rangle \\ &= C^{c}_{AB}(t,t) + C^{c}_{AB}(t',t') - 2\, C^{c}_{AB}(t,t') \;. \end{aligned}$$

6.2.3 The linear response

The application of an infinitesimal external force of strength h, possibly time-dependent, that couples linearly to a generic function B of the system's degrees of freedom, modifies the Hamiltonian according to

$$H \to H - h(t) B(\vec{x},t) \;.$$

We represent the instantaneous infinitesimal perturbation $h(t)$ as the kick between $t_2 - \delta/2$ and $t_2 + \delta/2$ in Fig. 8-left and a step-like perturbation that is continuously applied after time t_2 in the right panel of the same figure.

The variation of the average of another generic function A evaluated at time t, due to the applied force is

$$\delta\langle \, A(\vec{x},t) \, \rangle = \int_0^{\infty} dt' \; R_{AB}(t,t') \; \delta h(t') \;. \qquad (6.21)$$

This relation defines the *instantaneous linear response* $R_{AB}(t,t')$.

It is important to notice that the perturbation can modify the particle's position only at future times. This is the reason why the integration in (6.21) can only span the time interval going from the initial time $t' = 0$ to the measuring time $t' = t$. Otherwise stated, the instantaneous linear response is *causal* and hence proportional to the Heavide theta function:

$$R_{AB}(t,t') \equiv \left.\frac{\delta\langle A(t)\rangle}{\delta h(t')}\right|_{h=0} = r_{AB}(t,t')\,\theta(\tau) \qquad \tau \equiv t-t' \;. \tag{6.22}$$

Equation (6.21) then means that the change in the averaged function A due to the coupling of B to the force h exerted during a finite time-interval is a linear superposition of the changes due to spike-like perturbations applied at each instant in the same time-interval.

A very useful expression for the linear response of a Langevin process with additive noise is a consequence of Novikov's formula, which applies to general functions of Gaussian random variables with zero mean and states that

$$\langle\, x(t)\xi(t')\,\rangle = \int_0^\infty dt''\; \langle\, \xi(t'')\xi(t')\,\rangle \; \left\langle \frac{\delta x(t)}{\delta \xi(t'')} \right\rangle \;. \tag{6.23}$$

Using the noise-noise correlation (6.18) and the fact that the second factor in the integrand equals the linear response one finds

$$\langle\, x(t)\xi(t')\,\rangle = T\int_0^\infty dt''\; \gamma(t',t'')\, R(t,t'') \;. \tag{6.24}$$

For a white noise $\gamma(t',t'')=2\gamma\delta(t'-t'')$ and the relation becomes: $2\gamma T R(t,t') = \langle\, x(t)\xi(t')\,\rangle$. The Itô convention assumes that the noise term $\xi(t)$ has a strictly zero correlation time and is always slightly posterior to (and therefore independent of) any function $G(x,t)$, possibly determined by all previous values of $\xi(t')$, $t'<t$. Therefore, $\langle\, \xi(t)x(t)\,\rangle$ vanishes. The Stratonovich rule, on the other hand, assumes that the noise term $\xi(t)$ has a very small, but non zero correlation time, which is taken to zero *after* the continuous time limit of the Langevin equation has been taken. This implies that the linear responses take the boundary values:

$$\begin{array}{rll} \text{Stratonovich} & R(t,t) = 1/2\gamma \;, & R(t,t^-) = 1/\gamma \;, \\ \text{Itô} & R(t,t) = 0 \;, & R(t,t^-) = 1/\gamma \;. \end{array}$$

A function that we shall explore in great detail in future chapters is the *integrated linear response*

$$\chi_{AB}(t,t_w) \equiv \int_{t_w}^{t} dt''\; R_{AB}(t,t'') \;. \tag{6.25}$$

It represents the influence on the observable A of a perturbation applied at the *waiting-time* t_w and held constant until the measuring time t, as the one sketch in Fig. 8-right, normalized by the strength of the applied field.

The Fourier transform with respect to the time-difference τ of the response function at fixed time t' reads

$$\tilde\chi_{AB}(\omega,t') \equiv \int_{-\infty}^{\infty} d\tau\, r_{AB}(t'+\tau,t')\,\theta(\tau)\, e^{i\omega\tau} \;, \tag{6.26}$$

and defines a *generalized ac-susceptibility*. The real and imaginary parts of $\tilde\chi_{AB}$, are called χ'_{AB} and χ''_{AB}, respectively. The generalized ac-susceptibility satisfies $\tilde\chi^*_{AB}(\omega,t') = -\tilde\chi_{AB}(-\omega,t')$ and χ'_{AB} and χ''_{AB} are, respectively, even and odd functions of ω:

$$\chi'_{AB}(\omega,t') = \chi'_{AB}(-\omega,t') \;, \qquad \chi''_{AB}(\omega,t') = -\chi''_{AB}(-\omega,t') \;.$$

In equilibrium $\tilde\chi_{AB}(\omega,t') \to \tilde\chi_{AB}(\omega)$ since $\chi_{AB}(t,t') \to \chi^{st}_{AB}(\tau)$ for all observables A and B.

6.2.4 Multi correlations and responses

More general correlations and linear responses involve many time and possibly many point functions. More precisely, one defines

$$C_{A_n A_{n-1}\dots A_1}(t_n,\dots,t_1) = \langle\, A_n(\vec{x},t_n)\dots A_1(\vec{x},t_1)\,\rangle \ , \tag{6.27}$$

and the reaction of many observables to several fields linearly coupled to the system at different times; *i.e.* if $H \to H - h(t_k)A_k(t_k)$ or $H \to H - h(t_k)A_k(t_k) - h(t_{k-1})A_k(t_{k-1})$,

$$R_{A_n\dots A_1;A_k}(t_n,\dots,t_1;t_k) \;=\; \left.\frac{\delta\langle A_n(t_n)\dots A_1(t_1)\rangle}{\delta h_k(t_k)}\right|_{h=0} \tag{6.28}$$

$$R_{A_n\dots A_1;A_k,A_{k-1}}(t_n,\dots,t_1;t_k,t_{k-1}) \;=\; \left.\frac{\delta^2\langle A_n(t_n)\dots A_1(t_1)\rangle}{\delta h_k(t_k)\delta h_{k-1}(t_{k-1})}\right|_{h=0} \tag{6.29}$$

6.2.5 Absorption and dissipation

The power *absorbed* by a system is just the work done by a force applied on it measured by unit of time. More precisely, the differential work done on the system by a force $\vec{F}$ that is coupled to an observable $\vec{\mathcal{O}}$ when the latter changes by an amount $d\vec{\mathcal{O}}$ is

$$dW = -\vec{F}\cdot d\vec{\mathcal{O}} \ ,$$

and the absorbed power at time t is

$$P(t) = \frac{dW(t)}{dt} = -\vec{F}(t)\cdot\dot{\vec{\mathcal{O}}}(t) = -F_i(t)\,\frac{d}{dt}\int_0^\infty dt' R_{\mathcal{O}_i\mathcal{O}_j}(t,t')F_j(t') \ .$$

In Fourier representation, this formula becomes

$$\begin{aligned} P(t) = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega''}{2\pi}\;(-i\omega)\,\tilde{F}_i(\omega'')\tilde{F}_j(\omega')e^{-i(\omega+\omega'')t} \\ \times\int_0^\infty dt'\;e^{i(\omega-\omega')t'}\tilde{R}_{ij}(\omega,t') \ , \end{aligned}$$

where we have shorten the notation $R_{ij}\equiv R_{\mathcal{O}_i\mathcal{O}_j}$. In a steady state where $\tilde{R}_{ij}(\omega,t')$ does not depend anymore on t',

$$P(t) = -i\left(\frac{1}{2\pi}\right)^2\int_{-\infty}^{\infty}d\omega\int_{-\infty}^{\infty}d\omega''\;\omega\;\tilde{F}_i(\omega'')\tilde{R}_{ij}(\omega)\tilde{F}_j(\omega)\;e^{-i(\omega+\omega'')t} \ . \tag{6.30}$$

Dissipation represents the transfer of energy from the system to the incoherent degrees of freedom that form its environment. The sign of the friction term in the Langevin equation sets the direction of energy transfer from the system to the bath until equilibration sets in. Then, the averaged energy stabilizes and its fluctuations are only determined by the equilibrium measure. In this limit, using eqn (6.30) for the power and setting again the "initial time" to $-\infty$, one finds that the total energy dissipated is

$$E_\infty = \int_{-\infty}^{t}dt' P(t') = F_iF_j\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\omega{\chi''}_{ij}(\omega) \ ,$$

when the applied force is a percussion, $F_i(t) = F_i\delta(t)$ for all i, and

$$E_\infty = \frac{\omega_0 t}{2}\, F_i\, {\chi''}_{ij}(\omega_0)\, F_j \ ,$$

for an oscillatory force $F_i(t) = F_i \cos(\omega_0 t)$ that is applied from time $t = 0$ to time t. In these two typical cases the total energy dissipated depends only on the imaginary part of the Fourier transform of the response function, ${\chi''}_{ij}$. This is the reason why χ'' is usually called the dissipative part of the response.

6.2.6 The Kramers-Krönig relations.

The Kramers-Krönig relations connect the real and imaginary parts of the frequency dependent linear response $\tilde\chi_{AB}$. Usually they are derived in a stationary regime in which the response is invariant under translations of time. More general relations connecting the time dependent a.c. susceptibilities $\chi'(\omega, t')$ and $\chi''(\omega, t')$ can also be derived. The proof that we present below generalizes the one in Ref. [?] to the non-stationary case. The *waiting time* t' in the definition (6.26) will in fact be a dummy variable that plays no rôle in the derivation. The Fourier transform convolution theorem implies

$$\begin{aligned} \tilde\chi_{AB}(\omega, t') &= \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi}\, \tilde r_{AB}(\omega - \omega', t')\, \tilde\theta(\omega') \ , \\ \tilde r_{AB}(\omega - \omega', t') &= \int_{-\infty}^{\infty} d\tau\; e^{i(\omega-\omega')\tau}\, r_{AB}(t'+\tau, t') \ , \end{aligned}$$

and the Fourier transform of the Heaviside theta-function is

$$\tilde\theta(\omega') = \int_{-\infty}^{\infty} d\tau\; e^{i\omega'\tau}\, \theta(\tau) = \lim_{\epsilon\to 0^+} \frac{i}{\omega' + i\epsilon} = P\,\frac{i}{\omega'} + \pi\delta(\omega') \ ,$$

with P indicating the principal part of the integral. Hence,

$$\tilde\chi_{AB}(\omega, t') = \frac{1}{2\pi}\, P \int_{-\infty}^{\infty} d\omega'\, \frac{i\tilde r_{AB}(\omega', t')}{w - w'} + \frac{r_{AB}(\omega, t')}{2} \ . \tag{6.31}$$

Due to causality, the function $r_{AB}(\tau + t', t')$ is not determined for $\tau < 0$ (since it gets multiplied by $\theta(\tau) = 0$). One can take advantage of this freedom and choose it at convenience. Indeed, if one chooses r_{AB} to be symmetric in τ, $r_{AB}(-\tau + t', t') = r_{AB}(\tau + t', t')$, then $\tilde r_{AB}(\omega', t')$ is real and eqn (6.31) implies $\tilde r_{AB}(\omega', t') = 2\,\mathrm{Re}\,\tilde R_{AB}(\omega, t')$ that inserted back in eqn (6.31) yields

$$\chi''_{AB}(\omega, t') = -\frac{1}{\pi}\, P \int d\omega'\, \frac{\chi'_{AB}(\omega', t')}{\omega' - \omega} \ . \tag{6.32}$$

Instead, if one chooses r_{AB} to be antisymmetric in τ: $r_{AB}(-\tau + t', t') = -r_{AB}(\tau + t', t')$, $\tilde r_{AB}(\omega', t')$ is imaginary and eqn (6.31) implies $i\tilde r_{AB}(\omega', t') = -2\,\mathrm{Im}\,\tilde R_{AB}(\omega, t')$ that substituted in eqn (6.31) yields

$$\chi'_{AB}(\omega, t') = \frac{1}{\pi}\, P \int d\omega'\, \frac{\chi''_{AB}(\omega', t')}{\omega' - \omega} \ . \tag{6.33}$$

Equations (6.32) and (6.33) are the Kramers-Krönig relations, generalized to a non-stationary situation. Clearly, if R_{AB} reaches a stationary regime, the explicit t' dependence disappears and the usual relations are recovered.

6.3 Probabilistic formalism

In this Section we recall some features of the alternative approach to stochastic processes that consists in studying the evolution of the time-dependent probability distribution of the system's degrees of freedom, instead of writing a stochastic evolution equation for their dynamics.

6.3.1 Time-dependent probabilities

Let us consider a stochastic processes characterized by a dynamic variable $\vec{x}$. For simplicity we use a notation in which we encode all dynamic variables in a single vectorial degree of freedom $\vec{x} = (x_1, x_2, \dots, x_d)$. Later we shall make explicit the case in which one has position-like and momentum-like variables. The time-dependence of $\vec{x}$ can be dictated by a Langevin equation of the kind discussed in the previous Sections or it can be determined by other stochastic dynamic rules. Since its evolution is not deterministic, it can be known only in probabilistic sense. The knowledge of the joint time-dependent probability density

$$P(\vec{x}_n, t_n; \vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_0, t_0) \; ,$$

characterizes the process completely (we use a discretized description of all times, $t_0 \leq t_1 \leq \dots \leq t_n$, t_n being the total time). In **??** we shall use a short-hand notation denoting $\vec{X}$ the complete trajectory of the process, $\vec{X} \equiv (\vec{x}_n, t_n; \dots; \vec{x}_0, t_0)$.

The conditional probability density

$$\begin{aligned} P(\vec{x}_n, t_n; \vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_{k+1}, t_{k+1} | \vec{x}_k, t_k; \vec{x}_{k-1}, t_{k-1}; \dots \vec{x}_0, t_0) \\ = \frac{P(\vec{x}_n, t_n; \vec{x}_{n-1}, t_{n-1}; \dots; \vec{x}_{k+1}, t_{k+1})}{P(\vec{x}_k, t_k; \vec{x}_{k-1}, t_{k-1}; \dots; \vec{x}_0, t_0)} \end{aligned} \tag{6.1}$$

allows one to know the probability of finding future configurations given that the process took some specified previous values.

One can distinguish between processes for which their full history is needed to predict their future (in a probabilistic sense) and processes for which the knowledge of the present is enough to do so. The former are called *non-Markov* while the latter are *Markov*. The standard example of a Markov chain is a random walk on a d-dimensional square lattice: at each time-step the walker can jump to any of its $2d$ neighboring sites with equal probability $1/(2d)$ that is totally independent of how the walker got to its present position. A self-avoiding random walk on the same lattice is instead a non-Markovian stochastic process: since the chain cannot cross on the lattice one needs to know the positions the walker occupied at all previous times.[6] In more technical terms, a Markov process is such that

$$P(\vec{x}_n, t_n; \dots; \vec{x}_{k+1}, t_{k+1} | \vec{x}_k, t_k; \dots \vec{x}_0, t_0) = P(\vec{x}_n, t_n; \dots; \vec{x}_{k+1}, t_{k+1} | \vec{x}_k, t_k) \; .$$

This property implies that any joint probability density can be expressed in terms of conditional probabilities that depend only on the present and immediately subsequent configurations, called *transition probabilities*,

$$T(\vec{x}_k, t_k | \vec{x}_{k-1}, t_{k-1}) \equiv P(\vec{x}_k, t_k | \vec{x}_{k-1}, t_{k-1}) \; .$$

---

[6] Note however that we refer here to dynamically generated self-avoiding walks, which does not define the same statistical ensemble as that relevant for polymers in thermal equilibrium. On this point, see [?].

Hence,

$$P(\vec{x}_n, t_n; \ldots; \vec{x}_0, t_0) = T(\vec{x}_n, t_n|\vec{x}_{n-1}, t_{n-1}) \ldots T(\vec{x}_1, t_1|\vec{x}_0, t_0) P(\vec{x}_0, t_0) \; .$$

Clearly, the transition probabilities satisfy

$$\int d\vec{x}_k \; T(\vec{x}_k, t_k|\vec{x}_{k-1}, t_{k-1}) = 1 \; , \qquad T(\vec{x}_k, t_k|\vec{x}_{k-1}, t_{k-1}) \geq 0 \; .$$

Chapman-Kolmogorov integral equations.

The probability density P of a generic (and not necessarily Markov) stochastic process evaluated at t_2 is related to the one at t_1 by the integral equation

$$P(\vec{x}_2, t_2) = \int d\vec{x}_1 \; P(\vec{x}_2, t_2; \vec{x}_1, t_1) = \int d\vec{x}_1 \; P(\vec{x}_2, t_2|\vec{x}_1, t_1) P(\vec{x}_1, t_1)$$

where the integration runs over all possible values of $\vec{x}_1$. Note that t_1 and t_2 are not necessarily infinitesimally close to each other.

