Master ICFP 2020-2021

Lecture Notes on Advanced Statistical Physics and its New Applications

by Giulio BIROLI and Gregory SCHEHR

typed by Anthony GUILLEN and Louis GINABAT

The drawing on the front page, adapted from R. Wolstenholme, represents a flock of birds in the shape of a bird.

Contents

1 Thermal noise, coupling to the environment and irreversibility

The dynamics of isolated systems is well-described by NEWTON's or SCHRÖDINGER's equations. However, in general, systems are not isolated: they are coupled to their environment. As such, they can also be considered sub-systems of a much larger system, or equivalently as few degrees of freedom of a system which contains many. If the system as a whole can be considered isolated, the study of its sub-system yields effective dynamical equations. This notion of coupling to the environment was introduced by Albert EINSTEIN, when he initiated the theory of Brownlan motion; that is the erratic motion of pollen grains in a fluid carefully described by Robert BROWN in 1827, as shown in this figure:

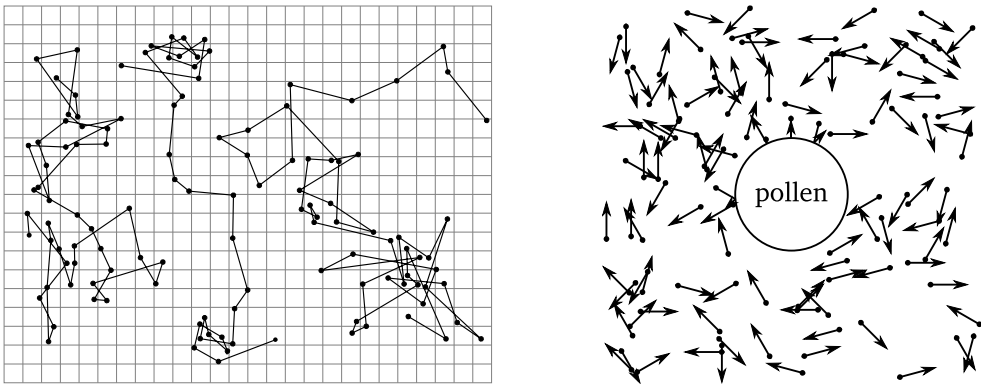

Figure 1: reproduced from *Les Atomes*, three trajectories of colloidal particles are displayed, and successive positions are joined by segments (left); grain of pollen in a fluid that undergoes collisions with the particles of the fluid (right).

This phenomenon found its first quantitative explanation in EINSTEIN's work, which proved the existence of atoms. In 1908, LANGEVIN proposed the equation bearing his name allowing the analytical study of the Brownian motion.

1.1 Inductive way *à la* EINSTEIN

Let us now consider a pollen grain in a fluid.

1.1.1 Dissipation

In presence of a viscous drag, NEWTON's equation reads $m\dot{\boldsymbol{v}} + \gamma \boldsymbol{v} = \boldsymbol{F}$, with $\boldsymbol{v}$ the velocity of the pollen grain, m its mass, $\boldsymbol{F}$ any external force acting on it and γ the viscous drag coefficient. If $\boldsymbol{F} = \boldsymbol{0}$:

$$\boldsymbol{v}(t) = \boldsymbol{v}_0 \exp\left(-\frac{\gamma}{m}t\right) \xrightarrow[t\to+\infty]{} \boldsymbol{0} \tag{1.1}$$

this is a problem, because if thermodynamics is right, from equipartition of energy:

$$\frac{1}{2}m\langle v^2\rangle = \frac{3}{2}k_B T \implies \langle v^2\rangle = \frac{3k_B T}{m} > 0 \tag{1.2}$$

1.1.2 Thermal noise *à la* LANGEVIN

LANGEVIN added a force in NEWTON's equation, to account for the erratic collisions of the particles of the fluid with the pollen grain (see figure 1 right):

$$m\dot{\boldsymbol{v}} + \gamma \boldsymbol{v} = \boldsymbol{F} + \boldsymbol{\xi}(t) \tag{1.3}$$

The term $\boldsymbol{\xi}(t)$ is called the thermal noise because of its relation to the temperature, which will become clear. It is a random Gaussian function, it can be seen as a collection of random variables:

$$\langle \xi_t^i \rangle = 0 \quad \text{and} \quad \langle \xi_t^i \xi_{t'}^j \rangle = B\delta_{i,j}\frac{\delta_{t,t'}}{\Delta} \tag{1.4}$$

where Δ is some time step. In the continuous[1], these relations become very simply:

$$\langle \xi^i(t) \rangle = 0 \quad \text{and} \quad \langle \xi^i(t)\xi^j(t') \rangle = B\delta_{i,j}\delta(t-t') \tag{1.5}$$

To fullfill EINSTEIN's argument, we want equation (1.2) to hold when $\boldsymbol{F} = \boldsymbol{0}$. We have:

$$\boldsymbol{v}(t) = \boldsymbol{v}(-\infty)\exp\left(-\frac{\gamma}{m}t\right) + \int_{-\infty}^{t}\exp\left(-\frac{\gamma}{m}(t-t')\right)\frac{\boldsymbol{\xi}(t')}{m}dt' \tag{1.6}$$

when $t \to +\infty$, the exponential multiplying $\boldsymbol{v}(-\infty)$ vanishes, and we compute the mean square of $\boldsymbol{v}$:

$$\langle \boldsymbol{v}(t)^2 \rangle = \left\langle \iint_{-\infty}^{t} \exp\left(-\frac{\gamma}{m}(t-t_1)\right)\exp\left(-\frac{\gamma}{m}(t-t_2)\right)\frac{\boldsymbol{\xi}(t_1)}{m}\cdot\frac{\boldsymbol{\xi}(t_2)}{m}dt_1dt_2 \right\rangle \tag{1.7}$$

$$= \frac{1}{m^2}\iint_{-\infty}^{t} \exp\left(-\frac{\gamma}{m}(t-t_1)\right)\exp\left(-\frac{\gamma}{m}(t-t_2)\right)\langle \boldsymbol{\xi}(t_1)\cdot\boldsymbol{\xi}(t_2)\rangle\, dt_1dt_2 \tag{1.8}$$

since the noise is isotropic:

$$\langle \boldsymbol{\xi}(t_1)\cdot\boldsymbol{\xi}(t_2)\rangle = 3\langle \xi^x(t_1)\xi^x(t_2)\rangle = 3B\delta(t_1-t_2) \tag{1.9}$$

Therefore we obtain that:

$$\langle \boldsymbol{v}(t)^2 \rangle = \frac{3B}{m^2}\int_{-\infty}^{t}\exp\left(-2\frac{\gamma}{m}(t-t_1)\right)dt_1 = \frac{3B}{2m\gamma} \tag{1.10}$$

For equation (1.2) to hold, B has to verify this relation, which is due to EINSTEIN [2]:

$$\boxed{B = 2\gamma k_B T} \tag{1.11}$$

By an analogous computation of $\langle (\boldsymbol{x}(t))^2 \rangle = 6Dt$ (setting $\boldsymbol{x}(t=0) = \boldsymbol{0}$), and using $\gamma = 6\pi\eta r$ for the viscous drag on a sphere of radius r, we obtain:

$$\boxed{D = \frac{k_B T}{6\pi\eta r} = \frac{RT}{6\mathcal{N}_A\pi\eta r}} \tag{1.12}$$

In a series of experiments conducted during the years 1907-1909, Jean PERRIN obtained several important results.

$$\langle x^2 \rangle \propto t, \quad D \propto \frac{1}{r}, \quad \text{and a measurement of the AVOGADRO number } \mathcal{N}_A \tag{1.13}$$

these definitely settled the debate on whether the matter was continuous or discrete, by proving the existence of atoms. This work earned him the NOBEL prize in 1926.

[1] the δ distribution is dimensionfull, since $\int f(t')\delta(t-t')dt' \sim \sum_{t'} f_{t'}(\delta_{t,t'}/\Delta)\Delta$

[2] EINSTEIN worked with some kind of FOKKER-PLANCK equation that describes diffusion processes, as will be seen later on in this chapter, introducing a field of random noise and a diffusion coefficient.

1.1.3 Discussion about LANGEVIN's equation

Let us consider a one-dimensional movement in a harmonic potential from which derives a force $F = -V'(x)$. The motion of a massive particle in this potential is pictured in figure 2, adding successively dissipation and noise.

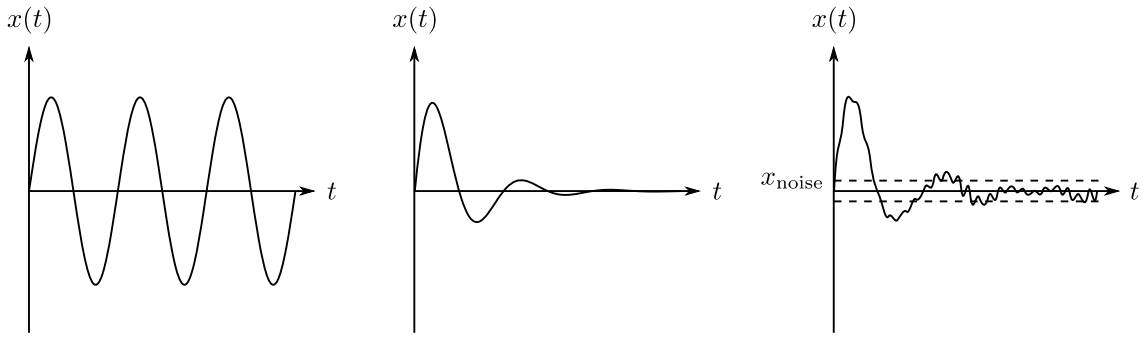

Figure 2: particle in a harmonic potential (left); with additionnal dissipation (middle); and noise (right).

In the precedent figure, x_{noise} is defined by $V(x_{\text{noise}}) = k_B T$. Stochastic dynamics are used in computer science to analyze coupling with degrees of freedom. The Brownian motion $\dot{x} = \xi(t)$, is very important in probability theory.

Exercise Consider $V(x) = x^2/2$ and solve everything. You will eventually obtain BOLTZMANN-GIBBS probability.

$$\mathbb{P}(x,t) = \frac{1}{Z} \exp\left(-\frac{x^2}{2k_B T}\right) \tag{1.14}$$

1.2 Derivation of the LANGEVIN equation, deductive way *à la* ZWANZIG

Robert ZWANZIG was a theoretical physicist and chemist, who made important contributions to the statistical mechanics of irreversible processes. In particular, he devised a deductive way to derive LANGEVIN's equation.

Let us consider a one-dimensional system interacting with an environment. The Hamiltonian is then composed of three parts: the system H_S, the environment H_E and the interaction between the two H_I. We model the environment as N independent harmonic oscillators of mass 1, interacting with the system via the following couplings.[3]

$$\gamma_k = \frac{\tilde{\gamma}_k}{\sqrt{N}} \quad \text{with} \quad \tilde{\gamma}_k \underset{N\to+\infty}{\sim} \mathcal{O}(1) \tag{1.15}$$

To formalize, the Hamiltonian of the problem is $H = H_S + H_E + H_I$, with:

$$H_S = \frac{p^2}{2m} + V(x), \quad H_E = \sum_{k=1}^{N} \frac{1}{2} P_k^2 + \frac{\omega_k^2}{2} Q_k^2 \quad \text{and} \quad H_I = -\sum_{k=1}^{N} \gamma_k Q_k x \tag{1.16}$$

These harmonic oscillators coupled with the system account for the particles interacting with it. An analogy with phonons in a crystal provides a good intuition about this interaction[4], so let us give some reminders on phonons.