The conditional probabilities satisfy a similar integral equation. Indeed,

$$\begin{aligned} P(\vec{x}_3, t_3; \vec{x}_1, t_1) &= P(\vec{x}_3, t_3|\vec{x}_1, t_1) \, P(\vec{x}_1, t_1) = \int d\vec{x}_2 \, P(\vec{x}_3, t_3; \vec{x}_2, t_2; \vec{x}_1, t_1) \\ &= \int d\vec{x}_2 \; P(\vec{x}_3, t_3|\vec{x}_2, t_2; \vec{x}_1, t_1) P(\vec{x}_2, t_2; \vec{x}_1, t_1) \\ &= \int d\vec{x}_2 \, P(\vec{x}_3, t_3|\vec{x}_2, t_2; \vec{x}_1, t_1) P(\vec{x}_2, t_2|\vec{x}_1, t_1) P(\vec{x}_1, t_1) \; , \end{aligned} \tag{6.2}$$

from where we get

$$T(\vec{x}_3, t_3|\vec{x}_1, t_1) = \int d\vec{x}_2 \; T(\vec{x}_3, t_3|\vec{x}_2, t_2) T(\vec{x}_2, t_2|\vec{x}_1, t_1) \; , \tag{6.3}$$

if the process is Markov. This is the *Chapman-Kolmogorov* equation that links the transition probabilities.

Generation and elimination of memory.

Two points about the differences between Markov and non-Markov processes are worth discussing before going on.

First, the minimum time-separation, $\delta = t_{k+1} - t_k$, that is physically observable may play a role in the classification of a process as Markov or non-Markov. In most realistic situations, if one investigated the dynamics with a sufficiently fine time grid, non-Markov effects would be observable. However, for many practical purposes one can assume that these fine details are overlooked by a sparse time-grid that is longer than the characteristic memory time. Thus, the processes can be considered to be effectively Markov, in the same way as the noise-noise correlations in the Langevin approach can usually be taken to be delta functions (in time).

Besides, as we have already noticed when we rewrote the Langevin equation for the Brownian particle as a function of the coordinate only, the elimination of some degrees of freedom in the system, by integrating their dynamic equations and replacing the result in the remaining ones, may transform a Markov process into a non-Markov one. The same can occur at the level of the probabilistic description that we are discussing in this Section.

Thus, we conclude that the Markov or non-Markov character of a process may depend on the level of description we want to obtain and on the approach we adopt.

Stationarity and ergodicity.

A stochastic process is *stationary* when the joint-probability distribution is *invariant under translations of time* (TTI):

$$P(\vec{x}_n, t_n; \dots; \vec{x}_0, t_0) = P(\vec{x}_n, t_n + \Delta; \dots; \vec{x}_0, t_0 + \Delta) \; . \tag{6.4}$$

In particular, this implies that the one-time probability is independent of time, $P(\vec{x}, t) = P_{st}(\vec{x})$, and the two-time joint probability depends on the time-difference only, $P(\vec{x}_1, t_1; \vec{x}_2, t_2) = P_{st}(\vec{x}_1, t_1 - t_2; \vec{x}_2, 0)$. This immediately implies that the conditional probabilities are also functions of the time-difference only: $P(\vec{x}_1, t_1 | \vec{x}_2, t_2) = P_{st}(\vec{x}_1, t_1 - t_2 | \vec{x}_2, 0)$. For a Markov process the latter properties are sufficient to determine the stationary character of the process since all joint probabilities can be expressed in terms of one-time and transition probabilities only.

Property (6.4) immediately implies that in a stationary process the the correlation between any number of observables, $A_1, \dots, A_n$, evaluated at different times $t_1 \leq \dots \leq t_n$, is invariant under translations of time, irrespective of the values of the time differences $t_2 - t_1, \dots, t_n - t_{n-1}$,

$$C_{A_n A_{n-1} \dots A_1}(t_n + \Delta, \dots, t_1 + \Delta) = C_{A_n A_{n-1} \dots A_1}(t_n, \dots, t_1) \; .$$

In particular, one has $C_{AB}(t, t') = C^{st}_{AB}(t - t')$. Similarly, one proves that the multi-time linear responses are also stationary when (6.4) holds.

An *ergodic* Markov chain is such that all possible configuration, *i.e.* all possible values of $\vec{x}$, are accessible from any initial condition.

6.4 Master equation

Let us now focus on a Markov chain. The evolution of the conditional probability density $P(\vec{x}, t_k | \vec{x}', t_0)$ on the time sequence $t_0, \dots, t_n$, is determined by the *master equation*

$$\begin{aligned} P(\vec{x}, t_{k+1} | \vec{x}'', t_0) - P(\vec{x}, t_k | \vec{x}'', t_0) = & - \sum_{\vec{x}'} T(\vec{x}', t_{k+1} | \vec{x}, t_k) P(\vec{x}, t_k | \vec{x}'', t_0) \\ & + \sum_{\vec{x}'} T(\vec{x}, t_{k+1} | \vec{x}', t_k) P(\vec{x}', t_k | \vec{x}'', t_0) \; . \end{aligned} \tag{6.5}$$

Sometimes, one simplifies the notation and writes $P(\vec{x}, t)$ for the unknown in this equation, with the initial condition $P(\vec{x}, t_0)$ kept implicit. The LHS is the definition of the change in the probability dentisty between the subsequent times t_k and t_{k+1}. The RHS has two contributions: the first (negative) term represents the process of leaving the configuration $\vec{x}$, the second (positive) term represents the process of reaching the configuration $\vec{x}$.

The master equation (6.5) can be written in a *matricial* form. Indeed, one recasts it as

$$P(\vec{x}, t_{k+1} | \vec{x}'', t_0) = \sum_{\vec{x}'} W(\vec{x}, t_{k+1} | \vec{x}', t_k) P(\vec{x}', t_k | \vec{x}'', t_0) \tag{6.6}$$

with the matrix W written in terms of the transition probabilities, T.

Balance and detailed balance.

Any stationary solution to the master equation satisfies

$$\sum_{\vec{x}'} T(\vec{x}', t_{k+1}|\vec{x}, t_k) P_{st}(\vec{x}) = \sum_{\vec{x}'} T(\vec{x}, t_{k+1}|\vec{x}', t_k) P_{st}(\vec{x}') \qquad \forall\, \vec{x}. \tag{6.7}$$

This equation is called *balance* and it is a necessary condition to allow for a steady state. The balance condition may admit many solutions. *Detailed balance* is a restatement of the invariance of the dynamics under time-reversal in a probabilistic sense. It states that, in the steady state, any transition is balanced by its time-reversed,

$$T(\vec{x}', \delta|\vec{x}, 0) P_{st}(\vec{x}) \;=\; T(\vec{x}^R, \delta|\vec{x}'^R, 0) P_{st}(\vec{x}'^R)\;, \tag{6.8}$$

and the stationary measure is also invariant under time-reversal,

$$P_{st}(\vec{x}) \;=\; P_{st}(\vec{x}^R)\;. \tag{6.9}$$

We classified the stochastic variables of a system depending on their parity properties under time-reversal

$$x_i^R(t) \equiv x_i(-t) = \epsilon_i x_i(t) \quad \text{with} \;\; \epsilon_i = \pm 1\;,$$

for even and odd variables, respectively. One can check that detailed balance implies balance.

When the number of possible ‘microstates’ $\vec{x}$ is finite, detailed balance is sufficient (though not necessary) to ensure that the stochastic process approaches its stationary distribution function for long times.

As an example let us consider a particle system characterized by the positions and momenta of each particle. A transition corresponds to modifying the position and momentum of one particle from $(\vec{r}, \vec{v})$ to $(\vec{r}', \vec{v}')$. Since the backwards motion from $\vec{r}'$ to $\vec{r}$ occurs in the opposite direction, the reversed transition corresponds to modifying $(\vec{r}', -\vec{v}')$ into $(\vec{r}, -\vec{v})$. Detailed balance requires that these two processes be equiprobable when the system attained its stationary state, *i.e.*

$$\begin{aligned} P_{st}(\vec{r}', \vec{v}', \delta; \vec{r}, \vec{v}, 0) &= P_{st}(\vec{r}, -\vec{v}, \delta; \vec{r}', -\vec{v}', 0) \\ T(\vec{r}', \vec{v}', \delta|\vec{r}, \vec{v}, 0) P_{st}(\vec{r}, \vec{v}) &= T(\vec{r}, -\vec{v}, \delta|\vec{r}', -\vec{v}', 0) P_{st}(\vec{r}', -\vec{v}')\;. \end{aligned} \tag{6.10}$$

Note that these conditions ensure that the master equation admits a stationary solution. We still do not know if this solution is the only one or whether the process converges, asymptotically, to it. We shall discuss these two questions in Sect. 6.5. Furthermore, we have to determine, in as much generality as possible, when the stationary solution coincides with the equilibrium measure $P_{st}(\vec{x}) = P_{eq}(\vec{x})$. We shall come back to these very important issues later. Finally, we have not discussed here the effect of external fields on the time-reversal properties of a stochastic process. As we discuss in **??**. the external fields should also be time-reversed in the RHS of the generalized (6.8).

6.5 Fokker-Planck equations

The integral Chapman-Kolmogorov equation can be transformed into a differential equation under certain assumptions. Several textbooks describe this derivation [22]. Here we summarize some of the forms that this equation can take.

Kramers-Moyal expansion.

Starting from the Chapman-Kolmogorov equation, Kramers and Moyal derived a general differential equation for the conditional probabilities. This equation involves a series in which the coefficients are related to the momenta of the conditional probability. For non-Markov systems these coefficients depend on the full history of the process. For Markov processes the differential equation is local in time but it might involve an infinite number of terms. When the Kramers-Moyal expansion stops after the second term, as happens for Langevin processes with additive noise, we are left with a *Fokker-Planck equation* (see also below). In the case where the process does not 'jump', *i.e.* all displacements become small in the limit $\delta \to 0$, one can write [22]:

$$\frac{\partial P(\vec{x},t|\vec{x}',t')}{\partial t} = -\frac{\partial}{\partial x_i}\left[D_i^{(1)}(\vec{x},t)P(\vec{x},t|\vec{x}',t')\right] + \frac{\partial^2}{\partial x_i \partial x_j}\left[D_{ij}^{(2)}(\vec{x},t)P(\vec{x},t|\vec{x}',t')\right] \tag{6.11}$$

The *drift* $D_i^{(1)}(\vec{x},t)$ and the *diffusion constants* $D_{ij}^{(2)}(\vec{x},t)$ are given by:

$$\begin{aligned} D_i^{(1)}(\vec{x},t) &= \lim_{\delta\to 0}\frac{1}{\delta}\int d\vec{x}'\,(x'-x)_i T(\vec{x}',t_{k+1}|\vec{x},t_k) \\ D_{ij}^{(2)}(\vec{x},t) &= \lim_{\delta\to 0}\frac{1}{2!\delta}\int d\vec{x}'\,(x'-x)_i(x'-x)_j T(\vec{x}',t_{k+1}|\vec{x},t_k). \end{aligned} \tag{6.12}$$

The names given to these coefficients reflect their underlying physical nature, as will be clear below. The above Fokker-Planck equation describes a Brownian motion process with a local and time dependent drift and diffusion constants. There are however a much wider class of processes that admit a continuous time limit, although they lose the property of being continuous (in $\vec{x}$). Suppose that the process can make jumps of non vanishing amplitude, but more and more as $\delta \to 0$, such that the contribution to the transition probabilities $T(\vec{x}',t_{k+1}|\vec{x},t_k)$ corresponding to these jumps are proportional to δ. In this case, the Kramers-Moyal expansion must contain non local spatial terms, corresponding to these jumps, and reads

$$\begin{aligned} \frac{\partial P(\vec{x},t|\vec{x}',t')}{\partial t} &= -\frac{\partial}{\partial x_i}\left[D_i^{(1)}(\vec{x},t)P(\vec{x},t|\vec{x}',t')\right] + \frac{\partial^2}{\partial x_i \partial x_j}\left[D_{ij}^{(2)}(\vec{x},t)P(\vec{x},t|\vec{x}',t')\right] \\ &+ \int d\vec{x}''\left[\frac{\hat{T}(\vec{x},t+\delta|\vec{x}'',t)}{\delta}P(\vec{x}'',t|\vec{x}',t') - \frac{\hat{T}(\vec{x}'',t+\delta|\vec{x},t)}{\delta}P(\vec{x},t|\vec{x}',t')\right]\ , \end{aligned} \tag{6.13}$$

where the last term describes the jumps in the trajectories of the stochastic process, and $\hat{T}$ the corresponding transition probabilities. That the above decomposition between a (continuous) Brownian diffusion component and a jump component is unique in the continuous time limit is the content of the work of P. Lévy on infinitely divisible processes, and is called the *Lévy decomposition.* In some cases, the jump component admits a fractional derivative representation, generalizing the second order diffusion term. This corresponds to what is known as *Lévy flight processes* [25].

When the jump processes vanish, the Fokker-Planck equation can be recast in the form of a *local conservation law.* Indeed, defining the *probability current*

$$J_i(\vec{x},t|\vec{x}',t') \equiv D_i^{(1)}(\vec{x},t)P(\vec{x},t|\vec{x}',t') - \sum_j \frac{\partial}{\partial x_j} D_{ij}^{(2)}(\vec{x},t)P(\vec{x},t|\vec{x}',t') \ ,$$

it becomes

$$\frac{\partial P(\vec{x},t|\vec{x}',t')}{\partial t} = -\sum_i \frac{\partial J_i(\vec{x},t|\vec{x}',t')}{\partial x_i} \ .$$

Two special cases: Kramers and Smoluchowski.

The Kramers and Smoluchowski equations are particular cases that hold exactly for a linear Langevin process with white additive noise. The probability distribution of the thermal noise, $P(\xi)$, induces a time-dependent probability distribution of the dynamic variables $\vec{x}$ and $\vec{v}$:

$$P(\vec{x},\vec{v},t) \equiv \int \mathcal{D}\xi \, P(\xi)\, \delta\left(\vec{x}-\vec{x}_\xi(t)\right) \delta\left(\vec{v}-\vec{v}_\xi(t)\right) \ , \tag{6.14}$$

where we adopted the short-hand notation $P(\vec{x},\vec{v},t)$ for the conditional probability $P(\vec{x},\vec{v},t|\vec{x}',\vec{v}',t')$.

In the overdamped limit, the velocity $\vec{v}$ is slaved to the local force and $P(\vec{x},\vec{v},t)$ is replaced by an exclusive function of $\vec{x}$:

$$P(\vec{x},t) \equiv \int \mathcal{D}\xi \, P(\xi)\, \delta\left(\vec{x}-\vec{x}_\xi(t)\right) \tag{6.15}$$

that is determined by the following Fokker-Planck equation (also called *Smoluchowski equation* in this context):

$$\gamma \frac{\partial P(\vec{x},t)}{\partial t} = \frac{\partial}{\partial x_i}\left[P(\vec{x},t)\, \frac{\partial V(\vec{x})}{\partial x_i}\right] + T \frac{\partial^2}{\partial x_i^2} P(\vec{x},t) \ , \tag{6.16}$$

with the initial condition $P(\vec{x}_0,t_0)$. [Note that this equation actually applies to the conditional probability $P(\vec{x},t|\vec{x}_0,t_0)$.] See 4 for a rather rigurous proof of this equation.

In the case where inertia cannot be neglected, and noise is additive, one can establish that the probability density $P(\vec{x},\vec{v},t)$ satisfies a first order differential *Kramers equation*:

$$\begin{aligned} \frac{\partial P(\vec{x},\vec{v},t)}{\partial t} &= -\frac{\partial}{\partial x_i}\left(v_i P(\vec{x},\vec{v},t)\right) \\ &\quad + \frac{1}{m}\frac{\partial}{\partial v_i}\left[\left(\gamma v_i + \frac{\partial V(x)}{\partial x_i} + \frac{\gamma T}{m}\frac{\partial}{\partial v_i}\right) P(\vec{x},\vec{v},t)\right] \end{aligned} \tag{6.17}$$

with the initial condition $P(\vec{x}_0,\vec{v}_0,t_0)$.

It is very important to note that the balancing of factors on the RHS of the Kramers and Smoluchowski equations is a direct consequence of the equilibration of the noise (see

Sect. 6.1.2) when the equations derive from a Langevin process. It is totally equivalent to the relation (6.10) between the strength of the noise-noise correlator and the friction coefficient. More generally, it is a particular case of the detailed-balance condition (6.8), that calling $H_{fp}(\vec{x})$ the operator acting on $P(\vec{x},t)$ in the RHS of the Fokker-Planck equation, reads

$$H_{fp}(\vec{x}) P_{st}(\vec{x}) = P_{st}(\{\epsilon_i x_i\}) H^\dagger_{fp}(\{\epsilon_i x_i\}) \ . \tag{6.18}$$

The Fokker-Planck equation for a stochastic field.