[3] the $1/\sqrt{N}$ in the coupling constant is important to make the sums converge when $N \to \infty$; physically one can argue that coupling of a system with a bath that contains many phonons modes is always weak

[4] the only difference is that here no spatial order is required

We can picture the problem as a chain of harmonic oscillators interacting with the system at $x = 0$. The Hamiltonian of this problem, without considering the system H_S, would be $H = H_E + H_I$, with [5]:

$$H_E = \sum_{l=-N/2+1}^{N/2+1} \frac{p_l^2}{2} + \frac{(q_l - q_{l+1})^2}{2} \quad \text{and} \quad H_I = -q_0 x \tag{1.17}$$

In FOURIER space:

$$Q_k = \frac{1}{\sqrt{N}} \sum_{l=1} \exp\left(-\frac{i2\pi kl}{N}\right) q_k \quad \text{and} \quad q_l = \frac{1}{\sqrt{N}} \sum_k \exp\left(\frac{i2\pi kl}{N}\right) Q_k \tag{1.18}$$

After calculation, we would recover the expressions in (1.16) for H_E and H_I, and we would get:

$$\omega_k^2 = 2(1 - \cos k) \quad \text{and} \quad \gamma_k = \frac{1}{\sqrt{N}} \tag{1.19}$$

then, using HAMILTON's equations, we obtain the following NEWTON's equations for x and Q_k:

$$m\frac{d^2x}{dt^2} = -V'(x) + \sum_k \gamma_k Q_k \quad \text{and} \quad \frac{d^2Q_k}{dt^2} = -\omega_k^2 Q_k + \gamma_k x \tag{1.20}$$

the solution for $Q_k(t)$ in (1.20) can be written as follows:

$$Q_k(t) = \frac{\gamma_k}{\omega_k^2} x(t) + \left(Q_k(0) - \frac{\gamma_k}{\omega_k^2} x(0)\right) \cos(\omega_k t) + \dot{Q}_k(0) \frac{\sin(\omega_k t)}{\omega_k} - \frac{\gamma_k}{\omega_k^2} \int_0^t \cos(\omega_k (t-s)) \dot{x}(s) ds \tag{1.21}$$

inserting this result into the equation for $x(t)$ gives:

$$m\frac{d^2x}{dt^2} + \sum_k \frac{\gamma_k^2}{\omega_k^2} \int_0^t \cos(\omega_k (t-s)) \dot{x}(s) ds \tag{1.22}$$

$$= -V'(x) + \sum_k \frac{\gamma_k^2}{\omega_k^2} x(t) + \sum_k \left[\dot{Q}_k(0) \frac{\gamma_k}{\omega_k} \sin(\omega_k t) + \gamma_k \left(Q_k(0) - \frac{\gamma_k}{\omega_k^2} x(0)\right) \cos(\omega_k t)\right] \tag{1.23}$$

this can be rewritten, without ambiguity on the definitions of the different terms:

$$\frac{d^2x}{dt^2} + \int_0^t K(t-s) \dot{x}(s) ds = -V'(x) - V_e'(x) + \xi(t) \tag{1.24}$$

We call $K(t-s)$ the retarder function, it enters into the dissipation term by acting as a memory term, taking the past into account. Actually, we will show that the integral in (1.24) is equal to $\gamma \dot{x}(t)$. The term noted $\xi(t)$ is a good candidate to be the thermal noise because it only depends on t. Let's understand why it is indeed the thermal noise.

Let us consider the environment at equilibrium, such that the distribution of its initial conditions $Q_k(0)$ and $\dot{Q}_k(0)$ is a BOLTZMANN-GIBBS distribution:

$$\mathbb{P} = \mathcal{N} \exp\left(-\frac{H}{k_B T}\right) \tag{1.25}$$

with $\mathcal{N}$ a normalization constant. To be more precise, we have:

$$\mathbb{P}[Q_k(0), \dot{Q}_k(0) | x(0)] = \mathcal{N} \prod_k \exp\left[-\beta(H_{E,k}(t=0) + H_{I,k}(t=0,x))\right] \tag{1.26}$$

$$= \mathcal{N} \prod_k \exp\left[-\frac{\dot{Q}_k^2(0)}{2k_BT} - \frac{\omega_k^2}{2k_BT} Q_k^2(0) + \frac{\gamma_k}{k_BT} Q_k(0) x(0)\right] \tag{1.27}$$

$$= \mathcal{N}' \prod_k \exp\left[-\frac{\dot{Q}_k^2(0)}{2k_BT} - \frac{\omega_k^2}{2k_BT} \left(Q_k(0) - \frac{\gamma_k}{\omega_k^2} x(0)\right)^2\right] \tag{1.28}$$

[5] in H_I we only keep the interacting part of $(q_0 - x)^2 = q_0^2 - 2q_0 x + x^2$

where $\dot{Q}_k(0)$ and $Q_k(0) - \gamma_k/\omega_k^2 x(0)$ are independent Gaussian variables. If the $\{X_k\}_k$ are independent Gaussian variables, we remind:

$$\langle X_k \rangle = 0, \quad \langle X_k^2 \rangle = \sigma_k^2, \quad \mathbb{P}(X_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{X_k^2}{2\sigma_k}\right) \tag{1.29}$$

moreover, $Y = a_k X_k$ is a Gaussian variable, with $\langle Y \rangle = 0$ and $\langle Y^2 \rangle = a_k^2 \sigma_k^2$ (using Einstein's convention).

This latter result shows that $\xi(t)$ is a Gaussian variable, so $\langle \xi(t) \rangle = 0$; also:

$$\langle \dot{Q}_k(0) \rangle = 0 \quad \langle \dot{Q}_k(0)^2 \rangle = k_B T \tag{1.30}$$

and we have:

$$\left\langle \left(Q_k - \frac{\gamma_k}{\omega_k^2} x(0) \right) \right\rangle = 0 \quad \left\langle \left(Q_k(0) - \frac{\gamma_k x(0)}{\omega_k^2} \right)^2 \right\rangle = \frac{k_B T}{\omega_k^2} \tag{1.31}$$

We compute:

$$\begin{aligned} \langle \xi(t)\xi(t') \rangle = \Bigg\langle \sum_k \left[\dot{Q}_k(0) \frac{\gamma_k}{\omega_k} \sin(\omega_k t) + \gamma_k \left(Q_k(0) - \frac{\gamma_k}{\omega_k^2} x(0) \right) \cos(\omega_k t) \right] \\ \times \sum_{k'} \left[\dot{Q}_{k'}(0) \frac{\gamma_{k'}}{\omega_{k'}} \sin(\omega_{k'} t) + \gamma_{k'} \left(Q_{k'}(0) - \frac{\gamma_{k'}}{\omega_{k'}^2} x(0) \right) \cos(\omega_{k'} t) \right] \Bigg\rangle \end{aligned} \tag{1.32}$$

here $\langle \dot{Q}_k(0) \dot{Q}_{k'}(0) \rangle = \delta_{k,k'} k_B T$, and for the two other terms:

$$\left\langle \left(Q_k(0) - \frac{\gamma_k x(0)}{\omega_k^2} \right) \left(Q_{k'}(0) - \frac{\gamma_{k'} x(0)}{\omega_{k'}^2} \right) \right\rangle = \delta_{k,k'} \frac{k_B T}{\omega_k^2} \quad \text{and} \quad \left\langle \dot{Q}_k(0) \left(Q_{k'}(0) - \frac{\gamma_{k'} x(0)}{\omega_{k'}^2} \right) \right\rangle = 0 \tag{1.33}$$

it finally gives:

$$\langle \xi(t)\xi(t') \rangle = \sum_k k_B T \frac{\gamma_k^2}{\omega_k^2} \sin(\omega_k t) \sin(\omega_{k'} t') + \sum_k k_B T \frac{\gamma_k^2}{\omega_k^2} \cos(\omega_k t) \cos(\omega_k t') \tag{1.34}$$

$$= \sum_k k_B T \frac{\gamma_k^2}{\omega_k^2} \cos(\omega_k (t - t')) = k_B T K(t - t') \tag{1.35}$$

this is EINSTEIN's relation between fluctuations $\langle \xi(t)\xi(t') \rangle$ and dissipation $K(t - t')$. If $\tau \ll \mathcal{O}(1)$, which is the case when $N \to +\infty$, then $K(t) \approx 2\gamma\delta(t)$ and we finally get LANGEVIN's equation:

$$m \frac{d^2 x}{dt^2} + \gamma \dot{x}(t) = -V'(x) - V_e'(x) + \xi(t) \tag{1.36}$$

with ξ the noise, defined by:

$$\langle \xi(t) \rangle = 0 \quad \text{and} \quad \langle \xi(t)\xi(t') \rangle = 2\gamma k_B T \delta(t - t') \tag{1.37}$$

If $N = 1$ then $K(t)$ is a pure cosine. When $N \to \infty$, there are so much cosines of different pulsations involved in the precedent sum that it seems to fluctuate randomly; it is only at very large timescales that we can observe periodicity. This illustrates a principle according to which something very complex can often be seen as random.

2 Methods and models of stochastic dynamics and stochastic processes

2.1 The over-damped LANGEVIN equation

In this section, we will focus on long timescales. According to LANGEVIN's equation, without force:

$$x(t) = x(0) + \tau\left(1 - \exp\left(\frac{-t}{\tau}\right)\right)\dot{x}(0) + \frac{1}{\gamma}\int_0^t \left(1 - \exp\left(-\frac{t-t'}{\tau}\right)\right)\xi(t')dt' \tag{2.1}$$

where we defined the characteristic $\tau = m/\gamma$. Then, we compute the mean quadratic displacement:

$$\langle x(t)^2\rangle = \tau^2\left(1 - \exp\left(\frac{-t}{\tau}\right)\right)^2\langle\dot{x}(0)^2\rangle \tag{2.2}$$

$$+ \frac{1}{\gamma^2}\int_0^t\int_0^t \left(1 - \exp\left(-\frac{t-t'}{\tau}\right)\right)\left(1 - \exp\left(-\frac{t-t''}{\tau}\right)\right)\langle\xi(t')\xi(t'')\rangle\, dt'dt'' \tag{2.3}$$

that is:

$$\langle x(t)^2\rangle = \tau^2\left(1 - \exp\left(\frac{-t}{\tau}\right)\right)^2\langle\dot{x}(0)^2\rangle + \frac{1}{\gamma^2}\left(t - 2\tau\left(1 - \exp\left(\frac{-t}{\tau}\right)\right) + \frac{\tau}{2}\left(1 - \exp\left(\frac{-2t}{\tau}\right)\right)\right) \tag{2.4}$$

here we used that $x(0) = 0$ and the fact that the initial speed is uncorrelated from noise : $\langle\dot{x}(0)\xi(t)\rangle = 0$. There are two interesting limits:

$$\langle x(t)^2\rangle \underset{t\ll\tau}{\sim} \langle\dot{x}(0)^2\rangle t^2 \quad \text{and} \quad \langle x(t)^2\rangle \underset{t\gg\tau}{\sim} \frac{1}{\gamma^2}t \tag{2.5}$$

This tells us that at small times ($t \ll \tau$), the displacement is ballistic: the particle moves in a straight line at a constant speed; while at long times ($t \gg \tau$), the displacement is diffusive: the initial speed is damped and the displacement is only due to the noise. In this regime, the inertia is not observable and it can be neglected.

Setting $\gamma = 1$ and $k_B = 1$, the resulting equation at long timescales is the over-damped LANGEVIN equation:

$$\dot{x}(t) = -V'(x) + \xi(t) \tag{2.6}$$

with $\langle\xi(t)\rangle = 0$ and $\langle\xi(t)\xi(t')\rangle = 2T\delta(t-t')$. Since in the discrete case, $\xi(t)$ was described by (1.4); in a sense:

$$\xi(t) \sim \sqrt{\frac{2T}{dt}} \tag{2.7}$$

The equation of Brownian motion that is used in mathematics comes from: $dx = -V'(x)dt + \xi(t)dt$, and is rather written as: $dx = -dV + dB_t$, where dV is deterministic while dB_t is random.

2.2 A paradox

Let us show a paradox, if $V'(x) = 0$ we have:

$$\frac{dx}{dt} = \xi(t) \tag{2.8}$$

which, if $x(0) = 0$, is obviously solved by:

$$x(t) = \int_0^t \xi(t')dt' \tag{2.9}$$

so, using that $\langle\xi(t_1)\xi(t_2)\rangle = 2T\delta(t_1 - t_2)$, we obtain the following correlation function:

$$\langle x^2(t)\rangle = \int_0^t\int_0^t \langle\xi(t_1)\xi(t_2)\rangle\, dt_1dt_2 = 2T\int_0^t dt_1 = 2Tt \tag{2.10}$$

this accounts for a diffusive process as pointed out previously, and it finally yields:

$$\frac{d}{dt}\langle x^2(t)\rangle = 2T \tag{2.11}$$

but one could also write:

$$\frac{d}{dt}\langle x^2(t)\rangle = \langle 2x(t)\frac{dx}{dt}\rangle = 2\langle x(t)\xi(t)\rangle \tag{2.12}$$

then, going back to a discrete version of the LANGEVIN equation:

$$\frac{x_{t+\Delta} - x_t}{\Delta} = \xi_t \tag{2.13}$$

we have that x_t and ξ_t are uncorrelated[6], so we obtain $\langle x_t \xi_t \rangle = 0$. Finally, there are two different results for the same quantity, so there must be an error somewhere. We will fix it in the following.

2.3 Discretization of the LANGEVIN equation

There are two common choices of discretization: the ITŌ and the STRATONOVICH conventions.

2.3.1 ITŌ discretization

Discretization of the LANGEVIN equation:

$$\frac{x_{t+\Delta} - x_t}{\Delta} = -V'(x_t) + \xi_t \tag{2.14}$$

with an associated discretization of the correlations:

$$\langle f(x(t))\rangle \to \langle f(x_t)\rangle\,, \quad \langle f(x(t))\xi(t)\rangle \to \langle f(x_t)\xi_t\rangle \quad \text{and} \quad \langle f(x(t))\dot{x}(t)\rangle \to \left\langle f(x_t)\frac{x_{t+\Delta} - x_t}{\Delta}\right\rangle \tag{2.15}$$

which leads to ITŌ's chain rule:

$$\frac{d}{dt}\langle f(x(t))\rangle = \left\langle f'(x(t))\frac{dx}{dt}\right\rangle + T\langle f''(x(t))\rangle \tag{2.16}$$

let us prove it. We start from:

$$\frac{d}{dt}\langle f(x(t))\rangle = \lim_{\Delta\to 0}\frac{\langle f(x(t+\Delta))\rangle - \langle f(x(t))\rangle}{\Delta} = \lim_{\Delta\to 0}\left\langle \frac{f(x_{t+\Delta}) - f(x_t)}{\Delta}\right\rangle \coloneqq \left\langle \frac{f(x_{t+\Delta}) - f(x_t)}{\Delta}\right\rangle \tag{2.17}$$

where we dropped the limit in the last equality by convention. In the discretized version of the LANGEVIN equation, we have $x_{t+\Delta} - x_t$ of the order of Δ, so let us TAYLOR expand f around x_t in powers of $x_{t+\Delta} - x_t$:

$$f(x_{t+\Delta}) = f(x_t) + f'(x_t)(x_{t+\Delta} - x_t) + \frac{f''(x_t)}{2}(x_{t+\Delta} - x_t)^2 + \cdots \tag{2.18}$$

then, inserting this into (2.17) and plugging in the discretized version of the LANGEVIN equation leads to:

$$\frac{d}{dt}\langle f(x(t))\rangle = \left\langle f'(x_t)\frac{x_{t+\Delta} - x_t}{\Delta}\right\rangle + \left\langle \frac{f''(x_t)}{2}\frac{(x_{t+\Delta} - x_t)^2}{\Delta}\right\rangle + \cdots \tag{2.19}$$

$$= \left\langle f'(x_t)\frac{x_{t+\Delta} - x_t}{\Delta}\right\rangle + \left\langle \frac{f''(x_t)}{2}\frac{(-V'(x_t) + \xi_t)^2\Delta^2}{\Delta}\right\rangle + \cdots \tag{2.20}$$

[6] to be precise, $x_{t+\Delta}$ correlated with ξ_t, but not x_t

we have seen that (1.4) implies in a sense that $\xi \sim \sqrt{2T/\Delta}$, so at highest order when $\Delta \to 0$, we have:

$$(-V'(x_t) + \xi_t)^2 \Delta \sim \xi_t^2 \Delta \tag{2.21}$$

finally, x_t and ξ_t being independent:

$$\frac{d}{dt}\langle f(x(t))\rangle = \left\langle f'(x_t)\frac{x_{t+\Delta}-x_t}{\Delta}\right\rangle + \left\langle \frac{f''(x_t)}{2}\right\rangle \langle \xi_t^2 \Delta\rangle \tag{2.22}$$

$$= \left\langle f'(x_t)\frac{x_{t+\Delta}-x_t}{\Delta}\right\rangle + T\langle f''(x_t)\rangle \tag{2.23}$$

and using reversed ITŌ's rules for the discretization of the correlations, we recover the announced chain rule.

2.3.2 STRATONOVICH discretization

Discretization of the LANGEVIN equation:

$$\frac{x_{t+\Delta}-x_t}{\Delta} = -V'(x_t) + \xi_t \tag{2.24}$$

but this time with a different discretization of the correlations:

$$\langle f(x(t))\rangle \to \left\langle \frac{f(x_{t+\Delta})+f(x_t)}{2}\right\rangle \tag{2.25}$$

$$\langle f(x(t))\xi(t)\rangle \to \left\langle \frac{f(x_{t+\Delta})+f(x_t)}{2}\xi_t\right\rangle \tag{2.26}$$

$$\langle f(x(t))\dot{x}(t)\rangle \to \left\langle \frac{f(x_{t+\Delta})+f(x_t)}{2}\frac{x_{t+\Delta}-x_t}{\Delta}\right\rangle \tag{2.27}$$

which leads to the usual chain rule:

$$\frac{d}{dt}\langle f(x(t))\rangle = \left\langle f'(x(t))\frac{dx}{dt}\right\rangle \tag{2.28}$$

that we can prove in a similar way:

$$\frac{d}{dt}\langle f(x(t))\rangle = \left\langle \frac{f(x_{t+\Delta})-f(x_t)}{\Delta}\right\rangle = \left\langle f'(x_t)\frac{x_{t+\Delta}-x_t}{\Delta}\right\rangle + \left\langle \frac{f''(x_t)}{2}\frac{(x_{t+\Delta}-x_t)^2}{\Delta}\right\rangle + \cdots \tag{2.29}$$

using the rule, we have on the other side:

$$\left\langle f'(x(t))\frac{dx}{dt}\right\rangle = \left\langle \frac{f'(x_{t+\Delta})+f'(x_t)}{2}\frac{x_{t+\Delta}-x_t}{\Delta}\right\rangle \tag{2.30}$$

finally, using a TAYLOR expansion of f' near x_t in powers of $(x_{t+\Delta}-x_t)$, namely:

$$f'(x_{t+\Delta}) = f'(x_t) + (x_{t+\Delta}-x_t)f''(x_t) + \cdots \tag{2.31}$$

we recover:

$$\left\langle f'(x(t))\frac{dx}{dt}\right\rangle = \left\langle f'(x_t)\frac{x_{t+\Delta}-x_t}{\Delta} + \frac{f''(x_t)}{2}\frac{(x_{t+\Delta}-x_t)^2}{\Delta}\right\rangle + \cdots \tag{2.32}$$

hence the announced chain rule.

2.3.3 Differences between the ITŌ and the STRATONOVICH conventions

For a classical LANGEVIN equation $\dot{x}(t) = -V(x) + \xi(t)$, the discretizations are the same and the two conventions are equivalent when $\Delta \to 0$. However, calculi are often easier with ITŌ calculus.

With a multiplicative noise (for instance $\dot{x}(t) = -V'(x) + g(x(t))\xi(t)$), the discretized equations are different in the two conventions. This means that the choice of convention changes the results.

Choosing a convention is equivalent to making an hypothesis. ITŌ's convention admits that the position at instant t is independent from the noise on the interval $[t, t+\Delta]$, whereas the STRATONOVICH convention has them correlated. In physical models where there is retarded friction, like in the example with $K(t-s)$, we obtain STRATONOVICH.

2.3.4 Fixing the paradox of the introduction

Using ITŌ's chain rule, with $f : x \mapsto x^2$

$$\frac{d}{dt}\left\langle x^2(t)\right\rangle = 2\left\langle x(t)\frac{dx}{dt}\right\rangle + 2T \tag{2.33}$$

so according to the LANGEVIN equation:

$$\frac{d}{dt}\left\langle x^2(t)\right\rangle = 2\left\langle x(t)\xi(t)\right\rangle + 2T = 2T \tag{2.34}$$

since $V'(x) = 0$ and $x(t), \xi(t)$ are uncorrelated. Doing the calculation in the STRATONOVICH:

$$\frac{d}{dt}\left\langle x^2(t)\right\rangle = 2\left\langle x(t)\frac{dx}{dt}\right\rangle = 2\left\langle \frac{x_{t+\Delta}+x_t}{2}\xi_t\right\rangle = \left\langle x_{t+\Delta}\xi_t\right\rangle + \left\langle x_t\xi_t\right\rangle \tag{2.35}$$

according to the LANGEVIN equation, we have $x_{t+\Delta} = x_t + \xi_t\Delta$, so using $\langle x_t\xi_t\rangle = 0$ we get:

$$\frac{d}{dt}\left\langle x^2(t)\right\rangle = 2\langle x_t\xi_t\rangle + \Delta\left\langle \xi_t^2\right\rangle = 2T \tag{2.36}$$

with the two conventions, we recover the result of the introduction. The error lied in the discretization.

2.4 FOKKER-PLANCK equation

2.4.1 Derivation from the over-damped LANGEVIN equation

Let $\mathbb{P}(x,t)$ be the probability density function to find a particle in $[x, x+dx]$ at time t, and let x satisfy:

$$\dot{x}(t) = -V'(x) + \xi(t) \tag{2.37}$$

if f is a function, we have:

$$\frac{d}{dt}\left\langle f(x(t))\right\rangle = \frac{d}{dt}\int \mathbb{P}(x,t)f(x)dx = \int \frac{\partial \mathbb{P}}{\partial t}(x,t)f(x)dx \tag{2.38}$$

but using ITŌ's chain rule:

$$\frac{d}{dt}\left\langle f(x(t))\right\rangle = \left\langle f'(x(t))\frac{dx}{dt}\right\rangle + T\left\langle f''(x(t))\right\rangle \tag{2.39}$$

with LANGEVIN's equation:

$$\frac{d}{dt}\left\langle f(x(t))\right\rangle = \left\langle f'(x(t))\left(-V'(x(t)) + \xi(t)\right)\right\rangle + T\left\langle f''(x(t))\right\rangle \tag{2.40}$$

so, since $\langle f'(x(t))\xi(t)\rangle = 0$:

$$\frac{d}{dt}\langle f(x(t))\rangle = \int \left[\frac{df}{dx}(x)\left(-\frac{dV}{dx}(x)\right) + T\frac{d^2f}{dx^2}(x)\right]\mathbb{P}(x,t)dx \tag{2.41}$$

performing an integration by parts, and using that $\mathbb{P}(x,t)$ is a probability density vanishing at $x \to \infty$:

$$\int \frac{\partial \mathbb{P}}{\partial t}(x,t)f(x)dx = \int \frac{\partial}{\partial x}\left[\frac{dV}{dx}(x) + T\frac{\partial}{\partial x}\right]\mathbb{P}(x,t)f(x)dx \tag{2.42}$$

this is true for any function f, so:

$$\boxed{\frac{\partial \mathbb{P}}{\partial t}(x,t) = \frac{\partial}{\partial x}\left[\frac{dV}{dx} + T\frac{\partial}{\partial x}\right]\mathbb{P}(x,t)} \tag{2.43}$$

it can be written as : $\partial_t \mathbb{P}(x,t) = -H_{FP}\mathbb{P}(x,t)$ with H_{FP} the FOKKER-PLANCK operator, defined by:

$$H_{FP} = -\frac{\partial}{\partial x}\left[\frac{dV}{dx} + T\frac{\partial}{\partial x}\right] \tag{2.44}$$

the precedent FOKKER-PLANCK equation is similar to a SCHRÖDINGER equation in imaginary time.

2.4.2 Stationary distribution of the FOKKER-PLANCK equation

We try to solve $\partial_t \mathbb{P}(x,t) = 0$. A good guess is the BOLTZMANN-GIBBS probability density, let us try:

$$H_{FP}\mathbb{P}_{BG}(x) = -\frac{\partial}{\partial x}\left[\frac{dV}{dx} + T\frac{\partial}{\partial x}\right]\frac{1}{Z}\exp\left(-\frac{V(x)}{T}\right) \tag{2.45}$$

$$= -\frac{\partial}{\partial x}\left[\frac{dV}{dx}\frac{1}{Z}\exp\left(-\frac{V(x)}{T}\right) + T\left(-\frac{1}{T}\frac{dV}{dx}\right)\frac{1}{Z}\exp\left(-\frac{V(x)}{T}\right)\right] = 0 \tag{2.46}$$

So $\partial_t \mathbb{P}_{BG} = -H_{FP}\mathbb{P}_{BG} = 0$, so $\mathbb{P}_{BG}$ is stationary. A theorem states that the BOLTZMANN-GIBBS probability distribution $\mathbb{P}_{BG}$ is the only stationary distribution if V is bounded on any finite interval on $\mathbb{R}$ and is such that:

$$\int_{-\infty}^{+\infty} \exp\left(-\frac{V(x)}{T}\right)dx < \infty \tag{2.47}$$

2.4.3 Discrete stochastic dynamics

Let us consider a system with M configurations $\mathbb{C}_i$, with $i \in [\![1, M]\!]$.