A Fokker-Planck equation can also be deduced for a stochastic variable that is actually a d-dimensional field, $\vec{\phi}(\vec{x},t) = (\phi_1(\vec{x},t),\dots,\phi_d(\vec{x},t)$. It the natural generalization of the equations presented in the previous section. In the Smoluchowski limit, and for additive noise it reads

$$\frac{\partial P(\vec{\phi},t)}{\partial t} = \frac{\delta}{\delta\phi_a(\vec{x},t)} \left[P(\vec{\phi},t) \frac{\delta V[\phi]}{\delta\phi_a(\vec{x},t)} \right] + T \frac{\delta}{\delta\phi_a^2} P(\vec{\phi},t) \ .$$

Approach to stationarity.

Under rather mild conditions one can prove that if a stochastic process governed by a generic master equation admits a stationary state asymptotically, this one is unique. We present a proof of this statement below. A very simple exception to this rule is given by "decomposable" systems made of two or more non-interacting systems in which case one can construct several stationary asymptotic solutions given by linear superpositions with arbitrary coefficients of the stationary solution for each subsystem. Another case of exceptions is given by certain problems with a continuous set of possible states. We shall not discuss these special cases here but focus on the more generic situation.

Possibly, the simplest way to test under which conditions the time-dependent solution to the Fokker-Planck equation approaches a stationary form asymptotically is to use the "$\mathcal{H}$-theorem" that we discuss, with other purposes, in Comp. **??** [?, ?, ?]. The $\mathcal{H}$-functional or Lyapunov-functional

$$\mathcal{H}_{P_1,P_2}(t) \equiv \int d\vec{x}\; C\left(\frac{P_1(\vec{x},t)}{P_2(\vec{x},t)}\right) P_2(\vec{x},t) \ , \tag{6.19}$$

measures a "distance" between the (normalised) PDFs, P_1 and P_2. $C(y)$ is any strictly convex function, i.e., it satisfies $\sum_i C(\omega_i y_i) > \sum_i C(y_i)\omega_i$ for $\sum_i \omega_i = 1$ and $\omega_i \geq 0$. Customarily one uses $C(x) = x\ln x$ which suggests to relate $\mathcal{H}$ to a non-equilibrium extension of the entropy concept via $S(t) = S_o - \mathcal{H}(t)$ with S_o the thermodynamic entropy.

Now, $\mathcal{H}_{P_1,P_2}(t)$ is bounded from below and it is a monotonic decreasing function of time when (a) the diffusion matrix $D^{(2)}_{ij}$ is positive definite[7], (b) the drift matrix does not have singularities (that correspond to infinite high barriers that render the problem decomposable) and (c) P_1 and P_2 are different from zero away from infinity. Thus, the distance between any two normalized solutions vanishes asymptotically.

When the drift and diffusion matrices do not depend on time, the Fokker-Planck equation may admit a stationary solution. Based on the above argument, any other normalized

---

[7] The argument needs to be slightly modified for the Kramers equation given that its diffusion matrix is not positive definite; still, one also proves that the asymptotic solution is unique in this case [?].

solution necessarily approaches the stationary one in the long-time limit. Two types of stationary solution can be identified. The simpler ones have constant current $J_i(\vec{x}) = J_i$ fixed by the boundary conditions. Natural boundary conditions demand $J_i = 0$, when $x_j \to \infty$ for all j, thus, $J_i = 0$ everywhere in space, in which case the stationary solution takes the *potential form*

$$P_{st}(\vec{x}) \propto e^{-\phi(\vec{x})} , \qquad \phi(\vec{x}) = -\int_{\vec{x}} dx'_k \left[D^{(2)^{-1}}_{ki}(\vec{x}') \left(D^{(1)}_i(\vec{x}') - \frac{\partial D^{(2)}_{ij}(\vec{x}')}{\partial x_j{}'} \right) \right]$$

where a sum over repeated indices is assumed. In the simple case $D^{(1)}_i(\vec{x}) = -\partial_{x_i} V(\vec{x})$ and $D^{(2)}_{ij}(\vec{x}) = T\delta_{ij}$ one easily checks that this expression becomes the Boltzmann factor. The question then arises as to whether an explicit solution is possible when the drift term is not of a potential form.

Another family of solution correspond to non-equilibrium steady states (NESS) with a divergenceless, time-independent current that is not necessarily constant in space. When the drift term itself is divergenceless and $D^{(2)}$ is constant – corresponding for example to tracer particles advected by a incompressible convection flow and subject to molecular diffusion – then the NESS is trivial and is given by $P_{st}(\vec{x}) = 1/V$, where V is the total volume accessible to the particle. The local current is then given by $J_i = D^{(1)}_i(\vec{x})/V$ and is by construction divergenceless. Another case that can be solved in full generality is when the drift term is the sum of a potential part and a divergenceless part that are everywhere orthogonal to each other. Then, the divergenless part simply advects the particles along equi-potential lines and does not modify the standard equilibrium Boltzmann factor. In the completely general case of an arbitrary drift and diffusion structure, there are in general no explicit construction of the NESS (for examples where this construction is possible, see Haake, Derrida et al., Jona Lasinio).

Approach to equilibrium.

Whether a stochastic process approaches equilibrium asymptotically depends on the nature of the forces applied, the boundary conditions, etc. In the following we shall focus on Fokker-Planck processes with no jumps. Among these, a subclass admit a spectral representation of the Fokker-Plank operator with generic complex eigenvalues. The asymptotic analysis has to be performed on a case-by-case basis.

A more restrictive class of systems are represented by self-adjoint FP operators. The potential case, $D^{(2)}_{ij} = T\delta_{ij}$ and $D^{(1)}_i = -\partial_{x_i} V$, is a special problem in this class and one proves that the process does indeed approach a stationary solution that is given by the canonical equilibrium measure.

An easy and elegant proof of this statement, for a Smoluchowski potential problem relies on a mapping of the Fokker-Planck equation to the Schrödinger equation [?]. Introducing

$$P(\vec{x}, t) \equiv c\, e^{-\frac{\beta}{2} V(\vec{x})} p(\vec{x}, t) \tag{6.20}$$

with c a positive constant the Fokker-Planck equation becomes

$$\begin{aligned} \frac{\partial p(\vec{x}, t)}{\partial t} &= \left[T \frac{\partial^2}{\partial x_i^2} - \left(-\frac{1}{2} \frac{\partial^2 V(\vec{x})}{\partial x_i^2} + \frac{\beta}{4} \left(\frac{\partial V(\vec{x})}{\partial x_i} \right)^2 \right) \right] p(\vec{x}, t) \\ &= -L_s p(\vec{x}, t) \; . \end{aligned} \tag{6.21}$$

This is a Schrödinger equation in imaginary time. The potential energy term, say V_s, is a function of the original potential V. $p(\vec{x},t)$ is related to the probability density and plays the role of a wave function, while in true quantum mechanics it is the modulus squared of the wave function which has a probability interpretation. The operator L_s admits a spectral representation in terms of its eigenvalues λ_n and its eigenvectors ψ_n. Since the potential V_s is independent of time, the eigenvectors, ψ_n, are time-independent, $\psi_n(\vec{x})$. If the potential V_s grows to infinity sufficiently fast when $x_i \to \pm\infty$, for all i, the spectrum is discrete and there is gap between the lowest eigenvalue, λ_0, and the first excited state, λ_1. The ground state wavevector is everywhere positive and can hence describe a probability. One can also show that the operator L_s can be written as the square of a certain operator, and is thus positive semi-definite in a Hilbert space[8] and the eigenvalues, λ_n, are real and satisfy $\lambda_n \leq 0$.

The solution to eqn (6.21) can be expressed as the series

$$p(\vec{x},t) = \sum_n c_n \psi_n(\vec{x}) e^{-\lambda_n t} \;, \tag{6.22}$$

with c_n arbitrary numerical constants. When a stationary solution exists, the operator has a vanishing eigenvalue, $\lambda_0 = 0$, that, if the conditions mentioned in the previous subsection are satisfied, is not degenerate. One has $p_{st}(\vec{x}) = c_0\psi_0(\vec{x})$. The eigenstates associated to non-zero eigenvalue can be degenerate in which case one needs a more refined notation to distinguish their associated eigenvectors. We skip this detail and we continue to use the simplified notation above. If there is a gap in the spectrum, $\lambda_1 > 0$, in the long-time limit only the contribution of the zero eigenvalue survives and $p(\vec{x},t) \to p_{st}(\vec{x}) = c_0\psi_0(\vec{x})$. The constant c_0 is fixed to one by the normalization of the probability. Indeed, $c_0 = \int d\vec{x}\psi_0(\vec{x})p(\vec{x},0) = \int d\vec{x}P(\vec{x},0) = 1$.

Going back to the original PDF, $P(\vec{x},t)$, one has

$$\lim_{t\to\infty} P(\vec{x},t) = \psi_0^2(\vec{x}) = c^2 e^{-\beta V(\vec{x})} = \frac{e^{-\beta V(\vec{x})}}{\int d\vec{x}\, e^{-\beta V(\vec{x})}}$$

where we used the conservation of probability to compute c^{-2}. Thus P_{eq} is indeed the asymptotic solution to the Fokker-Planck equation.

Note that this argument assumes that a sufficiently long t ($t > t_{eq}$) is reached such that only the $\lambda_0 = 0$ term survives in the sum. This hypothesis does not hold in the asymptotic analysis for the relaxing models we analyze in the next Sections. If the next eigenvalue λ_1 does not vanish, its inverse is the time-scale needed to equilibrate the model. If however, there is no gap in the spectrum, one does not have a simple argument to estimate how long one should wait until the asymptotic limit is reached. This is a question that will be raised regularly in the treatment of glassy dynamics. Moreover, when non-potential or time-dependent forces are exerted on the system the transformation (6.20) is not sufficient to deal with their effect and equilibrium cannot be established.

Since the eigenvectors corresponding to excited states are not everywhere positive, one cannot interpret $\psi_0\psi_n$ directly as a probability. Interestingly however, one can construct linear combination of these that can be interpreted as *metastable states*, that can be defined as states with lifetimes longer than a certain fixed but long time-scale.

---

[8] This space is defined via the scalar product $(f,g) \equiv \int d\vec{x} f(\vec{x})g(\vec{x})/P_{st}(\vec{x})$ and an operator $\mathcal{L}$ is self-adjoint if $(f,\mathcal{L}g) = (\mathcal{L}f,g)$.

6.6 Quantum mechanical formulation

Kadanoff and Swift proposed to use a quantum mechanical notation to represent Fokker-Planck equations. Note that this is not identical to the above mapping to a Schrödinger equation since the Fokker-Planck Hamiltonian is non-Hermitian. This notation is very useful to prove simple properties of stochastic processes as the ones holding near equilibrium that we discuss in Sect. **??**. Indeed, identifying $-i\partial/\partial x_i$ with the operator $\hat{p}_i$ the usual commutation relations between momenta and coordinates are recovered:

$$-i\frac{\partial}{\partial x_i} \to \hat{p}_i \qquad [\hat{p}_i, \hat{x}_j] = -i\delta_{ij} \ ,$$

while $[\hat{x}_i, \hat{x}_j] = 0$ and $[\hat{p}_i, \hat{p}_j] = 0$, for all i, j. The probability distributions $P(\vec{x}, t)$ is then identified with a quantum time-dependent "state" $|P(t)\rangle$ and the Fokker-Planck equation reads

$$\frac{\partial |P(t)\rangle}{\partial t} = \hat{H}_{fp} |P(t)\rangle \ ,$$

where $\hat{H}_{fp}$ is written in terms of the canonically conjugated operators $\hat{\vec{x}}$ and $\hat{\vec{p}}$, and the transition probability is

$$T(\vec{x}, t+\tau | \vec{x}', t) = \langle \vec{x} | e^{-\hat{H}_{fp}\tau} | \vec{x}' \rangle \ .$$

The detailed balance property (6.18) implies that

$$e^{\beta\hat{H}(\{\epsilon_i x_i\}})\hat{H}_{fp}(\vec{x})e^{-\beta\hat{H}(\vec{x})} = \hat{H}^{\dagger}_{fp}(\{\epsilon_i x_i\}) \ , \tag{6.23}$$

The form of the Fokker-Planck operator depends on the stochastic processes considered. In the case of a stochastic processes of Kramers' type, describing a one dimensional system with position x and momentum p, one works with two "position operators", $(\hat{x}, \hat{p})$, and two "momentum operators", $(\hat{p}_x, \hat{p}_p)$, and the Hamiltonian reads

$$\hat{H}_k = -i\hat{p}_x v + \frac{1}{m}\, i\hat{p}_v\, V'(\hat{x}) + i\hat{p}_v \left(\frac{\gamma}{m}\hat{v} + \frac{\gamma T}{m^2} i\hat{p}_v \right) \ .$$

For the Smoluchowski equation one has one pair of conjugated variables, $(\hat{x}, \hat{p})$, and

$$\hat{H}_s = \frac{1}{\gamma}\, \hat{p}\, \left(i\hat{V}'(x) - T\hat{p} \right) \ .$$

6.7 The fluctuation – dissipation theorem

The fluctuation-dissipation theorem (FDT) relates the correlations of spontaneous equilibrium fluctuations to the response induced by an external field. It is a *model independent* relation between the linear response and its associated correlation function that takes somewhat different forms for classical and quantum system. The latter reduces to the former when quantum fluctuations become irrelevant. In this Section we present several proofs of the FDT. When the system is out of equilibrium hypothesis, on the other hand, this relation in general does not hold, although a generalized version of the FDT can in some cases be established.

6.7.1 Static FDT

Many relations between correlations of fluctuations and susceptibilities are known in statistical mechanics. All these are different statements of the *static* FDT.

Take for instance a perfect gas. The fluctuations in the density $\rho = n/\tilde{V}$ where n is the number of particles within a sub-volume $\tilde{V}$ of a system with N particles and volume V, are defined as: $\sigma_\rho^2 \equiv \langle (\rho - \langle \rho \rangle)^2 \rangle$. In the thermodynamic limit $N \to \infty, V \to \infty$ with $N/V = \overline{\rho}$ fixed, these are related to the isothermal compressibility $\chi_T = -V^{-1} \partial V / \partial P|_T$ via $\sigma_\rho^2 = T \overline{\rho}^2 \chi_T / \tilde{V}$. This relation is a form of FDT.

For a system in equilibrium with a thermal reservoir at temperature T one has

$$\chi \equiv \left. \frac{\delta \langle A \rangle_h}{\delta h} \right|_{h=0} = \beta \langle (A - \langle A \rangle)^2 \rangle \tag{6.24}$$

for any observable A. The average $\langle \ \rangle_h$ is calculated with the partition function of the system in the presence of a small field coupled to A in such a way that the Hamiltonian reads $H = H_0 - hA$. For a magnetic system this equation relates the magnetic susceptibility to the magnetization fluctuations.

When the order parameter is a field, for instance a scalar one, described by a Hamiltonian or a Landau-type free-energy that is linearly modified by an external perturbation, $\mathcal{F}[\phi] \to \mathcal{F}[\phi] - \int d^dx \ h(\vec{x}) \phi(\vec{x})$, the static FDT (6.24) gets generalized to

$$\chi(\vec{x}, \vec{x}') \equiv \left. \frac{\delta \langle \phi(\vec{x}) \rangle}{\delta h(\vec{x}')} \right|_{h=0} = \beta \langle \, \phi(\vec{x}) \phi(\vec{x}') \, \rangle_c \equiv \beta C(\vec{x}, \vec{x}') \ .$$

6.7.2 Dynamic FDT

There are several proofs of this theorem. We present two of them that are particularly instructive since they use, very explicitely, the two necessary hypothesis which are *i.* detailed balance, and *ii.* equilibration.

FDT from the master equation.

For the sake of simplicity, let us consider a system with a single dynamical variable x that evolves according to a Markov process on a discrete time-grid $t_k = k\delta$. In what follows we simplify the notation and we eliminate the time-dependence from the transition probabilities, $T_h(x_k|x_{k-1}) \equiv T_h(x_k, t_k|x_{k-1}, t_{k-1})$. The subindex h indicates that the perturbation is present during the time-step $t_{k-1} \to t_k$. If the system reached equilibrium at time t_{k-1} one has

$$\sum_{x_l; 1 \leq l \leq k-1} \prod T(x_l|x_{l-1}) P(x_0, t_0) = P_{eq}(x_{k-1}) \ . \tag{6.25}$$

We will use a superscript h whenever equilibration has been achieved in the presence of a field, $P_{eq}^h(x_{k-1})$.