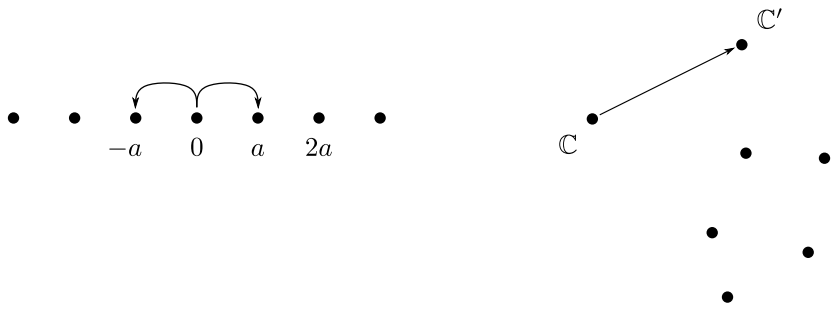

Figure 3: illustration of a one dimensional system (left) and of a transition in the configurations space (right)

We call $H(\mathbb{C})$ the energy of the system in the configuration $\mathbb{C}$. The probability of going from the configuration $\mathbb{C}$ to the configuration $\mathbb{C}'$ in a given time interval $[t, t+dt]$ is denoted $W(\mathbb{C} \to \mathbb{C}')$ and called the transition rate. If the environment is at equilibrium, the detailed balance condition gives, for all pair of configurations $\mathbb{C}$ and $\mathbb{C}'$:

$$\exp\left(-\frac{H(\mathbb{C})}{T}\right)W(\mathbb{C} \to \mathbb{C}') = \exp\left(-\frac{H(\mathbb{C}')}{T}\right)W(\mathbb{C}' \to \mathbb{C}) \tag{2.48}$$

This defines a reversible MARKOV chain if there is no influence from the past in the transition from a state to another. The probability $\mathbb{P}(\mathbb{C}, t)$ for the system to be on the configuration $\mathbb{C}$ at time t satisfies the master equation:

$$\frac{\partial \mathbb{P}}{\partial t}(\mathbb{C}, t) = \sum_{\mathbb{C}'} W(\mathbb{C}' \to \mathbb{C})\mathbb{P}(\mathbb{C}', t) - \sum_{\mathbb{C}'} W(\mathbb{C} \to \mathbb{C}')\mathbb{P}(\mathbb{C}, t) \tag{2.49}$$

applied to $\mathbb{P}_{BG}(\mathbb{C})$:

$$\sum_{\mathbb{C}'}\left[W(\mathbb{C}' \to \mathbb{C})\exp\left(-\frac{H(\mathbb{C}')}{T}\right) - W(\mathbb{C} \to \mathbb{C}')\exp\left(-\frac{H(\mathbb{C})}{T}\right)\right] = 0 \tag{2.50}$$

It is the only distribution to have this property if for all pair of configurations $\mathbb{C}$ and $\mathbb{C}'$, there exists an integer N and a sequence $\mathbb{C}_0, ..., \mathbb{C}_N$, with $\mathbb{C}_0 = \mathbb{C}$ and $\mathbb{C}_N = \mathbb{C}'$, such that : $W(\mathbb{C}_0 \to \mathbb{C}_1) > 0$, ..., $W(\mathbb{C}_{N-1} \to \mathbb{C}') > 0$.

2.5 Properties of the FOKKER-PLANCK operator

Let us resume what we just did; from the overdamped LANGEVIN equation:

$$\dot{x} = -V'(x) + \xi(t) \tag{2.51}$$

the thermal noise $\xi(t)$ satisfying:

$$\langle \xi(t)\xi(t') \rangle = 2T\delta(t - t') \tag{2.52}$$

we obtained the FOKKER-PLANCK equation on the probability density function:

$$\partial_t \mathbb{P}(x, t) = -H_{FP}\mathbb{P}(x, t) \tag{2.53}$$

here, the FOKKER-PLANCK operator is:

$$H_{FP} = -\partial_x\left[T\partial_x + V'(x)\right] \tag{2.54}$$

in what follows we shall study some remarkable properties of this operator.

2.5.1 Property 1 : H_{FP} can be associated with an hermitian operator H_S

The FOKKER-PLANCK operator is not Hermitian but it can be related to a Hermitian operator H_S:

$$H_S = \exp\left(\frac{V(x)}{2T}\right)H_{FP}\exp\left(-\frac{V(x)}{2T}\right) \tag{2.55}$$

Let us check that H_s is Hermitian; we start by replacing H_{FP} by its definition of equation (2.54):

$$H_S = -\exp\left(\frac{V(x)}{2T}\right)\partial_x\left[T\partial_x + V'(x)\right]\exp\left(-\frac{V(x)}{2T}\right) \tag{2.56}$$

we then commute the right exponential with the operator inside the brackets in the precedent expression:

$$\begin{aligned}\left[T\partial_x + V'(x)\right]\exp\left(-\frac{V(x)}{2T}\right) &= \exp\left(-\frac{V(x)}{2T}\right)\left[-\frac{V'(x)}{2} + T\partial_x + V'(x)\right] &(2.57)\\ &= \exp\left(-\frac{V(x)}{2T}\right)\left[T\partial_x + \frac{V'(x)}{2}\right] &(2.58)\end{aligned}$$

after that, we commute that same exponential with the partial derivative on x in equation (2.56):

$$\partial_x \exp\left(-\frac{V(x)}{2T}\right) = \exp\left(-\frac{V(x)}{2T}\right)\left[\partial_x - \frac{V'(x)}{2T}\right] \tag{2.59}$$

so that we can rewrite:

$$H_S = -\frac{1}{T}\left[T\partial_x - \frac{V'(x)}{2}\right]\left[T\partial_x + \frac{V'(x)}{2}\right] \tag{2.60}$$

Defining another operator:

$$A \coloneqq T\partial_x + \frac{V'(x)}{2} \tag{2.61}$$

and using the relation $\partial_x^\dagger = -\partial_x$ (which can be proved using an integration by parts on $\langle\phi_1|\partial_x\phi_2\rangle$):

$$H_S = \frac{1}{T}A^\dagger A \tag{2.62}$$

Under this form this is clear that H_S is Hermitian.

2.5.2 Property 2 : H_S is positive definite

The operator H_S is positive definite, which means that for any state $|\psi\rangle$:

$$\langle\psi|\, H_s\, |\psi\rangle = \frac{1}{T}\left\|A\,|\psi\rangle\right\|_2^2 \geq 0 \tag{2.63}$$

where we use the notations from quantum mechanics: the probability to find a particle in $[x, x+dx]$ at time t is denoted $\mathbb{P}(x,t) = \langle x|\mathbb{P}(t)\rangle$, with $\mathbb{P}(x,0) = \langle x|\mathbb{P}_0\rangle$ and $\mathbb{P}(x,t) = \langle x|\, e^{-H_{FP}t}\,|\mathbb{P}_0\rangle$.

2.5.3 Property 3 : H_{FP} and H_S have the same spectra

The spectral properties of the operators H_S and H_{FP} are related. Indeed, denoting by $|\alpha\rangle$ the eigenvectors of H_s and by λ_α the corresponding eigenvalues:

$$H_S = \sum_\alpha \lambda_\alpha\, |\alpha\rangle\, \langle\alpha| \tag{2.64}$$

so using equation (2.55):

$$H_{FP} = \sum_\alpha \lambda_\alpha \exp\left(-\frac{V}{2T}\right)|\alpha\rangle\,\langle\alpha|\exp\left(\frac{V}{2T}\right) = \sum_\alpha \lambda_\alpha\, |\alpha\rangle_{R\ L}\,\langle\alpha| \tag{2.65}$$

where $|\alpha\rangle_R$ and $|\alpha\rangle_L$ are respectively the right and left eigenvectors of the operator H_{FP}; they satisfy:

$${}_L\langle\alpha|\beta\rangle_R = \langle\alpha|\exp\left(\frac{V}{2T}\right)\exp\left(-\frac{V}{2T}\right)|\beta\rangle = \langle\alpha|\beta\rangle = \delta_{\alpha\beta} \tag{2.66}$$

and also:

$$H_{FP}\,|\beta\rangle_R = \sum_\alpha \lambda_\alpha\,|\alpha\rangle_{RL}\,\langle\alpha|\beta\rangle_R = \sum_\alpha \lambda_\alpha\,|\alpha\rangle_R\,\delta_{\alpha\beta} = \lambda_\beta\,|\beta\rangle_R \tag{2.67}$$

$${}_L\langle\beta|\,H_{FP} = \sum_\alpha \lambda_{\alpha L}\,\langle\beta|\alpha\rangle_{RL}\,\langle\alpha| = \sum_\alpha \lambda_\alpha\delta_{\beta\alpha L}\,\langle\alpha| = \lambda_{\beta L}\,\langle\beta| \tag{2.68}$$

$$\tag{2.69}$$

In conclusion, this shows that the eigenvalues of H_S and H_{FP} are the same; the corresponding eigenvectors being related by:

$$|\alpha\rangle_R = \frac{1}{\sqrt{Z}} \exp\left(-\frac{V}{2T}\right)|\alpha\rangle \quad \text{and} \quad {}_L\langle\alpha| = \langle\alpha| \exp\left(\frac{V}{2T}\right)\sqrt{Z} \tag{2.70}$$

here, Z is a factor allowing to keep the states normalized after the multiplication by the exponentials, it vanished in equation (2.65) and is given by:

$$Z = \int dx \exp\left(-\frac{V(x)}{T}\right) \tag{2.71}$$

2.5.4 Property 4 : H_{FP} has an unique null eigenvector

The eigenvalues of the FOKKER-PLANCK operator are either positive or null since H_s is positive definite as shown in equation (2.63). The only null eigenvector is the state $|0\rangle$ which we define as:

$$\langle x|0\rangle = \frac{1}{\sqrt{Z}} \exp\left(-\frac{V(x)}{2T}\right) \tag{2.72}$$

We show that $|0\rangle$ is a null eigenvector of H_s using the stationarity of the BOLTZMANN-GIBBS distribution:

$$H_{FP}\left[\frac{1}{Z}\exp\left(-\frac{V}{T}\right)\right] = 0 \tag{2.73}$$

so using the definition (2.55):

$$\sqrt{Z}\exp\left(\frac{V}{2T}\right)\exp\left(-\frac{V}{2T}\right)H_S\exp\left(\frac{V}{2T}\right)\frac{1}{Z}\exp\left(-\frac{V}{T}\right) = 0 \tag{2.74}$$

and then:

$$H_S\left[\frac{1}{\sqrt{Z}}\exp\left(-\frac{V}{2T}\right)\right] = 0 \tag{2.75}$$

which shows that $|0\rangle$ is a null eigenvector of H_S, in other words $\lambda_0 = 0$. Using a generalization of the Perron-Frobenius theorem, we could show that $\lambda_\alpha > 0$ for $\alpha > 0$ and that this null eigenvector of H_{FP} is unique

Let us compute the right and left eigenvectors of H_{FP} corresponding to $|0\rangle$, using equations (2.70) and (2.72):

$$\langle x|0\rangle_R = \langle x|\frac{1}{\sqrt{Z}}\exp\left(-\frac{V}{2T}\right)|0\rangle = \frac{1}{\sqrt{Z}}\exp\left(-\frac{V(x)}{2T}\right)\frac{1}{\sqrt{Z}}\exp\left(-\frac{V(x)}{2T}\right) = \mathbb{P}_{BG} \tag{2.76}$$

and:

$${}_L\langle 0|x\rangle = \langle 0|\sqrt{Z}\exp\left(\frac{V}{2T}\right)|x\rangle = \sqrt{Z}\exp\left(\frac{V(x)}{2T}\right)\frac{1}{\sqrt{Z}}\exp\left(-\frac{V(x)}{2T}\right) = 1 \tag{2.77}$$

2.5.5 Property 5 : the eigenvalues of H_{FP} are related to the relaxation time of the system

The eigenvalues of the FOKKER-PLANCK operator are related to **relaxation times** of the system; that is to say the characteristic it takes to thermalize and to converge towards the Boltzmann-Gibbs distribution. Indeed, let us show that for any $\mathbb{P}_0(x)$:

$$\mathbb{P}(x,t) \underset{t\to\infty}{\to} \mathbb{P}_{BG}(x) \tag{2.78}$$

Let us start by showing:

$$\exp(-H_{FP}t) = \sum_\alpha \exp(-\lambda_\alpha t)\,|\alpha\rangle_R\ {}_L\langle\alpha| \tag{2.79}$$

if we develop both side of this equation with respect to t, the precedent equality comes down to equality of each term, the first ones imply:

$$\sum_\alpha |\alpha\rangle_R \ {}_L\langle\alpha| = \sum_\alpha |\alpha\rangle\langle\alpha| = \mathbb{1} \tag{2.80}$$

which is true as soon as H_S is diagonalizable; the equality of the second terms gives:

$$H_{FP} = \sum_\alpha \lambda_\alpha |\alpha\rangle_R \ {}_L\langle\alpha| \tag{2.81}$$

this is exactly what is shown in equation (2.65); finally the equality of the third terms:

$$H_{FP}^2 = \sum_\alpha \lambda_\alpha^2 |\alpha\rangle_R \ {}_L\langle\alpha| \tag{2.82}$$

can be proven by writing:

$$H_{FP}^2 = \sum_{\alpha,\beta} \lambda_\alpha\lambda_\beta |\alpha\rangle_R \ {}_L\langle\alpha|\beta\rangle_R \ {}_L\langle\beta| = \sum_\alpha \lambda_\alpha^2 |\alpha\rangle_R \ {}_L\langle\alpha| \tag{2.83}$$

here we used equation (2.66); the equality of the higher order terms can be proved the same way. Now, using the definition of $\mathbb{P}(x,t)$ in terms of quantum mechanics notations given below equation (2.63) and the equation that we have just proved:

$$\mathbb{P}(x,t) = \langle x|\exp(-H_{FP}t)|\mathbb{P}_0\rangle = \sum_\alpha \exp(-\lambda_\alpha t)\langle x|\alpha\rangle_R \ {}_L\langle\alpha|\mathbb{P}_0\rangle \tag{2.84}$$

In **2.5.4** we proved that $\lambda_0 = 0$; then using equation (2.77) and various definitions, we show that:

$${}_L\langle 0|\mathbb{P}_0\rangle = \int dx \, {}_L\langle 0|x\rangle\langle x|\mathbb{P}_0\rangle = \int dx \mathbb{P}_0(x) = 1 \tag{2.85}$$

finally, with equation (2.76):

$$\mathbb{P}(x,t) = \mathbb{P}_{BG} + \sum_{\alpha>0} \exp(-\lambda_\alpha t)\langle x|\alpha\rangle_{RL}\langle\alpha|\mathbb{P}_0\rangle \tag{2.86}$$

since $\lambda_\alpha > 0$ if $\alpha \neq 0$, it implies:

$$\boxed{\mathbb{P}(x,t) \underset{t\to\infty}{\to} \mathbb{P}_{BG}(x)} \tag{2.87}$$

which is what was announced. This phenomenon is called the thermalization: given enough time, the system will always end up distributed according to the thermal Boltzmann distribution. And as we see in the precedent equation, the relaxation time is related to the eigenvalues of the FOKKER-PLANCK operator, if we denote them in such a way that $\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$:

$$\boxed{\tau_{\text{relaxation}} = \frac{1}{\lambda_1}} \tag{2.88}$$

Under certain conditions, the other eigenvectors are related to the probability distribution of metastable states and the associated eigenvalues to their lifetimes.

Metastable states are easily pictured as in Figure 4. The precedent result shows that even if the initial probability distribution is "trapped" in the well on the right, that is to say $T \ll \Delta$, at some point a fluctuation will happen and the probability distribution will move to the well on the left. The higher the barrier, the longer it can take.

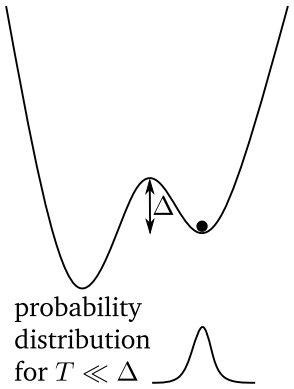

Figure 4: asymmetric double well potential; the initial condition is in the right well, with a probability distribution pictured on the bottom; it is metastable

2.6 Relationship with quantum mechanics

From equation (2.60):

$$H_S = -\frac{1}{T}\left[T\partial_x - \frac{V'(x)}{2}\right]\left[T\partial_x + \frac{V'(x)}{2}\right] \tag{2.89}$$

$$= -\frac{1}{T}\left[T^2\partial_x^2 + T\frac{V''(x)}{2} + T\frac{V'(x)}{2}\partial_x - T\frac{V'(x)}{2}\partial_x - \left(\frac{V'(x)}{2}\right)^2\right] \tag{2.90}$$

$$= -T\partial_x^2 - \frac{V''(x)}{2} + \frac{1}{T}\left(\frac{V'(x)}{2}\right)^2 \tag{2.91}$$

so we recognize that H_S is of the same form as the Hamiltonian operator:

$$\boxed{H_S = -\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2} + V_{\text{eff}}(x)} \tag{2.92}$$

with the identifications:

$$\hbar = \sqrt{T}, \quad m = \frac{1}{2} \quad \text{and} \quad V_{\text{eff}}(x) = \frac{1}{T}\frac{\left(V'(x)\right)^2}{4} - \frac{V''(x)}{2} \tag{2.93}$$

Let $\mathcal{O}$ be an observable; the average of this observable with the equilibrium Boltzmann probability density can be written:

$$\langle\mathcal{O}\rangle_{\text{eq}} = \int dx \frac{1}{Z}\exp\left(-\frac{V(x)}{T}\right)\mathcal{O}(x) = \int dx \,\langle 0|x\rangle\,\langle x|0\rangle\,\mathcal{O}(x) = \langle 0|\,\hat{\mathcal{O}}\,|0\rangle = \langle\hat{\mathcal{O}}\rangle_{\text{ground state}} \tag{2.94}$$

Therefore, we see that there is an analogy between stochastic dynamics governed by the FOKKER-PLANCK operator and quantum dynamics governed by the Hamiltonian operator H_S. In this analogy, averages in the equilibrium are replaced by averages in the ground state and dynamical correlation functions by quantum correlation functions.

3 Path integrals and stochastic dynamics

3.1 Derivation of the MSRDJ path integral

Let us consider a particle evolving according to the LANGEVIN's equation $\dot{x} = -V'(x) + \xi(t)$, and let us study the probability that this particle arrives in x_f at the time t_f, knowing that it was in x_0 at time $t = 0$. This particle can have moved along any of the numerous possible paths, see Figure 5 (left), each of which associated to a probability $\mathbb{P}[\text{path}]$, so the probability we study is:

$$\mathbb{P}(x_f, t_f | x_0, 0) = \sum_{\text{paths}} \mathbb{P}[\text{path}] \tag{3.1}$$

in what follows, we will work in a discretised space, as depicted in Figure 5 (right) for an arbitrary path.

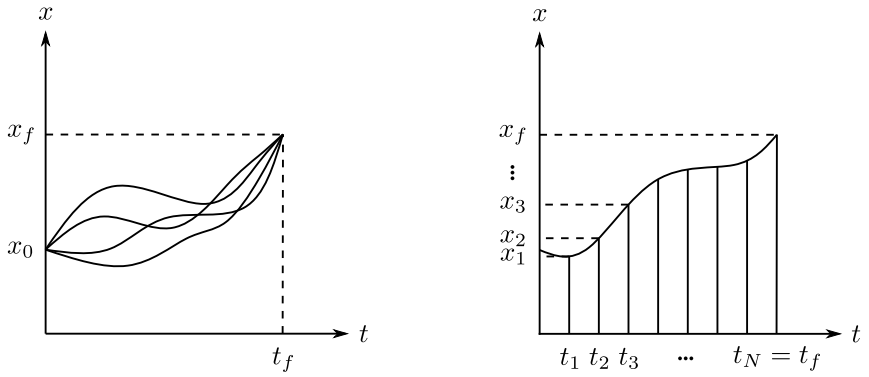

Figure 5: the particle goes from x_0 at $t = 0$ to x_f at t_f, following any path (left); the space is discretised (right)

The path is determined by the knowledge of x_0, all the $\{\xi_n\}$ and by the time step Δ. Indeed using the discretized version of the LANGEVIN equation:

$$\frac{x_{t+\Delta} - x_t}{\Delta} = -V'(x_t) + \xi_t \tag{3.2}$$

we obtain directly that for all n:

$$x_{n+1} = x_n + \Delta(\xi_n - V'(x_n)) \tag{3.3}$$

so after the right number N of steps, which depends on Δ, we obtain a value of x_N as a function of the $\{\xi_n\}$. We write this value $x_N\{\xi_n\}$. It can be defined for any set of $\{\xi_n\}$, but the path which contribute in the sum of equation (3.1) are those for which $x_N\{\xi_n\} = x_f$, and we rewrite the sum of equation (3.1) as the sum over all the possible $\{\xi_n\}$ of those contributing paths:

$$\mathbb{P}(x_f, t_f | x_0, 0) = \langle \delta[x_f - x_N\{\xi_n\}] \rangle_{\{\xi_n\}} \tag{3.4}$$

This can directly be rewritten as follows, in a form which explicitly contains the $\{\xi_n\}$ dependence of x_N:

$$\mathbb{P}(x_f, t_f | x_0, 0) = \left\langle \int dx_1 \delta[x_1 - x_0 + \Delta(V'(x_0) - \xi_0)] \times \int dx_2 \delta[x_2 - x_1 + \Delta(V'(x_1) - \xi_1)] \right. \tag{3.5}$$

$$\left. \times \ldots \times \int dx_N \delta[x_N - x_{N-1} + \Delta(V'(x_{N-1} - \xi_{N-1})] \times \delta[x_f - x_N] \right\rangle_{\{\xi_n\}} \tag{3.6}$$

now, using a classical identity:

$$\delta(f(x)) = \int \frac{d\hat{x}}{2\pi} \exp(-i\hat{x} f(x)) \tag{3.7}$$

we have:

$$\delta\left[x_{n+1} - x_n + \Delta(V'(x_n) - \xi_n)\right] = \int \frac{d\hat{x}_{n+1}}{2\pi} \exp(-i\hat{x}_{n+1}(x_{n+1} - x_n + \Delta(V'(x_n) - \xi_n))) \tag{3.8}$$

so:

$$\mathbb{P}(x_f,t_f|x_0,0)=\left\langle\int\prod_n\frac{dx_nd\hat{x}_n}{2\pi}\exp(-i\hat{x}_{n+1}(x_{n+1}-x_n+\Delta V'(x_n)))\exp(i\Delta\hat{x}_n\xi_n)\delta[x_f-x_N]\right\rangle_{\{\xi_n\}} \tag{3.9}$$

Here, x_n is just an integration variable, and the only term still containing the random variables over which we average is:

$$\left\langle\prod_n\exp(i\Delta\hat{x}_n\xi_n)\right\rangle_{\{\xi_n\}}=\prod_n\langle\exp(i\Delta\hat{x}_n\xi_n)\rangle_{\{\xi_n\}}=\prod_n\exp(-\Delta T\hat{x}_n^2) \tag{3.10}$$

which was obtained using the fact that the ξ_n are independent and identically distributed Gaussian random variables with correlations function:

$$\langle\xi_n\xi_m\rangle=\frac{2T}{\Delta}\delta_{n,m} \tag{3.11}$$

and the fact that for a Gaussian random variable X of width σ, by completing the square in the integral:

$$\langle e^{aX}\rangle=\frac{1}{\sqrt{2\pi\sigma^2}}\int dX\exp\left(aX-\frac{X^2}{2\sigma^2}\right)=\exp\left(\frac{a^2\sigma_X^2}{2}\right) \tag{3.12}$$

Now:

$$\mathbb{P}(x_f,t_f|x_0,0)=\int\prod_n\frac{dx_nd\hat{x}_n}{2\pi}\delta[x_f-x_N]\exp\left(-\Delta\sum_n i\hat{x}_{n+1}\left(\frac{x_{n+1}-x_n}{\Delta}+V'(x_n)-T(i\hat{x}_n)\right)\right) \tag{3.13}$$

doing:

$$\Delta\sum_n\to\int dt,\quad\frac{n_{n+1}-x_n}{\Delta}\to\dot{x}(t),\quad\int\prod_n\frac{dx_nd\hat{x}_n}{2\pi}\delta[x_f-x_N]\to\int_{x(0)=x_0}^{x(t_f)=x_f}dx(t)d\hat{x}(t) \tag{3.14}$$

we get:

$$\boxed{\mathbb{P}(x_f,t_f|x_0,0)=\int_{x(0)=x_0}^{x(t_f)=x_f}dx(t)d\hat{x}(t)\exp(-S\left[x(t),\hat{x}(t)\right])} \tag{3.15}$$

the action being:

$$\boxed{S\left[x(t),\hat{x}(t)\right]=\int dt(i\hat{x}(t))\left(\dot{x}(t)+V'(x(t))-T(i\hat{x}(t))\right)} \tag{3.16}$$

In the literature, our $i\hat{x}$ is often replaced by $\hat{x}$, it is an imaginary field called the response field, x being called the physical field:

$$\mathbb{P}(x_f,t_f|x_0,0)=\int_{x(0)=x_0}^{x(t_f)=x_f}dx(t)d\hat{x}(t)\exp\left(-\int dt\hat{x}(t)(\dot{x}(t)+V'(x(t))-T\hat{x}(t))\right) \tag{3.17}$$

What we obtained is called MARTIN-SIGGIA-ROSE-DE DOMINICIS-JANSEN (or MSRDJ for short) field theory.