The correlation between two observables $A(x)$ and $B(x)$ in the absence of any perturbation is given by

$$\langle A(x_j) B(x_k) \rangle = \sum_{x_l} A(x_j) T(x_j|x_{j-1}) \dots T(x_{k+1}|x_k) B(x_k) T(x_k|x_{k-1}) P_{eq}(x_{k-1}) \ .$$

The variation with respect to h of the mean value of $A(x)$ evaluated at time t_j after a step-like perturbation of strength h that modified the energy as $H \to H - hB(x)$ between t_{k-1} and t_k is

$$\frac{\delta \langle A(x_j)\rangle_h}{\delta h_k} = \sum_{x_l} A(x_j) T(x_j|x_{j-1}) \dots T(x_{k+1}|x_k) \left. \frac{\delta T_h(x_k|x_{k-1})}{\delta h} \right|_{h=0} P_{eq}(x_{k-1})$$

Let us assume that the variable x is even with respect to time-reversal. In this case, and if detailed balance holds, $T_h(x_k|x_{k-1}) P^h_{eq}(x_{k-1}) = T_h(x_{k-1}|x_k) P^h_{eq}(x_k)$, and

$$\frac{\delta T_h(x_k|x_{k-1})}{\delta h} = \frac{\delta}{\delta h} \left(\frac{P^h_{eq}(x_k)}{P^h_{eq}(x_{k-1})} \right) T_h(x_{k-1}|x_k) + \frac{P^h_{eq}(x_k)}{P^h_{eq}(x_{k-1})} \frac{\delta T_h(x_{k-1}|x_k)}{\delta h} \tag{6.26}$$

that has to be evaluated at $h = 0$. The first term is easily calculated by replacing $P^h_{eq}(x_k) = \exp\left(-\beta V(x_k) + \beta h B(x_k)\right)/Z_h$and yields

$$\beta (B(x_k) - B(x_{k-1})) \frac{P_{eq}(x_k)}{P_{eq}(x_{k-1})} T(x_{k-1}|x_k) \; . \tag{6.27}$$

The second term in eqn (6.26) vanishes identically; indeed, after simplification, x_{k-1} is seen to appear only in the derivative of the transition probability. Using the h independent normalisation condition, $\sum_{x_{k-1}} T_h(x_{k-1}|x_k) = 1$ one finally finds that this term is zero. Thus

$$\left. \frac{\delta \langle A(x_j)\rangle_h}{\delta h_k} \right|_{h=0} = \beta \, \langle \, A(x_j) \, (B(x_k) - B(x_{k-1})) \, \rangle \; , \qquad j \geq k \; . \tag{6.28}$$

This relation expresses the FDT in discrete time. Interestingly enough, we did not need to specify the h dependence of the transition probabilities T, provided this dependence is compatible with detailed balance.

Note that since in equilibrium the averages of one-time quantities are constant one can replace $\langle A(x_j) B(x_k)\rangle$ by the connected correlation $\langle A(x_j) B(x_k)\rangle - \langle A(x_j)\rangle \langle B(x_k)\rangle$ and the FDT also reads

$$\left. \frac{\delta \langle A(x_j)\rangle_h}{\delta h_k} \right|_{h=0} = \beta \, \langle \, A(x_j) \; (B(x_k) - \langle B(x_k)\rangle - B(x_{k-1}) + \langle B(x_{k-1})\rangle) \, \rangle \; , \quad j \geq k \; .$$

In its integrated form

$$\chi_{jl} \equiv \sum_{k=l}^{j} \left. \frac{\delta \langle A(x_j)\rangle_h}{\delta h_k} \right|_{h=0} = \beta \left(\langle \, A(x_j) B(x_j) \, \rangle - \langle \, A(x_j) B(x_l) \, \rangle \right) \; . \tag{6.29}$$

The RHS can also be written in terms of connected correlations.

The LHS in eqn (6.28) is just the linear response, $R_{AB}(t_j, t_k)$, while the RHS is a discretized time-derivative of the correlation, $C_{AB}(t_j, t_k)$, with respect to the earlier time t_k. We have already assumed that the system is equilibrated at time t_{k-1}. Thus, all correlations and responses are stationary and we have:

$$R^{st}_{AB}(t_j - t_k) = \beta \, (\, C^{st}_{AB}(t_j - t_k) - C^{st}_{AB}(t_j - t_{k-1}) \,) \; , \qquad j \geq k \; .$$

that in the continuous limit $t_k - t_{k-1} \to 0$ becomes

$$R^{st}_{AB}(t-t') = \beta\, \partial_{t'} C^{st}_{AB}(t-t') \;, \qquad t \geq t' \;. \tag{6.30}$$

One can also express the FDT in an equivalent form in which the correlation is replaced by the displacement:

$$R^{st}_{AB}(t-t') = -\frac{\beta}{2}\, \partial_{t'} \Delta^{st}_{AB}(t-t') \;.$$

In a problem that undergoes normal diffusion this expression implies the Einstein relation (6.13).

If one deals with a more complicated Markov chain with several degrees of freedom, and with both even and odd components of the transition rates with respect to time-reveral, the proof carries through in the same way *iff* the perturbation applied does not break time-reversal. In other words, perturbations that modify the Hamiltonian as

$$H \to H - hB(\{x_i\}) \quad \text{with } B(\{x_i\}(-t)) = B(\{x_i\}(t))$$

admit an FDT of the form given above. For instance, if we deal with a particle characterised by its position and velocity, the function B can have any depence on the position but it must be an even function of the velocity. An example would be to perturb the kinetic energy of the particle.

A fluctuation-dissipation relation for the velocities takes the form:

$$TR_{vx}(t,t') = T\frac{\delta\langle\, v(t)\,\rangle}{\delta h(t')} = \langle\, v(t)v(t')\,\rangle = C_{vv}(t,t') \;, \tag{6.31}$$

when the field h perturbs the system by being coupled to the coordinate at time t'. This property is a version of a *Kubo formula* and it can be useful in problems in which it is easy to measure a current. Being very sloppy, one can "derive" it from (**??**) by taking a derivative with respect to t on both sides and exchanging the average over the noise and the time-derivatives (a highly non-trivial operation).

FDT in the frequency domain.

Experimental devices often measure the correlation and response function in the frequency domain. The Fourier transform of the FDT relation (6.30) reads:

$$\tilde{\chi}_{AB}(\omega) = \beta\left(C^{st}_{AB}(0) + i\omega \int_0^\infty d\tau\; e^{i\omega\tau} C^{st}_{AB}(\tau) \right) \;. \tag{6.32}$$

The imaginary part of this expression leads to the FDT in the frequency domain:

$$\frac{\chi''(\omega)}{\omega} = \beta\, \mathrm{Re} \int_0^\infty d\tau\; e^{i\omega\tau} C^{st}_{AB}(\tau)$$

There is no extra information in the real part, since it is related to $\chi''(\omega)$ by the Kramers-Krönig relations (6.32) or (6.33).

When the FDT is written in the above form we see the justification of its name. As discussed in Sect. 6.2 $\chi''(\omega)$ quantifies the dissipation and the RHS is related to the power spectrum of the spontaneous fluctuations in the system.

6.8 Dynamics of discrete variables

6.8.1 Glauber dynamics

In Sect. **??** we presented the Langevin stochastic differential equation that describes the time evolution of continuous variables in contact with a heat bath. The question remains as to how to describe the time evolution of discrete, *e.g.* Ising spin variables, in contact with a reservoir. One way of doing it is to use the Langevin equation extending the discrete variables to become soft spins and, at the end of the calculation, imposing the hard constraint, $s = \pm 1$. Another way is to propose updates of the trully discrete spins that satisfy detailed balance and include the stochastic effect of the coupling to the environment.

Glauber proposed such a dynamic rule; his proposal is now known as *Glauber dynamics*. Let us first consider the simplest possible problem of a single Ising variable, $s = \pm 1$, and update it randomly. If there is no magnetic field, none of the two states $s = \pm 1$ is favoured. If the rate per unit time at which the spin makes a transition to its opposite state is $\alpha/2$,

$$s(t+dt) \;=\; \left\{ \begin{array}{lll} -s(t) & \text{with probability} & \frac{\alpha}{2}dt \\ \;\;\, s(t) & \text{with probability} & 1-\frac{\alpha}{2}dt \end{array} \right. , \qquad (6.33)$$

then the probability that the spin s takes the value σ at time $t + dt$ satisfies the equation

$$p(s, t+dt) = p(s,t) \times \left(1 - \frac{\alpha}{2}\, dt\right) + p(-s,t) \times \frac{\alpha}{2}\, dt \; . \qquad (6.34)$$

Taking the limit $dt \to 0$ one has

$$\frac{dp(s,t)}{dt} = p(-s,t)\frac{\alpha}{2} - p(s,t)\frac{\alpha}{2} \; . \qquad (6.35)$$

This is the master equation that determines the evolution of the probability distribution from an initial state characterized in probability. Randomly chosen initial conditions should satify

$$p(s,0) + p(-s,0) = 1 \; . \qquad (6.36)$$

One can easily check, by writing the evolution equation for $p(s,t)$ and $p(-s,t)$, that eq. (6.35) conserves the probability, *i.e.* given (6.36), $p(s,t)+p(-s,t) = 1$ for all t follows. $p(s,t)$ can be computed by noting that $\langle s \rangle(t) = p(1,t) - p(-1,t) = [p(1,0) - p(-1,0)]e^{-\alpha t}$ and $p(s,t) = \frac{1}{2}\{1 + s[p(1,t) - p(-1,t)]\}$.

Let us now consider an Ising model made of N spins $s_i = \pm 1$, for $i = 1, \dots, N$ in interaction. At each instant the system is characterized, in probability, by the joint probability function $P(s_1 = \sigma_1, \dots, s_N = \sigma_N; t)$. Taking a sufficiently short time interval dt such that only one spin among the N in the system can attempt a flip, and calling $T_j(s_j)dt$ the probability that spin j changes its sign in dt, in analogy with (6.37) one has

$$(s_1, \dots, s_N)(t+dt) \;=\; \left\{ \begin{array}{llll} (s_1, \dots, -s_j, \dots, s_N)(t) & \text{with prob} & T_j(-s_j)\, dt & \text{for all } j \\ \;\;(s_1, \dots, s_j, \dots, s_N)(t) & \text{with prob} & 1 - \sum_{j=1}^{N} T_j(s_j)\, dt & \end{array} \right. \qquad (6.37)$$

Proceeding as above, the joint probability distribution is given by

$$\begin{aligned} p(s_1,\dots,s_N;t+dt) &= p(s_1,\dots,s_N;t)\times\left[1-\sum_{j=1}^N T_j(s_j)\,dt\right] \\ &\quad +\sum_{j=1}^N p(s_1,\dots,-s_j,\dots,s_N;t)\times T_j(-s_j)\,dt\;. \end{aligned} \tag{6.38}$$

Taking the limit $dt \to 0$ one has

$$\frac{dp(s_1,\dots,s_N;t)}{dt} = -p(s_1,\dots,s_N;t)\sum_{j=1}^N T_j(s_j)+\sum_{j=1}^N p(s_1,\dots,-s_j,\dots,s_N;t)T_j(-s_j)\;. \tag{6.39}$$

There is now quite a big freedom to chose the transition probabilities $T_j(s_j)$. They can, in particular, depend on the configuration of the neighbouring spins to s_j and not only on s_j. The constraint

$$P_{eq}(s_1,\dots,s_N)\sum_{j=1}^N T_j(s_j) = \sum_{j=1}^N P_{eq}(s_1,\dots,-s_j,\dots,s_N)T_j(-s_j)\;, \tag{6.40}$$

on the transition probabilities, with P_{eq} the Boltzmann weight, ensures that this expression is a stationary solution of the dynamics. This condition is satisfied if, in particular, one asks the stronger *detailed balance condition*

$$P_{eq}(s_1,\dots,s_N)T_j(s_j) = P_{eq}(s_1,\dots,-s_j,\dots,s_N)T_j(-s_j) \qquad \text{for all} j\;. \tag{6.41}$$

For the Ising model with generic two-body interactions one can simplify these equation using

$$\frac{P_{eq}(s_1,\dots,s_j,\dots,s_N)}{P_{eq}(s_1,\dots,-s_j,\dots,s_N)} = \frac{e^{-s_j\sum_{i(j)}J_{ij}s_i}}{e^{s_j\sum_{i(j)}J_{ij}s_i}} = \frac{T_j(-s_j)}{T_j(s_j)} \tag{6.42}$$

where $i(j)$ means that the sum runs only over spins connected to the j-th spin: nearest neighbours on the lattice for finite dimensional systems, all other spins in the system for fully connecetd models, something in between for models defined on random lattices. $\sum_{i(j)} J_{ij}s_j$ is the *local field* acting on s_j. Now, one can check by simple enumeration of all the possible configurations that the *choice*

$$T_j(s_j) = \frac{1}{2}\left[1-s_j\tanh(\beta\sum_{i(j)}J_{ij}s_i)\right] \tag{6.43}$$

satisfies eq. (6.42).

6.8.2 Montecarlo dynamics

This is a numerical technique to simulate the temporal evolution of a system including the stochastic effect of its coupling to a heat bath. One uses a master equation with transition probabilities that satisfy detail balance. At each step a spin in the sample, say s_i, is chosen at random. It is then turned with probability

$$W_i = \begin{cases} e^{-\Delta E_i/T} & \Delta E_i \geq 0 \; , \\ 1 & \Delta E_i < 0 \; , \end{cases} \tag{6.44}$$

where $\Delta E_i = E(-s_i) - E(s_i)$ where E is the energy of the system and $-s_i$ is the reversed value of the spin s_i. The spin flips whether W_i exceeds a random number taken from a uniform distribution between 0 and 1. This rule implies that all updates that are favourable in energy ($\Delta E_i < 0$) are accepted while some updates that increase the energy of the system are also acceped [with probability (6.44)]. These *a priori* unfavourable moves might help the system get out of metastable configurations and reach equilibrium. The unit of time is defined as a Montecarlo sweep, that is to say, N attemps to flip a spin (note that due to the random choice of spins, some will appear more than once and others will not appear within the N chosen ones).

7 Interface and surface growth

An interface is a frontier separatiing two regions of space with two phases. It could be the border between water and oil in a liquid mixture, the border between regions with positive and negative magnetization in a magnet, the limit of a fluid invading porous media, *etc.* As a physicist one would like to characterize the static and dynamic properties of such a surface.

7.1 Generic

The morphology of an interface depends on the length scale of observation: the Alps look rough on Earth but thy look smooth seen from the Moon.

7.2 Domain growth

Take a magnetic system, such as the ubiquitous Ising model, and quench it into the low temperate phase starting from a random initial condition. In the course of time neighbouring spins realize that their preferred configuration is an ordered one and domains of the two ordered phases form and grow. At any finite time the configuration is such that both types of domains exist.

The domain growth kinetics in systems undergoing an ordering process after a quench is an important problem for material science applications but also for our understanding of pattern formation in nonequilibrium systems. The late stage dynamics is believed to be governed by a few properties of the systems whereas material details should be irrelevant. Among these relevant properties one may expect to find the number of degenerate gound states, the nature of the conservation laws and the hardness or softness of the domain walls. Thus, classes akin to the universality classes of critical phenomena have been identified.

If the ordering is governed by an order-parameter field $\phi(\vec{,} t)$ the physically relevant observable is the order-parameter correlation function

$$\langle \phi(\vec{r}, t)\phi(\vec{r}', t')\rangle \tag{7.1}$$

For simplicity, we consider a scalar field ϕ but cases in which vectorial or even tensorial order parameters are also of experimental relevance. The times t and t'

The simplest way to determine the domain growth properties of pure models is to use computer simulations of lattice models. The average domain size is determined using several criteria, the most common ones being:

- The 'inverse perimeter density'

  $$R(t) = \frac{1}{U(t) - U_{eq}} \;, \tag{7.2}$$

  where $U(t)$ is the energy at time t and U_{eq} is the equilibrium energy.

- The pair correlation function $C(\vec{r}, t)$ LOOK AT AMAR ET AL.

- $$R(t) = \langle N^{-1} (\sum_{i=1}^{N} s_i)^2 \rangle^{1/d} \tag{7.3}$$

In order to avoid finite-size effects, a rule-of-thumb states that the growth has to be stopped when R reaches 0.4 L, with L the linear size of the system. Another limitation is given by the fact that the true asymptotic behaviour may be veiled by crossover effects.