3.2 Properties of the path integral

3.2.1 Normalization

The path integral is normalized, in the sense that at time t_f the particle should be found somewhere:

$$\int dx_f\int_{x(0)=x_0}^{x(t_f)=x_f}dxd\hat{x}\exp(-S_{\text{MSRDJ}}[x,\hat{x}])=\int dx_f\mathbb{P}(x_f,t_f|x_0,0)=1 \tag{3.18}$$

3.2.2 Correlation function

The correlation function between two local operators at different times is naturally defined by:

$$\langle \mathcal{O}(x(t_1))\mathcal{O}(x(t_2))\rangle = \int dx d\hat{x} \exp(-S_{\text{MSRDJ}})\mathcal{O}(x(t_1))\mathcal{O}(x(t_2)) \tag{3.19}$$

$$= \sum_{\text{paths}} \mathbb{P}[\text{path}]\mathcal{O}(x(t_1))\mathcal{O}(x(t_2)) \tag{3.20}$$

they can be computed by "cutting the path", that is to say by summing the $\mathcal{O}(x_1)\mathcal{O}(x_2)$ for paths that go from $(0, x_0)$ to (t_f, x_f) while passing at (t_1, x_1) and (t_2, x_2), and by integrating over all possible values of x_1 and x_2:

$$\langle \mathcal{O}(x(t_1))\mathcal{O}(x(t_2))\rangle = \int dx_1 dx_2 \mathbb{P}(x_f, t_f|x_2, t_2)\mathcal{O}(x_2)\mathbb{P}(x_2, t_2|x_1, t_1)\mathcal{O}(x_1)\mathbb{P}(x_1, t_1|x_0, t_0) \tag{3.21}$$

as depicted:

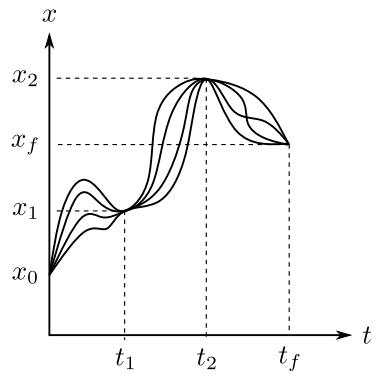

Figure 6: paths going from $(0, x_0)$ to (t_f, x_f) while passing at (t_1, x_1) and (t_2, x_2)

3.2.3 Response function

If we introduce a perturbation $V(x) \to V(x) - h(t)x(t)$. The response function:

$$R(t, t') = \left.\frac{\delta \langle x(t)\rangle}{\delta h(t')}\right|_{h=0} \tag{3.22}$$

is also given by the correlation function:

$$R(t, t') = \langle x(t)\hat{x}(t')\rangle \tag{3.23}$$

this is why we call $\hat{x}$ the response field.

3.3 Other examples of path integrals

3.3.1 ONSAGER-MACHLUP path integral

It is given by:

$$\mathbb{P}(x_f, t_f|x_0, 0) = \int_{x(0)=x_0}^{x(t_f)=x_f} dx \exp\left(-\frac{1}{4T}\int dt(\dot{x}(t) + V'(x))^2\right) \tag{3.24}$$

With $V(x) = 0$ (that is to say for the Brownian motion $\dot{x}(t) = \xi(t)$) and seeing t as a length, the term:

$$E = \frac{1}{4}\int_0^{l_f} dl \left(\frac{dx}{dl}\right)^2 \tag{3.25}$$

can be seen as an energy cost to deform an elastic line, in analogy with elastics and polymer physics.

3.3.2 FEYNMAN-KAC theory

In analogy with:

$$\mathbb{P}(x_f, t_f | x_0, 0) = \langle x_f | \exp(-H_{FP} t_f) | x_0 \rangle = \int_{x(0)=x_0}^{x(t_f)=x_f} dx e^{-S_{OM}} \tag{3.26}$$

we can write:

$$\langle x_f | \exp\left(-\frac{iHt}{\hbar}\right) | x_0 \rangle = \int_{x(0)=x_0}^{x(t_f)=x_f} dx \exp\left(-\frac{i}{\hbar}\int dt \mathcal{L}(x(t))\right) \tag{3.27}$$

here, H is the usual Hamiltonian of quantum mechanics, and $\mathcal{L}$ is the Lagrangian density, which is integrated to give the action.

4 Time-reversal symmetry and fluctuation-dissipation relations

4.1 Time-reversal symmetry

4.1.1 Change of field in the path integral

Let us consider a physical field $x(t)$ and the corresponding response field $\hat{x}(t)$. In this section we study the time reversal of these paths, as in Figure 7.

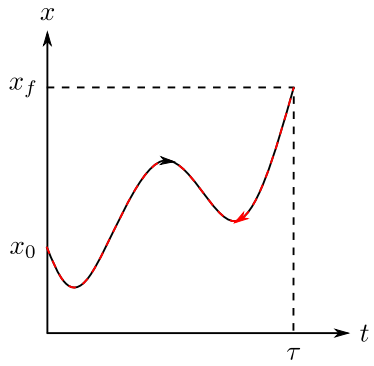

Figure 7: a path $x(t)$ from $(0, x_0)$ to (τ, x_f) in black, and the corresponding time-reversed path $x_R(t)$ in red.

The reversed paths are:

$$x_R(t) \coloneqq x(\tau - t) \quad \text{and} \quad \hat{x}_R(t) \coloneqq \hat{x}(\tau - t) + \frac{1}{T}\frac{d}{dt}[x(\tau - t)] \tag{4.1}$$

from which we easily obtain the reciprocal relations, being careful of a minus sign introduced by the derivative in the calculation:

$$x(t) = x_R(\tau - t) \quad \text{and} \quad \hat{x}(t) = \hat{x}_R(\tau - t) + \frac{1}{T}\frac{d}{dt}[x_R(\tau - t)] \tag{4.2}$$

Now the MSRDJ action that we previously obtained can be rewritten as follows by completing the square:

$$S[x(t), \hat{x}(t)] = \int_0^\tau dt \left(\hat{x}(t) V'(x(t)) - T \left(\hat{x}(t) - \frac{\dot{x}(t)}{2T} \right)^2 + \frac{\dot{x}^2(t)}{4T} \right) \tag{4.3}$$

then:

$$\begin{aligned} S[x(t), \hat{x}(t)] &= \int_0^\tau dt \left(\hat{x}_R(\tau - t) + \frac{1}{T}\frac{d}{dt}[x_R(\tau - t)] \right) V'(x_R(\tau - t)) && (4.4) \\ &- T \int_0^\tau dt \left(\hat{x}_R(\tau - t) + \frac{1}{T}\frac{d}{dt}[x_R(\tau - t)] - \frac{1}{2T}\frac{d}{dt}[x_R(\tau - t)] \right)^2 && (4.5) \\ &+ \frac{1}{4T} \int_0^\tau dt \left(\frac{d}{dt}[x_R(\tau - t)] \right)^2 && (4.6) \end{aligned}$$

so:

$$\begin{aligned} S[x(t), \hat{x}(t)] &= \int_0^\tau du \left(\hat{x}_R(u) - \frac{1}{T}\frac{d}{du}[x_R(u)] \right) V'(x_R(u)) - T \int_0^\tau du \left(\hat{x}_R(u) - \frac{1}{2T}\frac{d}{du}[x_R(u)] \right)^2 && (4.7) \\ &+ \frac{1}{4T} \int_0^\tau du \left(\frac{d}{du} x_R(u) \right)^2 && (4.8) \\ & && (4.9) \end{aligned}$$

and we recognize:

$$\begin{aligned} S[x,\hat{x}] &= S[x_R,\hat{x}_R] - \frac{1}{T}\int_0^\tau du \frac{d}{du}[x_R(u)]V'(x_R(u)) && (4.10)\\ &= S[x_R,\hat{x}_R] - \frac{1}{T}\int_0^\tau du \frac{d}{du}[V(x_R(u))] && (4.11)\\ &= S[x_R,\hat{x}_R] - \frac{1}{T}(V(x_R(\tau)) - V(x_R(0))) && (4.12)\\ & && (4.13) \end{aligned}$$

Finally, we obtain the relation:

$$\boxed{S[x,\hat{x}] = S[x_R,\hat{x}_R] + \frac{1}{T}(V(x_f) - V(x_0))} \tag{4.14}$$

The fact of going from $x(t)$ and $\hat{x}(t)$ to $x_R(t)$ and $\hat{x}_R(t)$ in the path integral can be seen as a change of variable, or more precisely as a change of field. The associated Jacobian is equal to 1 as can be seen going back to the discrete, where the time reversal amounts to a relabeling of the steps which does not change the integration measure:

$$\mathcal{D}x(t) \to \prod_n dx_n = \prod_n dx_{N-n} \tag{4.15}$$

Remark : $\hat{x}$ has been defined as a purely imaginary field, but x is not; so a glance at equation (4.1) tells us that x_R is not purely imaginary. However, it is possible to "shift the contour" in the path integral to make it purely imaginary.

Remark : $x(t)$ is a path obeying the LANGEVIN equation, so according to the ITŌ chain rule:

$$\frac{d}{dt}[f(x(t))] = \dot{x}(t)f'(x(t)) + Tf''(x(t)) \tag{4.16}$$

and in principle there should be an extra term in equation (4.11). But in fact, the path integral corresponds to an "anti-ITŌ" discretization given by:

$$\frac{x_{t+\Delta} - x_t}{\Delta} = -V'(x_{t+\Delta}) + \xi_{t+\Delta} \tag{4.17}$$

for which the extra term vanishes.

4.1.2 Detailed balance and time reversal symmetry

If the probability associated to a path $x(t)$ starting at x_0 is given by the path integral:

$$\mathbb{P}(x(t)|x(0) = x_0) = \int \mathcal{D}\hat{x}(t) \exp(-S[x(t),\hat{x}(t)]) \tag{4.18}$$

then the probability associated to the reversed path $x_R(t)$ starting at x_f is directly given by the following path integral:

$$\mathbb{P}(x_R(t)|x_R(0) = x_f) = \int \mathcal{D}\hat{x}_R(t) \exp\left(-S[x_R(t),\hat{x}_R(t)] - \frac{1}{T}(V(x_f) - V(x_0))\right) \tag{4.19}$$

from there we get:

$$\boxed{\mathbb{P}(x(t)|x(0) = x_0) \exp\left(-\frac{V(x_0)}{T}\right) = \mathbb{P}(x_R(t)|x_R(0) = x_f) \exp\left(-\frac{V(x_f)}{T}\right)} \tag{4.20}$$

this is the detailed balance condition, the analog to the discrete dynamics equation we gave previously:

$$\exp\left(-\frac{H(\mathbb{C})}{T}\right)W(\mathbb{C}\to\mathbb{C}') = \exp\left(-\frac{H(\mathbb{C}')}{T}\right)W(\mathbb{C}'\to\mathbb{C}) \tag{4.21}$$

which means that if we are at equilibrium, we remain at equilibrium. Now, if we divide by Z we have:

$$\mathbb{P}(x(t)|x(0)=x_0)\mathbb{P}(x(0)=x_0) = \mathbb{P}(x_R(t)|x_R(0)=x_f)\mathbb{P}(x_R(0)=x_f) \tag{4.22}$$

or in other words:

$$\boxed{\mathbb{P}(x(t)) = \mathbb{P}(x_R(t))} \tag{4.23}$$

and this is the time-reversal symmetry, telling us that there is no arrow of time in equilibrium dynamics.

4.2 Consequences of the time-reversal symmetry and fluctuation-dissipation relations

4.2.1 Time-reversal symmetry of correlation functions

By time-reversal symmetry:

$$\langle\mathcal{O}(x(t_1))\mathcal{O}(x(t_2))\rangle = \int \mathcal{D}x(t)\mathbb{P}(x(\tau-t))\mathcal{O}(x(t_1))\mathcal{O}(x(t_2)) \tag{4.24}$$

now, performing the change of path $x(t-\tau)=x_R(t)$ in the integral and using equation (4.15):

$$\langle\mathcal{O}(x(t_1))\mathcal{O}(x(t_2))\rangle = \int \mathcal{D}x_R(t)\mathbb{P}(x_R(t))\mathcal{O}(x_R(\tau-t_1))\mathcal{O}(x_R(\tau-t_2)) \tag{4.25}$$

so in other words:

$$\langle\mathcal{O}(x(t_1))\mathcal{O}(x(t_2))\rangle = \langle\mathcal{O}(x(\tau-t_1))\mathcal{O}(x(\tau-t_2))\rangle = \langle\mathcal{O}(x(-t_1))\mathcal{O}(x(-t_2))\rangle \tag{4.26}$$

The last equality translates that the equilibrium dynamics is stationary. The correlation functions do have time-reversal symmetry.