In pure and isotropic systems the domain growth is characterized by a power law

$$R(t) = At^{\phi} \tag{7.4}$$

with ϕ the *growth exponent.*

The averaged radious of the magnetic domains in clean ferromagnetic models has been determined using computer simulation studies of Ising (and Potts) models and with field-theoretical Langevin-like effective equations. One finds the so-called *Lifshitz-Allen-Cahn* growth law

$$R(t) = At^{1/2} \tag{7.5}$$

with A a weakly temperature dependent coefficient, independently of the number of equilibrium states (Ising or Potts models). The domain walls are sharp. The domains and their surface are compact (*i.e.* they have dimension d and $d-1$, respectively).

For systems with continuous variables, such as rotors or XY models, and no conserved order parameter, a number of computer simulations have shown that the growth law is

$$R(t) \sim t^{1/4} \; . \tag{7.6}$$

These models support the formation of wider domain walls. This result was hardly debated during some years, since several authors claimed that (7.6) was just a crossover towards the asymptotic regime (7.5), at least at non-zero temperature.

Another question one may would be interested in is characterizing the distribution of the sizes of these domains and its evolution. This is known in $d = 1$ but much less can be said about the higher dimensional problem.

A different type of dynamics occurs in the case of phase separation (the water and oil mixture). In this case, the material is locally conserved, *i.e.* water does not transform into oil but they just separate. Determining the growth and geometrical properties of the domains is already much harder in this case. After some discussion, it was established, as late as in the early 90s, that for systems with *conserved order parameter* as the example at hand, the growth is given by

$$R(t) \sim t^{1/3} \; . \tag{7.7}$$

In the late stages of the coarsening process the spherically averaged structure factor $S(k,t)$ that can be measured experimentally with small-angle scattering of neutrons, x-rays or light, has been found to satisfy scaling:

$$S(k,t) = N^{-1} \langle \left| \sum_{i=1}^{N} s_i e^{i\vec{k}\vec{r}} \right|^2 \rangle \sim R^d(t) F(kR(t)) \tag{7.8}$$

and correspondingly the real space correlation function is expected to behave as

$$C(r,t) = F(r/R(t)) \tag{7.9}$$

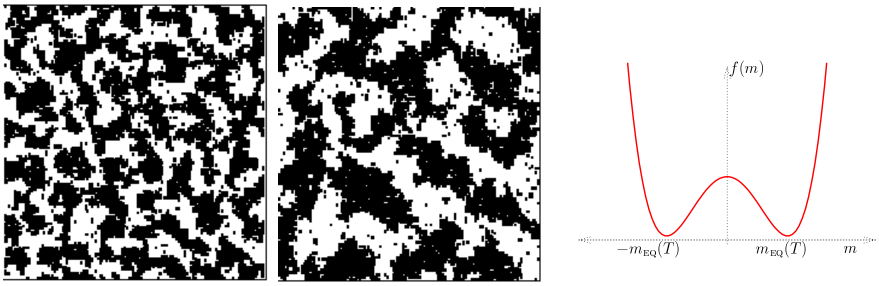


Figure 9: Time evolution of a spin configuration; two snapshots of a $2d$ slice in a $3d$ Ising model on a cubic lattice at $t_w = 1000$, $t_w = 10000$ MC steps in a simulation.

The situation becomes much less clear when there is quenched disorder in the form of non-magnetic impurities in a magnetic sample, lattice dislocations, residual stress, etc. Qualitatively, the dynamics is expected to be slower than in the pure cases. Indeed, disorder generates metastable states that trap the system in its evolution and thus render its relaxation slower. Determining the precise growth law becomes a difficult task.

Already for the random ferromagnet, for example, there is no consensus about the growth law. *** RIEGER GREG ***

In the $3d$ RFIM the curvature-driven growth mechanism that leads to (7.5) is impeded by the random field roughening of the domain walls. Indeeed, one observes that after a quench to low temperature the spins rapidly coalesce and form small domains, these domains expand and compact at the expense of their smaller neighbours but their growth is partially stopped by the random fields that pin the interfaces. Much longer time scales are needed to surmount the (free) energy barriers introduced by the local fluctuations in the fields, and eventually reach the long-range order. Comparing to the pure Ising model one notices that the initial growth follows a very similar time-dependence in the two cases but the subsequent coarsening is much slower in the presence of random fields. The precise behaviour of the growth law depends on time, temperature and the strength of the random field. In the early stages of growth, one expects the zero-field result to hold with a reduction in the amplitude

$$R(t) \sim (A - Bh^2)\, t^{1/2} \; . \tag{7.10}$$

The time-window over which this law is observed numerically is smaller, the larger the field strength. In the late time regime, where pinning is effective Villain deduced a logarithmic growth

$$R(t) \sim \frac{T}{h^2} \ln t \tag{7.11}$$

by estimating the maximum barrier height encountered by the domain wall and using the Arrhenius law to derive the associated time-scale.

In the case of spin-glasses, if the mean-field picture with a large number of equilibrium states is realized in finite dimensional models, the dynamics would be one in which all these states grow in competition. If, instead, the droplet model applies, there would be two types of domains growing and this phenomenological theory predicts that these will have an average radious growing as

$$R(t) = (\ln t)^{1/\psi} \; , \tag{7.12}$$

with the exponent ψ satisfying $0 \leq \psi \leq d-1$. Some refined arguments that we shall not discuss here indicate that the dimension of the bulk of these domains should be compact but their surface should be a rough with fractal dimension $d_s > d-1$.

Note that if one starts such a dynamics in a very large ($L \gg a$) system with no biasing magnetic field the system will never magnetize in finite times with respect to L. More explicitly, if the growth law is the power law (7.5) one needs times of the order of L^2 to grow a domain of the size of the system. For any shorter time, domains of the two types exist and the system is *out of equilibrium*. This systems constitutes a first example of a problem with *slow dynamics*. Whether all systems with slow dynamics, in particular structural and spin glasses, undergo some kind of sumple though slow domain growth is an open question.

7.2.1 Damage spreading

A useful method to isolate the effect of initial conditions on the subsequent evolution of a system, in paticular one undergoing coarsening dynamics, is *damage spreading*. The idea is simple. Take two copies of a system and at, say, time t^*, introducing damage into one of them, both systems are then evolved using idential dynamic rules (same thermal noise, same random numbers in a Montecarlo simulation). Any subsequent difference between the two copies is due to the difference in the configuration at t^*. A measure of damage spreading is the Hamming distance between the two configurations.

7.3 Phase space dynamics

It is commonly done in the study of disordered systems or, better said, in the applications of disordered systems and ideas in other areas as biology, economy, etc. to discuss the dynamics in phase space. Adopting a mean-field-like viewpoint in which disordered systems are characterized by many metastable states of all types, one pictures the dynamics of the system as the wandering of a representative point in a phase space characterized by a rough (free)-energy landscape. This viewpoint can be very useful if correctly used but it can also be misleading.

Indeed, phase space if a highly dimensional (2^N in a spin system) space and one can only *grasp* three dimensional geomtries. To show this problem, let us come back to a magnetic systems that is quenched to its low temperature phase. The initial, high temperature configuration, is not magnetized. It corresponds to the unstable maximum of the free-energy density $f(m)$. Naively, one would imagine that in time the point that sits on top of this barrier very quickly falls to one of the two mimima corresponding to $\pm m$, the equilibrium values. Instead, we know that this happens in a rather slow time scale and for all times $t < L^2$ the system is not magnetized and just sits on top of the maximum.

Moreover, in more complicated systems, as mean-field spin-glass models, the relevant free-energy density is not a function of just one parameter but N of them, the N local magnetizations. The dynamics is not just given by jumps from one minimum to another and saddle-points take an unexpected importance that might go beyond these solvable but unrealistic models.

7.4 Surface growth

Another interesting problem consists in characterizing the growth of a surface on a substrate due to material deposition combined (or not) with transverse diffusion of the material that reaches the surface.

7.4.1 Solid-on-solid models

These models are microscopic; they represent a trully experimental situation, such as atom deposition as in film growth by molecular beam epitaxy, and they are also advantageous to do numerical simulations.

A substrate d dimensional surface of size L^2 is divided into cells that can be occupied by columns of falling particles. Particles fall on this substrate and stick to it according to different rules that define different models.

In *ballistic deposition* particles are released from a randomly chosen position above the surface, and they follow a vertical trajectory and they simply stick to the nearest neighbour encountered.

In *random deposition* the particles fall vertically until they reach the top of their column.

In *random deposition with surface relaxation*, after reaching the top of its column the particle can further diffuse along the surface until it finds the position with the lowest height.

*** Depositing a particle at site $\vec{r}$ increases the height by one unity. In a time-step few particles reach the surface and they can then move to nearest neighbours sites subject to the condition that its new height is not higher than the original one. One translates the numerical time-scale into an experimental time-scale in which $d = 2$ as follows: the deposition of L^2 particles per Montecarlo time unit corresponds to the growth of one layer per second for a system with a lattice constant os 10øA and a flux of 10^{14}cm^2 s. ***

7.4.2 Continuous models

Continuous models often describe the surfaces at larger length-scales. A coarsening process is employed in such a way that the surface can be described by a continuous function.

The simplest model for the growth and rearrangement of a surface is due to Edwards and Wilkinson who showed that the continuum limit of the process of sedimentation of granular particles under gravity on an initial flat substrate and their further diffusion on the surface leads to

$$\frac{\partial h(\vec{x},t)}{\partial t} = \nu \nabla^2 h(\vec{x},t) + \xi(\vec{x},t) \ , \tag{7.13}$$

where $\vec{x}$ is a d-dimensional spatial vector denoting position on the substrate, t is time and h is a scalar function taking real values and measuring the height with respect to its average value. The last term is a thermal noise, typically chosen to have a Gaussian probability distribution a zero average and correlations

$$\langle\, \xi(\vec{x},t)\xi(\vec{x}',t')\,\rangle = 2T\delta(\vec{x}-\vec{x}')\delta(t-t') \ . \tag{7.14}$$

Equation (7.13) is a stochastic field equation. The first term on the RHS of eqn (7.13) penalizes rapid variations of the surface in space and ν is then a measure of the surface tension. The noise term describes the randomness in the deposition process. Note that the Edwards-Wilkinson equation can in fact be seen as describing the equilibrium fluctuations of an interface (for example liquid vapor), with the noise term describing the thermal fluctuations. The stochastic equation (7.13) is linear and can be easily solved using a Fourier transform of the space coordinate $\vec{x}$.

When the interface is driven out-of-equilibrium by an external force (for example a non zero average flux of particles impinging on the surface, or a difference in chemical potential between the two sides of the interface), then a new, non linear term of the form $[\nabla h(\vec{x},t)]^2$ should be added. This leads to the well-studied *Kardar-Parisi-Zhang equation* with many applications in various fields.

7.5 Scaling

The kinetics of a surface is characterised by the width of the surface that is defined as the mean-squared displacement of the total height and depends on time and the linear size of the substrate, L. Asymptotically, this quantity has a scaling limit given by:

$$W(L,t) \equiv \langle\, (h(\vec{x},t) - \langle\, h(\vec{x},t)\,\rangle)^2\,\rangle^{1/2} = L^\alpha \mathcal{F}(tL^z)\ , \tag{7.15}$$

where the roughness exponent α, the dynamical exponent z and the scaling function $\mathcal{F}$ are defined by the above equation in the large L limit. The average denotes either an average over thermal histories or an average over space, $\langle\, \cdot\,\rangle = A_L^{-1}\sum_{\vec{x}} \cdot$ with the area of the substracte of linear dimension L indicated by A_L. The scaling function $\mathcal{F}$ is such that

$$W(L,t) \sim \begin{cases} t^{1/2}\ , & \text{for short times} \\ t^{\alpha/z}\ , & \text{for intermediate times} \\ L^\alpha\ , & \text{for very long times} \end{cases} \tag{7.16}$$

The exponents α and z characterize the universality class of the growth problem.

FIGURE

Recently, there has been growing interest in characterising the complete dynamic probability distribution $P_L(W^2,t)$ since it has been suggested that it might be used to define universality classes for non-equilibrium steady states *via* the scaling function:

$$\langle W^2\rangle_{L\to\infty} P_L(W^2,t) = \Phi\left(\frac{W^2}{\langle W^2\rangle_{L\to\infty}}\right)\ .$$

Note the similarity between this problem and the one of the fluctuations of the global magnetization in the critical $2d$ XY model that we discussed in Sect. 4.8.5. Indeed, the two problems are intimately connected since the dynamics of the $2d$ XY model in the spin-wave approximation is given by eq. (7.13).

The above definition leads to what can be called a "global" characterisation of the roughness of the interface, averaged over the whole surface. Another, more local, characterisation of the roughness of the surface, is defined as

$$\Delta(\vec{x},\vec{x}',t,t') \equiv \langle\, (h(\vec{x},t) - h(\vec{x}',t')\,)^2\rangle = |\vec{x}-\vec{x}'|^\alpha\, \mathcal{F}\left(|t-t'||\vec{x}-\vec{x}'|^z\right)\ , \tag{7.17}$$

where we are now looking at the limit $\Lambda^{-1} \ll |\vec{x}-\vec{x}'| \ll L$. Often, the exponents and the scaling function defined globally over the whole surface [as in eq. (7.15)] coincide with their local counterpart, but this is not always the case. Note that in (7.17) one measures at two different times t and t'.

A Probability and statistics

We are used to reading the word "statistics" in many contexts of science but also everyday life. The word statistics has its origin in the Latin "status" as well as the word "state", suggesting that governments used statistical concepts since long ago. In a few words, statistics is the theory that allows one to make sense out of a list of numbers; it is the branch of mathematics that deals with the analysis of data. One identifies two sub-branches:

1. *Descriptive statistics.* The goal is to obtain useful information from a series of *raw data* that is typically too large to deal with directly. For instance, present experiments in molecular biology present the "difficulty" of yielding too much raw data that need processing before becoming useful. Descriptive statistics is a set of tools, or mathematical manipulations of the raw data, that convert them into a few numbers and plots that are easy to understand.

2. *Inferential statistics.* The goal is to obtain useful information about a very large *population* being able to test only a *sample*, that is to say, a small portion of the total population. The typical example of the application of inferential statistics are election polls. Clearly, one is not able to ask every citizen for which candidate he or she is going to vote. However, one can ask a sample of the population and infer from the result which is going to be the global one. One of the main difficulties in this case is related to the choice of the sample. Clearly, if the sample is taken exclusively from a posh neighbourhood the result will be different from the one obtained from a poor one. Similarly, if only aged people are consulted, the result might be different from the one obtained using a sample of young people.

Probability Theory is at the basis of Statistics. It allows to identify generic behaviour of *random events.* The mathematical theory of Statistics developed mostly in England, during the beginning of the XXth century. It was also the time when it became clear that a purely deterministic description of physics was not feasible with the development of Statistical Mechanics and Quantum Mechanics. [Interestingly enough, many developments in the Theory of Statistics were the consequence of research in Agriculture (the analysis of the effect of fertilizers!) by Sir R. Fisher.]

Let us here recall the definition and main properties of probabilities. If one tosses a coin it is practically impossible to predict if it will fall heads or tails (try to solve Newton's equations for the coin and air...) However, repeating the experiment *many times* one can predict general patterns of the head-tail sequence. This is a general feature of *random events*: one is unable to predict the outcome of a single experiment but one can predict some general features of a sequence of them. The same feature applies to determine the position and velocity of a particle in a gaz, and so many other physical systems involving many agents in interaction.

1.1 Frequency and probability

If we toss an unbiased coin we expect all sequences of heads (H) and tails (T) to be equally likely. Why? This is a very important *assumption*! How can we justify it?

Imagine that we toss the same coin a very large number of times, $n \to \infty$, we count the number of H and Ts that we obtain, and we construct the *probability or large N limit frequency*:

$$P(H) \equiv \lim_{n\to\infty} f_n(\#\text{H}) = \lim_{n\to\infty} \frac{\#\text{H}s}{n} \; . \tag{1.1}$$

$$P(T) \equiv \lim_{n\to\infty} f_n(\#\text{T}) = \lim_{n\to\infty} \left(\frac{n - \#\text{H}s}{n}\right) = 1 - P(\text{H}) \; . \tag{1.2}$$

This definition quantifies the subjective idea 'what do we expect to get after tossing a coin?' If the coin is not biased, if n is sufficiently large, we should get one half of Hs and one half of Ts. Thus, for $n \to \infty$:

$$p \equiv P(H) = \frac{1}{2} = 1 - P(T) \qquad q \equiv P(T) = \frac{1}{2} \; . \tag{1.3}$$

If, instead, the coin is biased, H (or T) will appear more frequently. Then $p > \frac{1}{2}$ and $q < \frac{1}{2}$ or *viceversa.*

1.2 "Ideal experimental' definition of probability

Based on the toin coss example it is natural to define in general the *probability of an event* as the value taken by the frequency in the limit $n \to \infty$:

$$P(\mathcal{E}) \equiv \lim_{n\to\infty} f_n(\mathcal{E}) \tag{1.4}$$

where $\mathcal{E}$ denotes the event we are interested in, e.g. getting 3Hs after tossing $n \to \infty$ coins.

Let us mention that one is used to listen to weather reports in which people talk about the "probability of having rain tomorrow". In this case, the definition of probability as a limiting procedure is much less clear and becomes much more subjective. We stick to objective cases here where the above definition can be safely applied.

But, it is actually not necessary to go to such extreme example to realize that sometimes the definition above may be ambiguous. This problem has been discussed in the past. Kolmogorov gave a rigurous definition of probability that we shall not discuss here. We shall simply say that it is based on using the list of properties that we discuss below as a definition of probability.

1.3 Properties

We now list a series of properties of probabilities. Some of them are obvious and follow simply from the frequency-based definition given above. Some other are not and constitute new definitions.

Semi-positive definite. It is clear that the probability of an event is a quantity that can take only positive or zero values:

$$P(\mathcal{E}) \geq 0 \; . \tag{1.5}$$

Bound. It is also clear from its definition that the probability of an event is bounded by one:

$$P(\mathcal{E}) \leq 1 \; . \tag{1.6}$$

Normalization. Note that the sum of the number of occurrences of each event is equal to the total number of possible event. Thus, the sum of the frequencies over the events is normalised to one, and the sum of probabilities is also normalised to one:

$$\sum_{\mathcal{E}} P(\mathcal{E}) = \sum_{\mathcal{E}} \lim_{n\to\infty} f_n(\mathcal{E}) = 1 \tag{1.7}$$

(assuming that one can exchange sum and limit).

Probability of the complementary event. If the probability of occurrence of an event, $\mathcal{E}$, is $P(\mathcal{E})$, the probability of non-occurrence of the same event (its *complementary event*, $\overline{\mathcal{E}}$) is

$$P(\overline{\mathcal{E}}) = 1 - P(\mathcal{E}) \; . \tag{1.8}$$

This is just a consequence of the normalization of the probabilities.

For example, when we introduced the coin toss problem before we said that when tossing a normal coin one finds H one half of the times and T the other half of the times. In toher words, we *assumed* that the probability of getting H in a toss is 1/2 and the probability of finding the complemetary event T is $1 - 1/2 = 1/2$ too.

Addition principle. For any pair of events $\mathcal{E}_1$ and $\mathcal{E}_2$,

$$P(\mathcal{E}_1 \bigvee \mathcal{E}_2) = P(\mathcal{E}_1) + P(\mathcal{E}_2) - P(\mathcal{E}_1 \bigwedge \mathcal{E}_2) \; . \tag{1.9}$$

The symbol $\bigvee$ represents the logical (not exclusive) "or" while the symbol $\bigwedge$ represents the logical "and". The meaning of the logical or is that the event $\mathcal{E}_1 \bigvee \mathcal{E}_2$ is true whenever $\mathcal{E}_1$ is true, $\mathcal{E}_2$ is true or both $\mathcal{E}_1$ and $\mathcal{E}_2$ are true. The meaning of the logical and is that the event $\mathcal{E}_1 \bigwedge \mathcal{E}_2$ is true only if $\mathcal{E}_1$ and $\mathcal{E}_2$ are true.

Let us illustrate this property with an example. Imagine that one is playing with two dices, a red and a blue one (to make them distinguishable). What is the probability for getting 1 with the red one or 2 with blue one?

This problem is represented mathematically as follows. The event $\mathcal{E}_1$ is getting 1 with the red dice. The event $\mathcal{E}_2$ is getting 2 with the blue dice. There are $6\times6 = 36$ possible outcomes of throwing the two dices, i.e. getting the pairs $(1,1); (1,2); \ldots; (2,1); (2,2); \ldots; (6,6)$ where the first element in each pair is the result of the red dice and the second one is the result of the blue dice. The event $\mathcal{E}_1 \bigvee \mathcal{E}_2$ is true in the cases $(1,1);(1,2);(1,3);(1,4);(1,5);(1,6);$ $(2,2);(3,2);(4,2);(5,2);(6,2)$, so there are 11 true realizations and, since the dices are not biased, after $n \to \infty$ repetitions of the experiment "throwing the two dices" we expact to find 11/36 successful tries. Thus, the probability of this event is $P(\mathcal{E}_1 \bigvee \mathcal{E}_2) = 11/36$.

Now, the number of positive outcomes for the event $\mathcal{E}_1$ is just 6. The same applies to the event $\mathcal{E}_2$. The sum of these two numbers is 12 and it does not coincide with result above. The reason why is that the event $(1,2)$, that solves the problem, is counted twice. Once in the number of positive outcomes for $\mathcal{E}_1$ and once in the number of positive outcomes for $\mathcal{E}_2$. One needs to correct this double counting and this is why there is the last term in (1.9). Taking this into account, one finds that the right-hand-side of this equation predicts $P(\mathcal{E}_1 \bigvee \mathcal{E}_2) = (6 + 6 - 1)/36 = 11/36$ which is the correct result.

Particular case: mutually exclusive events. When the events $\mathcal{E}_1$ and $\mathcal{E}_2$ are mutually exclusive, $\mathcal{E}_1 \bigwedge \mathcal{E}_2 = 0$ and the formula (1.9) simplifies to

$$P(\mathcal{E}_1 \bigvee \mathcal{E}_2) = P(\mathcal{E}_1) + P(\mathcal{E}_2) \; . \tag{1.10}$$

Working again with dices, mutually exclusive events are 'getting 1 and 2 with one dice after one throw. There are plenty of other examples of this sort.

Joint probability. One calls in this way the probability of simultaneous occurrence of a number of events,

$$P(\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n) \; . \tag{1.11}$$

The comma means here the same as the symbol $\bigwedge$ in (1.9).

Independence principle. If two events can occur independently of the realisation of the other, the probability of the simultaneous occurrence of them is simply the product of the individual probabilities:

$$P(\mathcal{E}_1, \mathcal{E}_2) = P(\mathcal{E}_1) P(\mathcal{E}_2) \; . \tag{1.12}$$

The generalisation to n independent events is straightforward:

$$P(\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n) = P(\mathcal{E}_1) P(\mathcal{E}_2) \dots P(\mathcal{E}_n) \; . \tag{1.13}$$

As an example one can imagine playing with dice and cards and asking about events that are associated to the dice and to the cards, independently.

Conditional probability. The probability of the occurrence of an event event $\mathcal{E}_1$ conditioned to the occurrence of another event $\mathcal{E}_2$ is

$$P(\mathcal{E}_1 | \mathcal{E}_2) = \frac{P(\mathcal{E}_1, \mathcal{E}_2)}{P(\mathcal{E}_2)} \; . \tag{1.14}$$

Clearly, if $\mathcal{E}_1$ and $\mathcal{E}_2$ are independent events, $P(\mathcal{E}_1 | \mathcal{E}_2) = P(\mathcal{E}_1)$

In the TDs we shall see plenty of examples that illustrate these properties.

1.4 Discrete and continuous random variables

A careful experimentalist performs a measurement many times in identical conditions, then calculates the frequency of each result and from them, in the limit of a very large number of measurements (that in real life are never infinite!) *estimates* the probability of each result. We call *random variable* the result of the experiment.

For instance, in the coin toss problem the random variable, let us call it x, takes only two possible values, H and T. We can associate these two exclusive results with the numbers 0 and 1 and then call the random variable x *bimodal.* Bimodal random variables are *discrete* since they can only take values on a discrete set $(0, 1$ in this case).

Other random variables can take values on continuous sets, as the real numbers, and are hence called *continuous* random variables. Imagine that we look at the absolute value of the velocity v of a particle moving within a gas. In principle, this value can be any real number (most probably bounded...) and hence v is a continuous random variable.

When dealing with continuous random variables we have to be more precise about what we mean by the probability of an event. In this case, the quantity that is well defined is the probability for the event to take values within a given interval. In the case of the velocity of a particle in the gas, we ask what is the probability that the particle has an absolute velocity comprised between v_1 and $v_1 + \Delta v$. In the limit in which Δv is infinitesimally small ($\Delta v \to dv$) this allows us to define the *probability density* as the probability for the random variable to take values within the infinitesimal interval of length dv starting at v_1:

$$p(v)dv \tag{1.15}$$

We use lower cases to denote probability densities and upper cases to indicate probabilities.

Next we define quantities that characterise the behaviour of random variables.

Probability distribution. The set of probabilities associated to a random variable is called a *probability distribution*. The probability distributions may be represented by tables but it is far more convenient to represent them with formulæor diagrams that generalize the frequency plot in Fig. **??**-right. These are just plots with $P(x)$ in the y-axis and x in the x-axis.

Cumulative probability. With the probability distribution of a discrete random variable one constructs a cumulative probability that is simply the probability for the random variable to take a value that is larger than some chosen one:

$$F(x_1) = \sum_{x \geq x_1} P(x) \; . \tag{1.16}$$

Similarly, for a continuous random variable,

$$F(x_1) = \int_{x_1}^{\infty} dx \; p(x) \; . \tag{1.17}$$

Expected value, mean or average[9] The expected values of a discrete and a continuous random variable are

$$E(x) = \sum_x x P(x) \; , \qquad E(x) = \int dx \; p(x) x \; , \tag{1.18}$$

respectively.

Median. It is the “middle” value:

$$x_{median} \quad \equiv \quad \int_{-\infty}^{x_{median}} dx \, p(x) = \int_{x_{median}}^{\infty} dx \, p(x) \tag{1.19}$$

The word median originates in the Latin *medius*.

Mode. It is the most probable value:

$$x_{mode} : \qquad \left\{ \left. \frac{dp(x)}{dx} \right|_{x_{mode}} = 0 \; , \; \left. \frac{d^2p(x)}{dx^2} \right|_{x_{mode}} > 0 \; . \right. \tag{1.20}$$

The mode is not necessarily unique.

Variance. The variance of a discrete and a continuous random variable are defined as

$$\sigma_x^2 \equiv \sum_x (x - E(x))^2 P(x) \qquad \sigma_x^2 \equiv \int dx \; p(x)(x - E(x))^2 \; , \tag{1.21}$$

respectively. It measures the spreading of the pdf, *i.e.* the variance is small if all data are concentrated around the mean and it is large otherwise.

The variance is well-suited to distinguish rare events.

Standard deviation. It is just given by

$$\sigma_x \equiv \sqrt{\sigma_x^2} \; ; \tag{1.22}$$

since it is measured in the same units as the average it is hence directly comparable to it.

Relative variability. It is a (adimensional) comparison between the standard deviation and the average:

$$rv_x \equiv \frac{\sigma_x}{\mu_x} \; . \tag{1.23}$$

It is relevant to compare the spread to the average of the data.

Momenta The k-th momentum of a random variable is defined as

$$\mu^{(k)} \equiv \sum_x x^k P(x) \qquad\qquad \mu^{(k)} \equiv \int dx \; x^k \, p(x) \; . \tag{1.24}$$

Indeed, knowing all momenta one can reconstruct the functional form of P or p.

Correlation. The correlation quantifies the similarity between two (or more) random variables. There are many possible definitions of correlation and each may be better adapted to some problem.

Let us take two random variables x and y described by the jopint probability distrobution $p(x, y)$. For example, x and y may represent two different observables of a physical system.

The correlation between the two sets X and Y is then defined as

$$C_{xy} = \frac{1}{\sigma_x \sigma_y} \int dx \int dy p(x, y)(x - \mu_x)(y - \mu_y) \tag{1.25}$$

with the averages μ_x and μ_y, and the standard deviations, σ_x and σ_y defined above. C_{xy} takes values between -1 (complete decorrelation) and 1 (complete correlation).

Another example can be drawn from physics: a magnetic system modelled by what is called the Ising model. Within this model the magnetic system is represented by spins (see the Quantum Mechanics course!) on a lattice. Each spin is a little vector that can only point up and down and is hence represented by a variable s that takes values ± 1 (for up and down, respectively). Each spin is labelled by an index i that represents the site it occupies on the lattice. If the lattice is cubic and d dimensional and one then has $n = L^d$ spins in the system, with L the linear length of the d dimensional cube. the set we want to study is then $X = \{s_1, \dots, s_n\}$. The magnetized state is represented by a configuration such that the magnetization density, $m = n^{-1} \sum_{i=1}^n s_i$, takes a non-zero value while the paramagnetic state is such that $m = 0$. Note that the magnetization density m is just the average of the set X, $m = \mu_x$. Now, one can imagine that the configuration evolves in time, meaning that each spin changes its configuration as time evolves. And one may be interested in comparing the configurations of the total system at different times. This is achieved by computing the correlation (1.25) where the X is the set of n spin values at one time, say t_1, and the Y is the set of n spin values at another time, say t_2. In other words, $X = \{s_1(t_1), s_2(t_1), \dots, s_n(t_1)\}$, $Y = \{s_1(t_2), s_2(t_2), \dots, s_n(t_2)\}$ and

$$C_{xy} = \frac{1}{\sigma(t_1)\sigma(t_2)} \sum_{i=1}^{n} (s_i(t_1) - m(t_1))(s_i(t_2) - m(t_2)) \; . \tag{1.26}$$

Covariance. The numerator in (1.25) is called the covariance of x and y.

1.5 Probability distributions

In this Section we present some probability distributions that appear recurrently in the study of physical, biological, and sociological systems among others.

The binomial

A simple question one can pose now is: how many Hs do we expect after a finite number n of tosses? The binomial (or Bernoulli) distribution gives an asnwer to this question. The probability of getting h Hs after n tosses of the coin is

$$P_n(h) = \left(\begin{array}{c} n \\ h \end{array} \right) p^h (1-p)^{n-h} \; ,$$

with p the probability of getting an H is one toss, $1-p$ the probability of getting a T in one toss, and the combinatorial number defined as

$$\left(\begin{array}{c} n \\ h \end{array} \right) \equiv \frac{n!}{(n-h)!h!} \qquad n! = n(n-1)\dots 1 \; .$$

A simply way of understanding this result is the following.

The form of the bimodal distribution and its evolution with n is shown in Figs. 10 ($p = q = 1/2$) and 13($p \neq q$). Note that when n incrases the figures look more and more peaked about their maximum that occurs at $\approx n/2$. This statement can be proven rigorously.

With a simple calculation one proves

$$E(h) = np \; , \qquad \sigma_h^2 = np(1-p) \tag{1.27}$$

(see Problem Set 1). Note that $rv_h = \sigma_h / E(h) \propto 1/\sqrt{n}$ and this tends to 0 when $n \to \infty$. This means that the distributions look more and more smooth and peaked when n increases.

The bimodal distribution is a 'two-parameter' one. n and p are the parameters controlling its form.

To summarize, we have just studied a problem, the tossing of a coin and we have derived an equation that gives us the probability of occurrence of an event (h Hs) in a series of n experiments. This formula is valid, obviously, only for this problem (and others that can be mapped onto it). Even if it is just an example, it is illuminating and it allows us to define the concept of probability.

A typical physical realization of the toss coin problem is the random walk problem. Take a one dimensional lattice with spacing a and a drunken walker that can occupy the

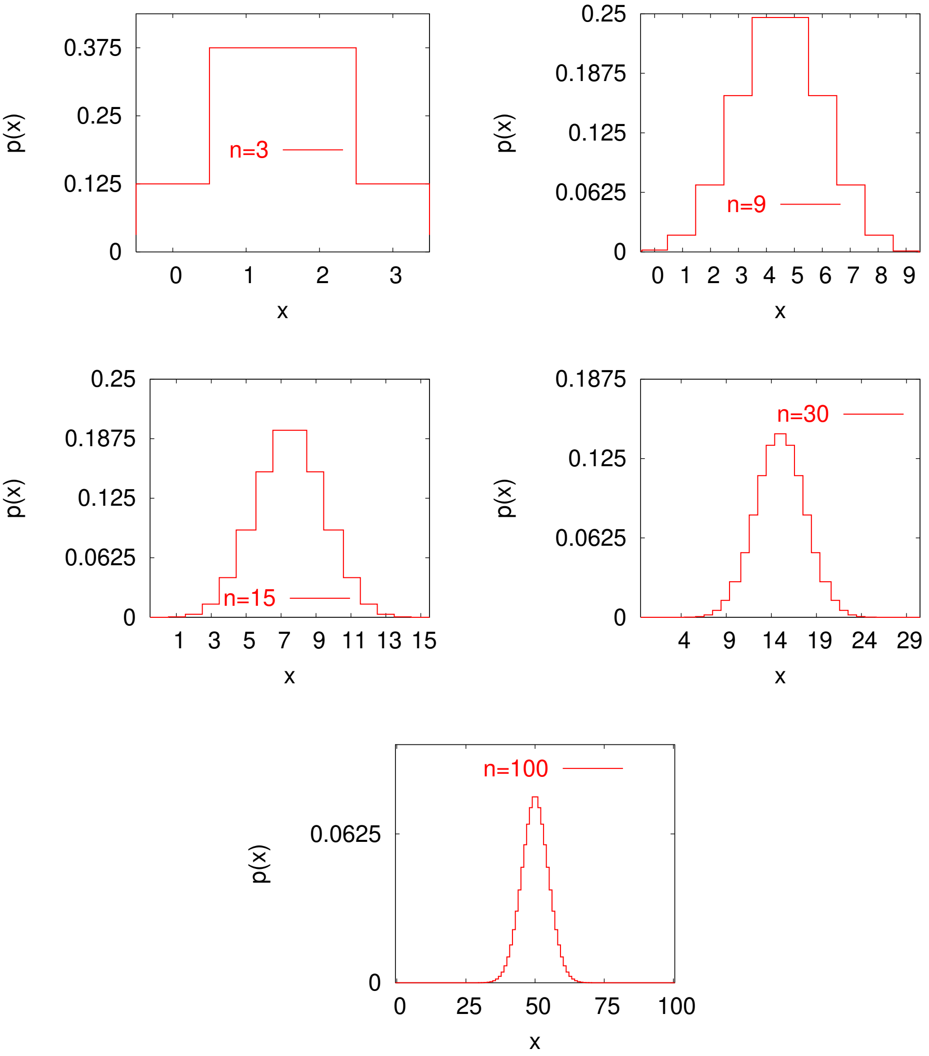


Figure 10: Evolution with n of the binomial distribution function. Note how the form is more and more continuous as n increases. One sees from these plots that the location of the maximum is at $\approx n/2$ while the width of the curve at half height is $\approx 2\sigma_h \approx \sqrt{n}$.

sites on this lattice. One knows that in each time step the walker moves right one half of the times and it moves left the other half. After n time steps, the walker might have taken h steps to the right and $n-h$ steps to the left. $P_n(h)$ represents the probability of this composed event. This rather simple problem has many application in physics and biology and it is at the basis of the theory of diffusion.