4.2.2 Close to equilibrium: linear response and fluctuation-dissipation relation

Let us introduce a perturbation $V(x)\to V(x)-h(t)x(t)$ and recall the response function given in 3.2.3:

$$R(t,t') = \left.\frac{\delta\langle x(t)\rangle}{\delta h(t')}\right|_{h=0} = \langle x(t)\hat{x}(t')\rangle \tag{4.27}$$

using (4.2) we rewrite:

$$\begin{aligned} R(t,t') &= \left\langle x_R(\tau-t)\left(\hat{x}_R(\tau-t') + \frac{1}{T}\frac{d}{dt'}[x_R(\tau-t')]\right)\right\rangle & (4.28)\\ &= R(\tau-t,\tau-t') + \frac{1}{T}\frac{d}{dt'}[C_{xx}(\tau-t,\tau-t')] & (4.29)\end{aligned}$$

Stationarity close to the equilibrium allows us to say that $R(t,t')$ only depends on $t'-t$, we thus write it $R(t'-t)$; the time reversal symmetry of the correlation functions allows us to do the same for C_{xx} Then, we obtain:

$$R(\tau) = R(-\tau) - \frac{1}{T}\frac{d}{d\tau}C_{xx}(\tau) \tag{4.30}$$

where we defined $\tau = t-t'$ (this is not the same τ as before). By causality, there can be no response before any event happens, so $R(-\tau)=0$ and for $\tau>0$:

$$\boxed{R(\tau) = -\frac{1}{T}\frac{d}{d\tau}C_{xx}(\tau)} \tag{4.31}$$

This is the fluctuation-dissipation relation. The time-reversal symmetry imposes a relation between the response close to equilibrium, represented by the function $R(\tau)$; and fluctuations, represented by the correlation $C_{xx}(\tau)$.

4.2.3 ONSAGER reciprocity relation

More generally, if:

$$V(x) \to V(x) - h_A(t)A(x(t)) \quad \text{and} \quad V(x) \to V(x) - h_B(t)B(x(t)) \tag{4.32}$$

in other words if the field V couples linearly to two fields A and B; and introducing response functions:

$$R_{AB}(t-t') = \left.\frac{\delta \langle A(t)\rangle}{\delta h_B(t')}\right|_{h=0} \quad \text{and} \quad R_{BA}(t-t') = \left.\frac{\delta \langle B(t)\rangle}{\delta h_A(t')}\right|_{h=0} \tag{4.33}$$

we get the relation:

$$\boxed{R_{AB}(t-t') = R_{BA}(t-t')} \tag{4.34}$$

This is ONSAGER's reciprocity relation, which awarded him a NOBEL prize in 1968.

4.3 Fluctuation relations out of equilibrium

Let us consider the following generalization of the previous LANGEVIN equation:

$$\dot{x}(t) = -V'(x(t), \lambda(t)) + f + \xi(t) \tag{4.35}$$

where the $'$ is a derivation with respects to the x variable, f is a non-conservative force, and $\lambda(t)$ is an exterior parameter tuned by the experimentalist (for instance a piston). The same calculation that led us to equation (4.20) would now give:

$$\mathbb{P}_R(x_R(t)|x_R(0) = x_f) = \mathbb{P}(x(t)|x(0) = x_0) \exp\left(-\frac{1}{T}\int_0^\tau dt \frac{d}{dt}[x(t)]F(t)\right) \tag{4.36}$$

defining $F(t) = -V'(x(t), \lambda(t)) + f$. As a fundamental hypothesis, we consider that the system is coupled to an environment which is at equilibrium, and according to the first law of thermodynamics:

$$dW = dV + dq \tag{4.37}$$

with dW the internal energy of the system (which is only mechanical in the case of the particle we consider), dV the mechanical work and dq the heat received from the environment. But from equation (4.35) we see that the variations of energy of the system can only come from the experimentalist varying $\lambda(t)$ or from the non conservative force f, so we write:

$$dW = \frac{\partial V}{\partial \lambda} d\lambda + f dx = dV - \frac{\partial V}{\partial x} dx + f dx \tag{4.38}$$

and we identify:

$$dq = -\frac{\partial V}{\partial x} dx + f dx = F dx \tag{4.39}$$

so that finally:

$$\boxed{\mathbb{P}_R(x_R(t)|x_R(0) = x_f) = \mathbb{P}(x(t)|x(0) = x_0) \exp\left(-\frac{1}{T}\int_0^\tau dt \frac{dq}{dt}\right)} \tag{4.40}$$

even if the variation over time of heat transferred dq/dt is fluctuating, it has to satisfy this equation. We will see in problem class that this implies the following relation, established by Jarzynski :

$$\left\langle \exp\left(-\frac{W}{T}\right)\right\rangle = \exp\left(-\frac{\Delta F}{T}\right) \tag{4.41}$$

where the average is over the stochastic process, the work being a stochastic quantity.

5 Polymer models and path integrals

5.1 Random walks, Brownian motion and path integrals

Let us study a random walk on a d-dimensional square lattice, indexed by $\mathbb{Z}^d$, as depicted in Figure 8.

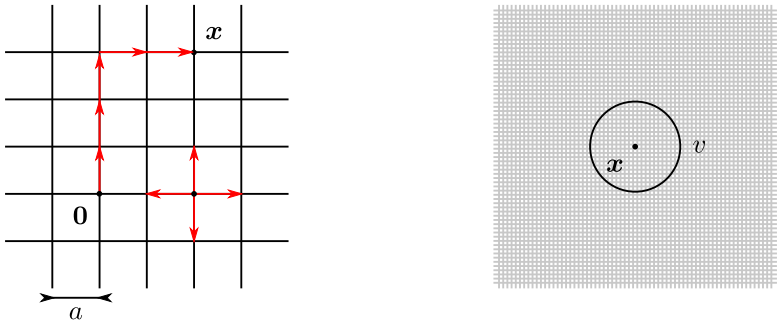

Figure 8: square lattice where a walker moves along the edges (left); volume v in the mesoscopic limit (right)

The walker starts in position $\boldsymbol{x} = \mathbf{0}$ at time $t = 0$, and at every time step, it jumps to an arbitrary neighboring site with uniform probability $1/2d$. The probability to find the walker in $\boldsymbol{x}$ at time t is given by:

$$\mathbb{P}(\boldsymbol{x}, t) = \frac{\text{number of paths ending in } \boldsymbol{x} \text{ at time } t}{\text{number of paths with } t \text{ steps}} \tag{5.1}$$

or by the recursion relation:

$$\mathbb{P}(\boldsymbol{x}, t) = \frac{1}{2d} \sum_{\boldsymbol{x}' \in \partial \boldsymbol{x}} \mathbb{P}(\boldsymbol{x}', t-1) \tag{5.2}$$

where $\partial \boldsymbol{x}$ denotes the neighboring sites of $\boldsymbol{x}$, and with the initial condition: $\mathbb{P}(\boldsymbol{x}, 0) = \delta_{\boldsymbol{x},\mathbf{0}}$

5.1.1 Scaling limit

As in Figure 8 we call the lattice spacing a; and we call τ the time interval between two consecutive jumps. When take them to the continuous $a \to 0$ and $\tau \to 0$ we need to specify a scaling, for instance:

$$a^2 = 2dD\tau \tag{5.3}$$

where D is a diffusion constant. If we do not do so, the limit is ill defined. For instance if a goes to 0 much faster than τ (in the most extreme case, a goes to 0 while τ remains fixed), then after the limit is taken the walker does not move anymore. Conversely, if τ goes to 0 much faster than a, the walker instantly goes to infinity.

Let us consider a volume $v = (\Delta x)^d$ in the mesoscopic limit $\Delta x \gg a$ as shown in Figure 8. The probability $\mathbb{P}_v(\boldsymbol{x}, t)$ to be around $\boldsymbol{x}$ within a volume v at time t is proportional to v. We define the probability density $\rho(\boldsymbol{x}, t)$ as:

$$\lim_{v \to 0} \frac{1}{v} \left(\lim_{a \to 0,\ \tau \to 0,\ a^2 = 2dD\tau} \mathbb{P}_v(\boldsymbol{x}, t) \right) = \rho(\boldsymbol{x}, t) \tag{5.4}$$

it has the dimension of a length^{-d}. The recursion relation of equation (5.2) is translated on ρ as:

$$\partial_t \rho = \triangle_{\boldsymbol{x}} \rho \tag{5.5}$$

with the initial condition $\rho(\boldsymbol{x}, t = 0) = \delta^d(\boldsymbol{x})$; so ρ is the usual propagator of Brownian motion:

$$\rho(\boldsymbol{x}, t) = \frac{1}{(4\pi D t)^{d/2}} \exp\left(-\frac{\boldsymbol{x}^2}{4Dt} \right) \tag{5.6}$$

5.1.2 Path integral for $\rho(\boldsymbol{x}, t)$

Now we will consider the transition probability $(\boldsymbol{x}_0, 0) \to (\boldsymbol{x}, t)$. First, the diffusion equation is translation invariant, so:

$$\rho(\boldsymbol{x}, t|\boldsymbol{x}_0, 0) = \frac{1}{(4\pi D t)^{d/2}} \exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_0)^2}{4Dt}\right) \tag{5.7}$$

this density also enjoys MARKOV's property, which states that the future of a path is independent from its past; to be precise for all t' in $(0, t)$:

$$\rho(\boldsymbol{x}, t|\boldsymbol{x}_0, 0) = \int_{\mathbb{R}^d} d^d\boldsymbol{x}' \rho(\boldsymbol{x}, t|\boldsymbol{x}', t')\rho(\boldsymbol{x}', t'|\boldsymbol{x}_0, 0) \tag{5.8}$$

We represent this in Figure 9 (left), where we have drawn two paths that have different x'; the density ρ includes all such paths, for all x'.

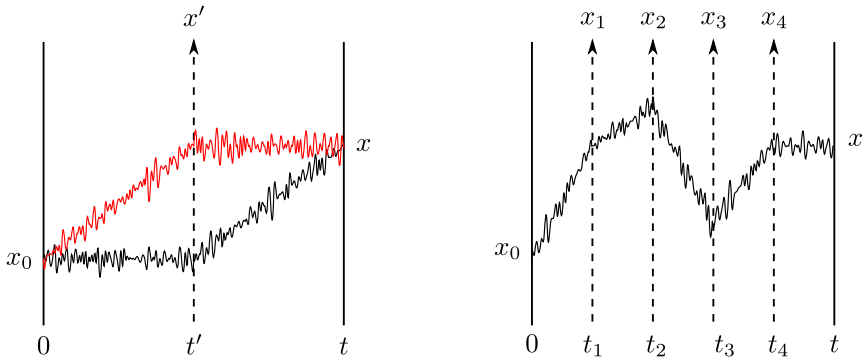

Figure 9: paths going from $(x_0, 0)$ to (x, t) passing at (x', t') (left); the interval $(0, t)$ sliced in sub-intervals (right)

Going a step further, we can slice $(0, t)$ in $(N+1)$ sub-intervals as depicted in Figure 9 (right); the density then writes:

$$\rho(\boldsymbol{x}, t) = \int \prod_{i=1}^{N} d^d\boldsymbol{x}_i \rho(\boldsymbol{x}_{i+1}, t_{i+1}|\boldsymbol{x}_i, t_i) = \int \prod_{i=1}^{N} \frac{d^d\boldsymbol{x}_i}{(4\pi D(t_{i+1}-t_i))^{d/2}} \exp\left(-\sum_{i=0}^{N} \frac{(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i)^2}{4D(t_{i+1}-t_i)}\right) \tag{5.9}$$

doing:

$$(t_{i+1}-t_i) \to 0, \quad \boldsymbol{x}_i = \boldsymbol{x}(i\Delta t) \quad \text{and} \quad \exp\left(-\sum_{i=0}^{N} \frac{(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i)^2}{4D(t_{i+1}-t_i)}\right) \to \exp\left(-\frac{1}{4D}\int_0^t d\tau \left(\frac{\partial \boldsymbol{x}}{\partial \tau}\right)^2\right) \tag{5.10}$$

and:

$$\prod_{i=1}^{N} \frac{d^d\boldsymbol{x}_i}{(4\pi D(t_{i+1}-t_i))^{d/2}} \to \mathcal{D}\boldsymbol{x}(\tau) \tag{5.11}$$

where we assume that $\mathcal{D}\boldsymbol{x}(\tau)$ is a measure on the space of paths, so we finally obtain:

$$\rho(\boldsymbol{x}, t) = \int_{\boldsymbol{x}(0)=\boldsymbol{0}}^{\boldsymbol{x}(t)=\boldsymbol{x}} \mathcal{D}\boldsymbol{x}(\tau) \exp\left(-\frac{1}{4D}\int_0^t d\tau \left(\frac{\partial \boldsymbol{x}}{\partial \tau}\right)^2\right) \tag{5.12}$$

The measure $\mathcal{D}\boldsymbol{x}(\tau)$ should be such that the path integral is normalized, in the sense that after any time t, the particle should be found somewhere:

$$\int d^d\boldsymbol{x} \rho(\boldsymbol{x}, t) = 1 \tag{5.13}$$

one can show that the measure $\mathcal{D}\boldsymbol{x}(\tau)$ defined by the limit in equation (5.11) satisfies this property.