In general, the bimodal disctribution applies to problems that can be cast in the form of a yes or no answer. For example, given a population of n patients one can wonder what is the probability that h among them have asthma if one knows that the probability of each patient having asthma is p. This is given by $P_n(h)$ with parameters n and p.

The Poisson distribution

The distribution

$$P(x) = \frac{\mu^x}{x!} e^{-\mu} \,, \tag{1.28}$$

with μ a parameter and x a discrete random variable, taking values $x = 0, 1, 2, \ldots$, is called Poisson probability distribution function (PDF). See Fig. 11.

With a simple calculation one proves that the parameter μ is the expected value of the random variable, $E(x) = \mu$ and $\sigma_x^2 = \mu$.

Poisson PDF can be obtained as the limit of the binomial distribution when n is very large, p is small and np is kept fixed.

Two examples of random variables described by the Poisson distribution are the following.

Take a gas confined to a volume V with average density ρ. If one divides the total volume into small boxes with equal volume v, the local density fluctuations from box to box. The number of particles in each little volume v is described by a Poisson distribution.

Another example is the one of the local connectivity in the so-called Erdos-Renyi random graph with average connectivity c. The number of links reaching a vertex fluctuate according to a Poisson distribution with parameter $\mu = c$.

The normal or Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \;, \tag{1.29}$$

where x takes all values on the real axis and μ and σ^2 are two parameters, represents the Gaussian probability density. One easily checks, by direct integration, that the expected value of x, $E(x)$, equals μ and its variance, σ_x^2, equals σ^2. See Fig. 12.

The Gaussian probability density can also be viewed as a limit of the binomial distribution. In this case one keeps p finite and takes n to infinity.

The Gaussian distribution in its *normal form* has zero mean and unit variance. Given a generic Gaussian distribution, the normal form is achieved by defining $y = (x - \mu)/\sigma$ and transforming $p(x)$ into $p(y)$:

$$p(y) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{y^2}{2}} \;. \tag{1.30}$$

The cumulative probability of a random variable with a Gaussian distribution are tabulated or can be computed numerically. They are of great use in Inferential Statistics.

Approximating the binomial

The calculation of the Poisson and Gaussian distributions is much simpler than the calculation of the binomial. Of course, these distributions are not identical.

A rule of thumb tells that the Poissonian approximation to the binomial is rather accurate when

$$n \geq 20 \qquad p \leq 0.05 \;. \tag{1.31}$$

Let us check this statement with an example. Take a binomial distribution with $n = 70$ and $p = 0.02$. The probability of three successes is

$$P(h = 3, n = 70) = \frac{70!}{67!\,3!}\, 0.02^3 (0.98)^{67} = 0.1151 \;. \tag{1.32}$$

while the Poissonian approximation is

$$P(\mu = np = 1.4, x = 3) = \frac{1.4^3}{3!}\, e^{-1.4} = 0.1128 \;, \tag{1.33}$$

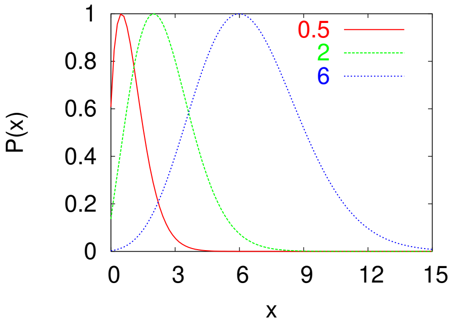


Figure 11: Poisson distribution for three values of μ given in the key.

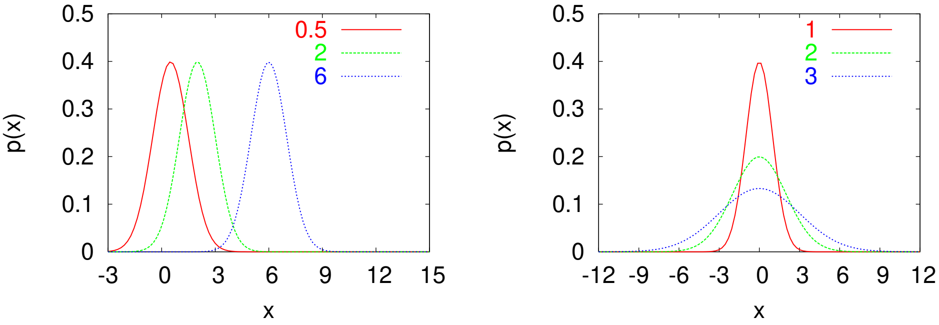


Figure 12: Gaussian probability density. Left: fixed $\sigma^2 = 1$ and three values of μ given in the key. Right: fixed $\mu = 0$ and three values of σ given in the key.

a very good result.

The binomial distribution is discrete while the Gaussian is continuous. To approximate the former by the latter one has to be precise what one really means. Thus, the binomial probability of having h will be approaximated by the Gaussian cumulative probability of finding the continuous variable x within the interval $[h - 0.5, h + 0.5]$:

$$P(h, n) \approx \int_{h-0.6}^{h+0.5} dx\, p(x) \ . \tag{1.34}$$

Let us test this hypothesis with an example. Suppose that 15% of the cars coming out of an assembly plant have some defect. In a delivery of 40 cars what is the probability that exactly five cars hace defects? The actual answer is given by the binomial formula

$$\frac{40!}{35!5!}\, 0.15^5 0.85^{35} = 0.1692 \ . \tag{1.35}$$

But the calculation of the factorials is quite complicated. What about the Gaussian approximation to this result? The mean and average are of the binomial $\mu = np = 40 \times 0.15 = 6$ and $\sigma = \sqrt{np(1-p)} = \sqrt{40 \times 0.15 \times 0.85} = 2258$. Thus, the Gaussian approximation is

$$\int_{4.5}^{5.5} dx \frac{1}{\sqrt{2\pi}2258}\, e^{-\frac{-(x-6)^2}{2\times 2258^2}} = 0.1583 \tag{1.36}$$

and, again, this result is quite close to the exact one. Note that in this case, $p = 0.15$ is not

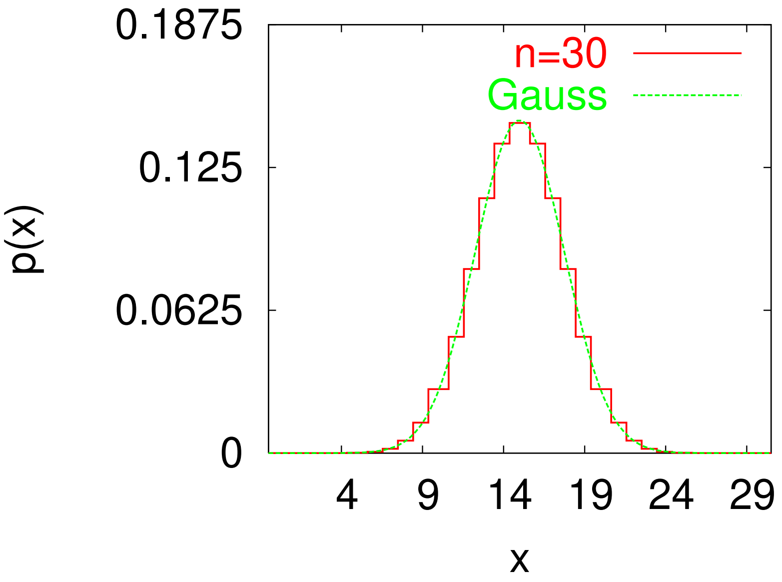


Figure 13: The binomial compared to the Gaussian.

small (it is not smaller than 0.05 as in the above example where we used the Poissonian approximation).

Lévy distribution

Lévy distributions play an important rôl ein the description of a number of statistical problems of current interest (finance, anomalous diffusion, etc). They are pdf's that falls off at infinity as a power law:

$$L_\mu(x) \sim \frac{\mu\, A_\pm}{|x|^{1+\mu}}\ , \qquad |x| \to \infty\ . \tag{1.37}$$

μ is a positve parameter (to ensure the normalization to one of the pdf). The n-th moment diverges as soon as $n \geq \mu$. In particular, if $0 < \mu \leq 2$ then the variance diverges.

1.6 The law of large numbers

We have already mentioned that if we toss a coin a large number of times, the number of Hs found will be close to $n/2$. This result is a particular form of the *law of large numbers*.

Take a random variable x and measure it n times, i.e. draw the numbers $x_1, x_2, \ldots, x_n$. Construct the average of these numbers, as done in Sect. 3:

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i\ . \tag{1.38}$$

The law of large numbers tell us that the probability for this value to be different from the expected value of x, $E(x)$, tends to zero when $n \to \infty$:

$$P(|x-\mu_x|) > \epsilon) \to_{n\to\infty} 0\ . \tag{1.39}$$

This statement can be rigorously proven. We shall not discuss the proof here. It can be found in the literature.

1.7 The central limit theorem

The *central limit theorem* states that in the large n limit the random variable constructed with the sum of independent, identically distributed random variables,

$$y \equiv \frac{1}{n}\sum_{i=1}^{n} x_i \ , \tag{1.40}$$

has a Gaussian distribution with expectation value equal to the expectation value of the original random variable, $E(x)$, and variance given by σ_x^2/n. In other words, y is a continuous random variable (even if x might have been a discrete one) distributed according to

$$p(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}}\, e^{-\frac{(y-E(y))^2}{2\sigma_y^2}} = \sqrt{\frac{n}{2\pi\sigma_x^2}}\, e^{-\frac{n(y-E(x))^2}{2\sigma_x^2}} \ . \tag{1.41}$$

Since y is a sum over n random variables, divided by n, its expectation value is

$$E(y) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1} E(x) = E(x) \ , \tag{1.42}$$

i.e. it is equal to the expectation value of the variables x_i that are supposed to be equally distributed.

What about the variance of the random variable y? We have shown that the variance of a random variable z built with the sum of two random variables, x and y, is equal to the sum of the variances, $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$, $z = x + y$. Now, in our case, each element in the sum is equal to x/n. And we also showed that the variance of x/n, $\sigma_{x/n}^2$, equals the variance of x divided by n^2: $\sigma_{x/n}^2 = \sigma_x^2/n^2$. Hence,

$$\sigma_y^2 = (\sigma_{x/n}^2 + \ldots + \sigma_{x/n}^2) = n\sigma_{x/n}^2 = n\sigma_x^2/n^2 = \sigma_x^2/n \tag{1.43}$$

Note the importance of this theorem for developing experiments. An experiment consists of the measurement of an observable. A single measurement does not make sense since the result found will be subject to many sources of noise. The result of an experiment only has a statistical sense. If one repeats the measurement n times and constructs the average of the results, $x_1, \ldots, x_n$, the central limit theorem ensures that the y will be a normal distributed random variable with expected value $E(x)$ and variance σ_x^2/n. Incresing the number n one then reduces the width of the Gaussian and for sufficiently large n one is sure to approach the actual expected value of the observable x with the average of the data, y. We shall come back to this issue below.

Generalization in which the variables x_i are not equally distributed (but have finite expectation value and variance) are possible but we shall not discuss them here.

In the case in which the variables to be added are still iid but have a probability distribution with a tail that falls as a power law with $\mu < 2$, as in the Lévy case with $A_{\pm}^{\mu}$, the limit law *is* a symmetric Lévy law (if the right and left tails of $p(x)$ have different amplitudes one obtains an asymmetric limit law with amplitude that are a combination of the original ones).

2 Some useful formulæ

Stirling formula for the factorial of a large number reads:

$$\ln N! \sim N \ln N - \ln N \;, \qquad \text{for} \qquad N \gg 1 \;. \tag{2.1}$$

3 The saddle-point method

Imagine one has to compute the following integral

$$I \equiv \int_a^b dx \; e^{-Nf(x)} \;, \tag{3.1}$$

with $f(x)$ a positive definite function in the interval $[a, b]$, in the limit $N \to \infty$. It is clear that due to the rapid exponential decay of the integrand, the integral will be dominated by the minimum of the function f in the interval. Assuming there is only one absolute minimum, x_0, one then Taylor expands $f(x)$ upto second order

$$f(x) \sim f(x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 \tag{3.2}$$

and obtains

$$I \sim e^{-Nf(x_0)} \int_a^b dx \; e^{-N\frac{1}{2}f''(x_0)(x-x_0)^2} = e^{-Nf(x_0)} [Nf''(x_0)]^{-1/2} \int_{y_a}^{y_b} dy \; e^{-\frac{1}{2}(y-y_0)^2} \;, \tag{3.3}$$

with $y_0 \equiv \sqrt{Nf''(x_0)} x_0$ and similarly for y_a and y_b. The Gaussian integral is just an error function that one can find in Tables.

This argument can be extended to multidimensional integrals, cases in which there is no absolute minimum within the integration interval, cases in which the function f is not positive definite, etc.

4 the Fokker-Planck equation

Let us show how to obtain the above Fokker-Planck equation from the Langevin equation. The following derivation encompasses the multiplicative noise case; this allows one to see how the difference between the Itô convention and the Stratonovich convention, discussed in Sect. **??**, translates onto the evolution of the probability density. For simplicity, we will only consider here the one dimensional Smoluchowski case.

Suppose first that one works with the Itô convention. Then, an arbitrary function $G(x)$ of the random variable x, and that does not depend explicitly on time, evolves according to

$$\begin{aligned} dG &= \frac{\partial G}{\partial x} dx + e^2(x) \frac{T}{\gamma} \frac{\partial^2 G}{\partial x^2} dt \\ &= \frac{\partial G}{\partial x} \frac{1}{\gamma} [F(x) + e(x)\xi(t)] dt + e^2(x) \frac{T}{\gamma} \frac{\partial^2 G}{\partial x^2} dt \;, \end{aligned} \tag{4.4}$$

where $e(x)$ is the function multiplying the noise term ξ in the Langevin equation (*cfr.* eqn **??**). Averaging the above equation over the thermal noise leads to

$$\langle G(x(t+dt)) - G(x(t)) \rangle = \frac{1}{\gamma}\left\langle \frac{\partial G}{\partial x}F(x)\right\rangle + \frac{T}{\gamma}\left\langle e^2(x)\frac{\partial^2 G}{\partial x^2}\right\rangle \ , \tag{4.5}$$

where we have used the fact that $x(t)$ and $\xi(t)$ are independent in the Itô convention. Now, as an identity, any average $\langle O(x) \rangle$ of a function $O(x)$ can be written in terms of the probability density $P(x,t)$ as

$$\langle O(x) \rangle = \int O(x)P(x,t)\, dx \ . \tag{4.6}$$

Hence, eqn (4.5) is equivalent to

$$\int dx\, G(x)\frac{\partial P(x,t)}{\partial t} \ = \ \frac{1}{\gamma}\int dx\, \frac{\partial G}{\partial x}F(x)P(x,t) + \frac{T}{\gamma}\int dx\, e^2(x)\frac{\partial^2 G}{\partial x^2}P(x,t) \ .$$

After integrating by parts and noticing that the resulting equation holds for an arbitrary function $G(x)$, one finally finds:

$$\gamma\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\, F(x)P(x,t)\,] + T\frac{\partial^2}{\partial x^2}[\, e^2(x)P(x,t)\,] \ , \tag{4.7}$$

which is identical to the Smoluchowski equation (6.16) with the identification $F(x) = -dV/dx$ and $e(x) = 1$.

If we now consider the Stratonovich prescription, the starting point is simply

$$dG = \frac{\partial G}{\partial x}dx = \frac{\partial G}{\partial x}\frac{1}{\gamma}[F(x) + e(x)\xi(t)]dt \ , \tag{4.8}$$

but now one cannot assume that $\langle O(x(t))\xi(t) \rangle = 0$ when averaging the above equation over the noise ξ. Rather, one has, using Novikov's formula in a discrete time setting:

$$\begin{aligned} \langle O(x(t))\xi(t) \rangle \ &= \ \sum_{l \le k}\left\langle \frac{\partial O(x_k)}{\partial \xi_l}\right\rangle \langle \xi_k \xi_l \rangle \\ &= \ \sum_{l \le k}\left\langle O'(x_k)\frac{1}{2\gamma}[e(x_k)\delta_{k,l} + e(x_{k-1})\delta_{k-1,l}]\right\rangle 2\gamma T \delta_{k,l} \ . \end{aligned} \tag{4.9}$$

Therefore, taking $O(x) = e(x)\partial G/\partial x$, one finally obtains

$$\left\langle e(x)\frac{\partial G}{\partial x}\xi(t)\right\rangle = T\left\langle e(x)\frac{\partial}{\partial x}\left[e(x)\frac{\partial G}{\partial x}\right]\right\rangle \ . \tag{4.10}$$

A manipulation to the above allows then one to establish the following evolution equation for the probability density:

$$\gamma\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\, F(x)P(x,t)\,] + T\frac{\partial}{\partial x}\left\{e(x)\frac{\partial}{\partial x}[e(x)P(x,t)]\right\} \ . \tag{4.11}$$

Note the difference with the above equation established in the Itô framework: the function $e(x)$ appears differently in the diffusion term. It can easily be checked that this equation is equivalent to the Itô case provided the true force $F(x)$ is replaced by an effective force given by $F(x) - Te(x)e'(x)$. Hence, the different conventions lead in general to different results, because they describe different microscopic processes. The two Fokker-Planck equations become the same in the additive noise case $e(x) = 1$.

Mécanique Statistique TD 1 :
Notions de base

Le but de ce TD est de faire un rappel des notions de base en probabilité of mécanique statistique.

1 Probabilité

Soit la fonction $f(x) = axe^{-\alpha x}$, définie pour des valeurs de x réels et positifs.

1. Que sont les conditions que les paramètres a et α doivent satisfaire pour que cette fonction soit une densité de probabilité ?
2. Calculer sa moyenne – en introduisant une source – et l'emplacement de son maximum. Sont ces deux valeurs les mêmes ?
3. Que peut-on dire sur les propriètés de symétrie de cette densité de probabilité ?
4. Calculer l'écart-type, σ.
5. Relier la valeur d'expectation de x et de $x + \lambda$ avec λ une constante.
6. Relier la valeur d'expectation de x et de λx avec λ une constante.

2 Paramagnetisme

Soit un ensemble de N spins, $s_1 = 1, \ldots, N$, en présence d'un champ magnétique externe. L'énergie de l'ensemble est donnée par

$$E = -h \sum_{i=1} s_i \ . \tag{1}$$

L'ensemble est en contact avec un reservoir à température T.

1. Calculer la fonction de partition et l'énergie libre par spin.
2. Calculer la densité d'aimantation du système.
3. Calcule la susceptibilité magnétique.
4. Identifier un paramètre naturel pour ce système. Tracer la densité d'énergie libre, la densité d'aimantation du système, et sa susceptibilité magnétique en fonction de ce paramètre. Discuter.

3 Théorème de fluctuation - dissipation

Prenons un système caractérisé par une énergie E en équilibre thermique avec un reservoir à température T.

1. Montrer le théorème de fluctuation – dissipation

$$\left.\frac{\partial \langle A \rangle_h}{\partial h}\right|_{h=0} = \beta \langle (A - \langle A \rangle)^2 \rangle \;, \tag{2}$$

   où h est un champs qui se couple linéairement à A:

$$E \to E - hA \;. \tag{3}$$

   Ce résultat ne dépend pas du système consideré.

4 Modèle d'Ising en une dimension

On étudiera le modèle d'Ising où les spins, $s = \pm 1$ sont placés sur un anneau de longueur L (*i.e.* avec des conditions au bord periodiques), sans champ magnétique externe.

1. Calculer l'énergie libre. Aide: définir les variables de lien, $\eta_i = s_i s_{i+1}$.
2. Calculer la fonction de corrélation $G(\vec{r}_i - \vec{r}_j) \equiv \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$. Aide: calculer la fonction de corrélation $G(\vec{r}_i - \vec{r}_{i+1})$ comme la variation de l'énergie libre par rapport à la contante de couplage J_i entre les spins s_i et s_{i+1} (on fixera tous les $J_i = J$ à la fin du calcul).

Mécanique Statistique TD 2 :
Modèles avec des interactions à portée infinie

1 Le modèle d'Ising

On considère le modèle d'Ising avec des interactions ferromagnétiques de portée infinie entre p spins : la constante de couplage $J_{i_1 \dots i_p}$ vaut $J_{i_1 \dots i_p} = J_0$ pour <u>touts</u> les groupes de p spins :

$$H = -J_0 \sum_{i_1 \dots i_p} s_{i_1} \dots s_{i_p} - h \sum_i s_i . \tag{1}$$

1. Comment doivent se comporter J_0 et h en fonction de N pour que le modèle soit bien défini dans la limite thermodynamique ?
2. Calculer l'énergie et l'entropie moyennes pour le cas cas $h = 0$ dans les limites $T = 0$ et $T \to \infty$. Que peut-on remarquer ?
3. Écrire la fonction de partition du système comme une somme sur les différentes valeurs de l'aimantation par spin m. Définir une densité énergie libre à m fixée.
4. En utilisant la méthode du col, trouver une équation pour l'aimantation par spin, m.
5. Dans le cas $p = 2$ comparer avec l'équation obtenue par l'approximation de Weiss (champ moyen) pour le modèle d'Ising avec interactions entre plus proche voisins sur un réseau carré de dimension d. Que peut-on alors conclure sur l'approximation de champ moyen pour le modèle d'Ising avec des interactions de portée infinie ?
6. Étudier graphiquement l'équation pour l'aimantation par spin m avec ou sans champ magnétique. Distinguer les cas $p = 2$ et $p > 2$.
7. Idem pour l'énergie libre par spin $f(m)$.
8. Dans le cas $p = 2$, étudier l'aimantation près de la transition et déterminer les exposants critiques β, γ et δ.

Mécanique Statistique TD 3 : Systèmes désordonnés

1 Frustration

1. Considerer trois spins, s_1, s_2 et s_3, avec fonction d'énergie

$$E = -s_1 s_2 - s_1 s_3 - s_2 s_3 \ . \tag{1}$$

Calculer la fonction de partition, Z, l'aimantation moyenne de chaque spin, $m_i = \langle\, s_i \,\rangle$, avec $i = 1, 2, 3$, ainsi que la valeur moyenne de l'énergie, $\langle\, E \,\rangle$. Calculer l'énergie et l'entropie moyennes à température nulle.

2. Repeter ces calculs pour l'énergie

$$E = -s_1 s_2 - s_1 s_3 + s_2 s_3 \ . \tag{2}$$

3. Que constatez-vous ?

2 Désordre

1. Considerer deux spins s_1 et s_2 couplés selon

$$E = -J s_1 s_2 \ . \tag{3}$$

Calculer la fonction de partition, l'aimanation moyenne de chaque spin ainsi que l'énergie moyenne en fonction de J.

2. Prenons maintenant J aléatoire distribué selon une densité de probabilité continue, $p(J) = 1/(J_+ - J_-)$ pour $J_- \leq J \leq J_+$ et $p(J) = 0$ autrement. Calculer la moyenne sur le désordre des aimantations: $[m_1] = [\langle\, s_1 \,\rangle]$ et $[m_2] = [\langle\, s_2 \,\rangle]$ et de l'énergie, $[\langle\, E \,\rangle]$. Quel est la valeur de l'énegie moyenne de l'état fondamental ?

3 Propriété d'auto-moyennement

1. Considerez une chaine de spins avec des couplages aléatoires :

$$E = -\sum_{i=1}^{N} J_i s_i s_{i+1} \ . \tag{4}$$

2. Calculer la fonction de partition Z_J.

3. Calculer l'aimantation locale moyenne $m_i = \langle\, s_i \,\rangle$ pour une réalisation donnée du désordre. Est-elle invariante par rapport aux traslations ?

4. Calculer l'énergie libre F_J. Quelle est la distribution de probabilité de F_J ?
5. Montrer que F_J est une quantité auto-moyennante.
6. Est Z_J auto-moyennante ?
7. Calculer la fonction de corrélation $\langle s_i s_j \rangle$ et discuter ses propriétés.

4 Le magnet de Mattis

Montrer que le modèle de Mattis n'est pas vraiment frustré.

Mécanique Statistique TD 4 :
Dynamique stochastique

Nous allons travailler, par souci de simplicité, en dimension $d = 1$. La généralisation au cas d dimensionnel est immédiate. Dans cette section nous étudierons quelques équations de Langevin pour le mouvement d'une particule sur un axe sous l'effet d'un bain thermique.

1 Transformation de l'équation de Langevin

La forme générique de l'équation de Langevin est

$$\gamma \frac{dx}{dt} = F + \xi \, , \tag{1}$$

où ξ répresente un bruit thermique qu'on prend, typiquement, Gaussien et blanc :

$$\langle\, \xi(t) \,\rangle = 0 \, , \qquad \langle\, \xi(t)\xi(t') \,\rangle = 2k_B T \gamma \, \delta(t - t') \, , \tag{2}$$

pout tous les temps t et t'. F est la force totale appliquée sur la particule.

Montrer qu'avec une transformation du temps, $\tau \equiv g(t)$, l'équation dévient

$$\frac{dx}{dt} = F + \eta \, , \tag{3}$$

avec η un bruit blanc Gaussienne avec moyenne et corrélation :

$$\langle\, \eta(\tau) \,\rangle = 0 \, , \qquad \langle\, \eta(\tau)\eta(\tau') \,\rangle = 2k_B T \, \delta(\tau - \tau') \, . \tag{4}$$

Trouver $g(t)$ et η.

2 Le mouvement Brownien forcé

Dans plusieurs application d'intérêt on 'tire' de la particule avec un force constante; un exemple est l'électrophorèse où on fait avancer les particules chargées (typiquement de lignes) sous l'effet d'un champ électrique.

1. Trouver la solution de l'équation stochastique

$$\frac{dx}{dt} = \xi + f \tag{5}$$

ou ξ est un bruit blanc et f est une force qu'on prendra constante, $f(t) = f$. On prend comme condition initielle $x(0) = x_0$.

2. Calculer la position moyenne de la particule $\langle\, x(t) \,\rangle$. Discuter la perte de mémoire de la condition initielle.

3. Calculer le déplacement moyen quadratique $\Delta^2(t,t') \equiv \langle\, (x(t) - x(t'))^2 \,\rangle$.

   Avez vous trouvé un résulat stationnaire ? Dans quelle limite ?

   Comparer le résultat à la forme diffusive normale. Peut-on identifier une constante de diffusion ?

4. Calculer la fonction de corrélation connexe $\langle\, (x(t) - \langle\, x(t)\,\rangle)(x(t') - \langle\, x(t')\,\rangle)\,\rangle$.

5. Calculer la réponse linéaire à une force infinitesimel $h(t')$, $\delta\langle x(t)\,\rangle/\delta h(t')|_{h=0}$.

6. Comparer les deux résultats.

2.1 L'oscillateur harmonique

On considère une particule en contacte avec un bruit blanc et sous l'effet d'une force harmonique:

$$\frac{dx}{dt} = -kx + \xi \tag{6}$$

1. Calculer la position moyenne $\langle\, x(t)\,\rangle$.

2. Calculer la fonction de corrélation connexe $\langle\, (x(t) - \langle\, x(t)\,\rangle)(x(t') - \langle\, x(t')\,\rangle)\,\rangle$.

3. Calculer le déplacement moyen quadratique $\Delta^2(t,t') \equiv \langle\, (x(t) - x(t'))^2 \,\rangle$.

   Avez vous trouvé un résulat stationnaire ?

   Comparer le résultat à la forme diffusive normale. Peut-on identifier une constante de diffusion ? Discuter la limite $k \to 0$.

4. Calculer la réponse linéaire à une force infinitesimel $h(t')$, $\delta\langle x(t)\,\rangle/\delta h(t')|_{h=0}$.

5. Comparer la réponse linéaire à la derivée temporelle de la fonction de corrélation. Sont-elles réliées par le théorème de fluctuation-dissipation ?

   Que se passe-t-il dans la limite $k \to 0$. Discuter.

Mécanique Statistique TD 5 :
Phénomènes de croissance de surfaces

References

[1] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill Series in Fundamentals of Physics).

[2] K. Huang, *Statistical mechanics*, (J. Wiley & sons, 2nd ed., 1987).

[3] L. E. Reichl, *A modern course in statistical physics* (E. Arnold publ., 1980).

[4] W. Pauli, *Pauli lectures on physics*, Volume 4 Statistical Mechanics (MIT Press, 1973).

[5] S. E. Luria, M. Delbrück, *Mutations of bacteria from virus sensitivity to virus resistance*, *Genetics* **28**, 491 (1943). This article is available online at www.esp.org

[6] W. Feller, *An introduction to probability theory and its applications*, Wiley, 1971.

[7] J-P Bouchaud, M. Potters *Theory of financial risk and Derivative Pricing : From Statistical Physics to Risk Management.*

[8] H. E. Stanley, *Introduction to phase transitions and critical phenomena* (Oxford University Press, New York, 1971).

[9] N. Goldenfeld, *Lectures on phase transitions and the renormalization group* (Addison-Wesley, 1992).

[10] J. Cardy, *Scaling and renormalization in Statistical Physics*, Cambridge Lecture notes in physics (Cambridge University Press, 1996).

[11] D. J. Amit, *Field theory, the renormalization group and critical phenomena*(World Scientific, Singapore, 1984).

[12] G. Parisi, *Statistical field theory* (Addison-Wesley, 1988).

[13] H. J. Maris, L. P. Kadanoff, *Teaching the renormalization group*, Am. J. Phys. **46**, 652 (1978).

[14] A. Gabrielli, F. Sylos-Labini, M. Joyce, L. Pietronero *Statistical physics for cosmi structures* (Springer Verlag, 2004).

[15] G. Fletcher, *A mechanical analog of first- and second-order phase transitions*, Am. J. Phys. 65, 74 (1997). E. Guyon, *Second-order phase transitions: Models and analogies*, Am. J. Phys. 43, 877 (1975). R. V. Mancuso, *A working mechanical model for first- and second-order phase transitions and the cusp catastrophe*, Am. J. Phys. 68, 271 (2000).

[16] M. Mézard, G. Parisi, M. A. Virasoro, *Spin-glass theory and beyond* (Springer-Verlag, 1987).

[17] *Spin Glasses* (Cambridge Studies in Magnetism, 1993).

[18] O. Martin, R. Monasson, R. Zecchina, *Statistical mechanics methods and phase transitions in optimization problems*, http://xxx.lanl.gov/cond-mat/0104428.

[19] Vik. Dotsenko, *Introduction to the Replica Theory of Disordered Statistical Systems* (Cambridge University Press)

[20] Vik. Dotsenko, *Introduction to the Theory of Spin Glasses and Neural Networks* (World Scientific, 1995).

[21] R. da Silveira, An introduction to breakdown phenomena in disordered systems, http://xxx.lanl.gov/cond-mat/0008174. Am. J. Phys.

[22] C. W. Gardiner, *Handbook of Stochastic Methods*.

[23] N.G. Van Kampen *Stochastic Processes in Physics and Chemistry* (North-Holland Personal Library).

[24] A-L Barabasi, H. E. Stanley, *Fractal concepts in surface growth*, (Cambridge University Press, 1995).

[25] J-P Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).

[26] L. F. Cugliandolo, *Dynamics of glassy systems* Lecture notes in "Slow Relaxation and non equilibrium dynamics in condensed matter", Les Houches Session 77 July 2002, J-L Barrat, J Dalibard, J Kurchan, M V Feigel'man eds., cond-mat/0210312

[27] F. Portelli, *FLuctuations des grandeurs globales dans les systèmes corrélés*, ENS-Lyon thesis under the advise of P. W. Holdsworth.