5.2 Polymer models and path integrals

A polymer is a macro-molecule which is basically a chain of molecules called monomers, as depicted in Figure 10.

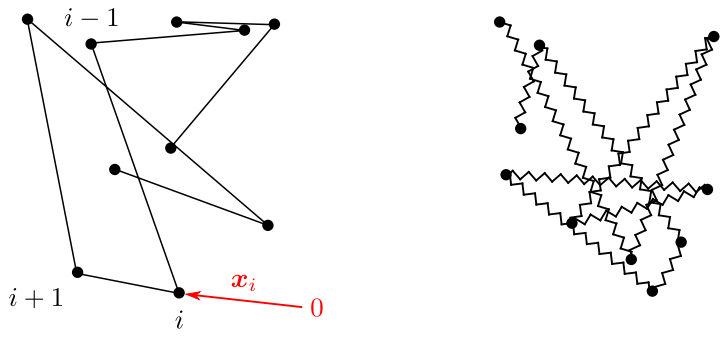

Figure 10: definition of a polymer, the junctions between the monomers are indexed by i and are located at $\boldsymbol{x}_i$ (left); a polymer with harmonic potential between consecutive junctions corresponds to Brownian motions (right)

Our goal when we study polymers is to understand their geometric structure and the universal large scale properties that emerge from their microscopic interactions. We will model them taking into account only interactions between nearest neighbors along the chain; indeed, molecular bonds are way stronger than electrostatic interactions.

5.2.1 Connection to random walks

The energy of a chain of $(N+1)$ monomers is:

$$E = \sum_{i=1}^{N} V(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i) \tag{5.14}$$

so the BOLTZMANN weight associated with a given configuration $(\boldsymbol{x}_1, \dots \boldsymbol{x}_N)$ of the polymer is simply:

$$\mathbb{P}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \frac{1}{Z} \exp\left(-\beta \sum_{i=1}^{N} V(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)\right) = \frac{1}{Z} \prod_{i=1}^{N} \exp\left(-\beta V(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)\right) \tag{5.15}$$

We define:

$$\psi(\boldsymbol{x} \to \boldsymbol{x}') = \frac{\exp\left(-\beta V(\boldsymbol{x}' - \boldsymbol{x})\right)}{\int d\boldsymbol{y} \exp\left(-\beta V(\boldsymbol{y})\right)} \tag{5.16}$$

this is the probability density associated to the transition from $\boldsymbol{x}$ to $\boldsymbol{x}'$, using this we can rewrite the probability measure as:

$$\mathbb{P}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \prod_{i=1}^{N} \psi(\boldsymbol{x}_i \to \boldsymbol{x}_{i+1}) \tag{5.17}$$

Here, the correspondence between i and Brownian motion's time is manifest, and from it we deduce some properties of the polymer, for instance:

$$\langle (\boldsymbol{x}_N - \boldsymbol{x}_0)^2 \rangle \propto N \tag{5.18}$$

Brownian chain: if we want to make the correspondence even more manifest we should consider a harmonic potential as depicted in Figure 10 (right), for instance the potential of the ROUSE model:

$$V(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i) = \frac{k_B T}{2l^2}(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)^2 \tag{5.19}$$

so that we recover the same kind of exponential as in equation (5.9) in the previous probability measure.

Brownian limit: if we now take $N \to \infty$ and $l \to 0$ with the scaling $Nl^2 = s$, where s is equivalent to time in the Brownian motion case, we obtain a probability weight for a configuration of the polymer which is exactly the same as in equation (5.12):

$$\mathbb{P}((\boldsymbol{x}(s'))_{0 \leq s' \leq s}) \propto \exp\left(-\frac{1}{2}\int_0^s ds' \left(\frac{\partial \boldsymbol{x}}{\partial s'}\right)^2\right) \tag{5.20}$$

where the configuration is defined as a function $\boldsymbol{x}(s')$, with s' going from 0 to s, as depicted in Figure 11.

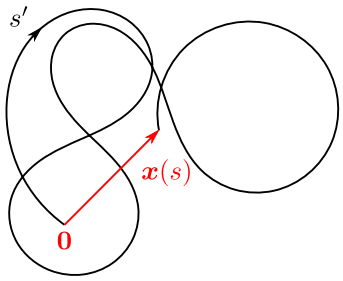

Figure 11: a polymer in the continuous limit

5.2.2 Path integral and FEYNMAN-KAC formula

Let us now define the partition function $Z(\boldsymbol{x}, s)$, for a polymer of length s attached between $\mathbf{0}$ and $\boldsymbol{x}$, as the following path integral:

$$Z(\boldsymbol{x}, s) = \int_{\boldsymbol{x}(0)=\mathbf{0}}^{\boldsymbol{x}(s)=\boldsymbol{x}} \mathcal{D}\boldsymbol{x}(s') \exp\left(-\frac{1}{2}\int_0^s ds' \left(\frac{\partial \boldsymbol{x}}{\partial s'}\right)^2\right) \tag{5.21}$$

where as in equation (5.13) the measure $\mathcal{D}\boldsymbol{x}(s')$ is such that the partition function is normalized:

$$\int d^d\boldsymbol{x} Z(\boldsymbol{x}, s) = 1 \tag{5.22}$$

This path integral is the same as the one that defined $\rho(\boldsymbol{x}, t)$ for the Brownian motion in equation (5.12), with the identification $s \sim t$ that was pointed out before. So this partition function must satisfy the same kind of equation as (5.5) with the same initial condition:

$$\partial_s Z = \triangle_{\boldsymbol{x}} Z \quad \text{and} \quad Z(\boldsymbol{x}, 0) = \delta^d(\boldsymbol{x}) \tag{5.23}$$

Now, let us suppose that the chain is subject to an additional external potential V, for instance gravity or a confining potential:

$$Z(\boldsymbol{x}, s) = \int_{\boldsymbol{x}(0)=\mathbf{0}}^{\boldsymbol{x}(s)=\boldsymbol{x}} \mathcal{D}\boldsymbol{x}(s') \exp\left(-\frac{1}{2}\int_0^s ds' \left(\frac{\partial \boldsymbol{x}}{\partial s'}\right)^2 - \int_0^s ds' V(\boldsymbol{x}(s'))\right) \tag{5.24}$$

then, going back to the the discrete using equations (5.10)-(5.11) in reverse, and denoting $s_{i+1} - s_i = \delta$:

$$Z(\boldsymbol{x}, s) = \lim_{\delta \to 0} \left(\int \prod_{i=1}^{N} \frac{d^d \boldsymbol{x}_i}{(2\pi\delta)^{d/2}} \exp\left(-\sum_{i=0}^{N} \frac{(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)^2}{2\delta} - \delta \sum_{i=0}^{N+1} V(\boldsymbol{x}_i) \right) \right) \tag{5.25}$$

and from (5.8):

$$Z(\boldsymbol{x} = \boldsymbol{x}_{N+1}, s) = \exp(-\delta V(\boldsymbol{x}_{N+1})) \int \frac{d^d \boldsymbol{x}_N}{(2\pi\delta)^{d/2}} \exp\left(-\frac{(\boldsymbol{x}_{N+1} - \boldsymbol{x}_N)^2}{2\delta} \right) Z(\boldsymbol{x}_N, s - \delta) \tag{5.26}$$

Let us introduce $\boldsymbol{y}$ such that : $(\boldsymbol{x}_{N+1} - \boldsymbol{x}_N)^2 + \delta \boldsymbol{y}^2$ and Taylor expand $Z(\boldsymbol{x}_N, s - \delta)$ around $(\boldsymbol{x}_{N+1}, s)$:

$$Z(\boldsymbol{x}_{N+1} - \delta^{1/2}\boldsymbol{y}, s - \delta) = Z(\boldsymbol{x}_{N+1}, s) - \delta^{1/2}\boldsymbol{y} \cdot \boldsymbol{\nabla} Z - \frac{\delta}{2} \sum_{i,j} y_i y_j \frac{\partial^2 Z}{\partial y_i \partial y_j} - \delta \frac{\partial Z}{\partial s} + \mathcal{O}(\delta^{3/2}) \tag{5.27}$$

so:

$$Z(\boldsymbol{x}_{N+1}, s) \approx (1 - \delta V(\boldsymbol{x}_{N+1})) \int \frac{d^d \boldsymbol{y}}{(2\pi)^{d/2}} \exp\left(-\frac{\boldsymbol{y}^2}{2} \right) \left(Z(\boldsymbol{x}_{N+1}, s) - \frac{\delta}{2} \sum_{i,j} y_i y_j \frac{\partial^2 Z}{\partial y_i \partial y_j} - \delta \frac{\partial Z}{\partial s} \right) \tag{5.28}$$

where the term $\delta^{1/2}\boldsymbol{y} \cdot \boldsymbol{\nabla} Z$ has vanished from the integral because it is an odd function of $\boldsymbol{y}$; for the same reason we have:

$$\int d^d \boldsymbol{y} \exp\left(-\frac{\boldsymbol{y}^2}{2} \right) y_i y_j \propto \delta_{ij} \tag{5.29}$$

finally, using:

$$\int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2} \right) = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} \frac{dy}{\sqrt{2\pi}} y^2 \exp\left(-\frac{y^2}{2} \right) = 1 \tag{5.30}$$

one obtain a partial differential equation which is very similar to equation (5.5) as was expected:

$$\partial_s Z = \frac{1}{2} \triangle_{\boldsymbol{x}} Z - V Z \tag{5.31}$$

the initial condition being $Z(\boldsymbol{x}, t = 0) = \delta^d(\boldsymbol{x})$. The equation (5.24) which gives its solution as a path integral is known as FEYNMAN-KAC formula. Note that this differential equation looks like SCHRÖDINGER's in imaginary time.

5.2.3 Quantum mechanics in imaginary time

Indeed, if we let $\hat{H}$ be the Hamiltonian associated to the SCHRÖDINGER equation in position basis:

$$\hat{H} = -\frac{1}{2} \triangle_{\boldsymbol{x}} + V \tag{5.32}$$

the precedent equation becomes directly:

$$-\partial_s Z = \hat{H} Z \tag{5.33}$$

and it is the same as SCHRÖDINGER's equation if we make the identification $s = it$, that is to say if we consider an imaginary time, because then $-\partial_s = i\partial_t$. This change of variable is often called a WICK rotation.

6 Phase transitions and collective phenomena

A phase transition is a dramatic change in the macroscopic behavior of a system which happens when a given external parameter (for instance the temperature) crosses some critical value. This dramatic change can either be abrupt (as water turning liquid above $0^o\ C$) or it can happen smoothly (as some magnetic materials going from paramagnetic to ferromagnetic above their Curie temperature); the distinction between the two is related to the order of the transition. A very wide range of systems are subject to this phenomenon in physics, but also beyond physics; as long as it is composed of many elements in interaction, any system can show this kind of collective behavior. Here we give some examples of subjects where phase transitions can appear :

- **In physics:** we mentioned solid-liquid and ferromagnetic-paramagnetic phase transitions, but there are many others, for instance fluids turn superfluid or materials become superconductors under a certain temperature.
- **In computer science:** in signal processing; for instance, if we want to denoise a signal, there is a transition between a phase where the noise is sufficiently low for the signal to be recovered and a phase where it cannot.
- **In social sciences:** they appear naturally in the study of crowds.
- **In mathematics:** they emerge in some combinatorics related problems.
- **In economy:** crashes are good examples of phase transitions and collective behaviors.
- **In biology:** in the study of flocks of birds or of colonies of bacteria.

In principle, we need nothing more than the law of interaction between the elements of the system to explain phase transitions, but they are often not easy to grasp directly. In the following, we will explain methods to study them.

6.1 Order of a phase transition (EHRENFEST)

The order of a phase transition is related to the analyticity of the free energy:

$$-\beta f = \frac{1}{N} \ln \left(\sum_{\mathbb{C}} \exp(-\beta E(\mathbb{C})) \right) \tag{6.1}$$

where we used the notation $\beta = 1/k_B T$ with k_B the Boltzmann constant. Here, $\mathbb{C}$ stands for a configuration of the system and $E(\mathbb{C})$ is the energy of this configuration. The free energy f can depend on multiple parameters that can be controlled in experiments, and which are appropriately called "control parameters", such as the temperature T.

Theorem : $f : T \mapsto f(T)$ is a continuous function of T.

6.1.1 First order phase transition

A transition at critical temperature T_c is said to be of first order if the derivative of $f(T)$ is discontinuous at T_c. Then, from the expression of the free energy:

$$f = e - Ts \tag{6.2}$$

we obtain directly the following expressions for entropy and internal energy:

$$s = -\frac{\partial f}{\partial T} \quad \text{and} \quad e = -T^2 \frac{\partial}{\partial T} \left(\frac{f}{T} \right) \tag{6.3}$$

the second one can be checked by developing the derivative. This shows that e and s are discontinuous at critical temperature for first order phase transitions.

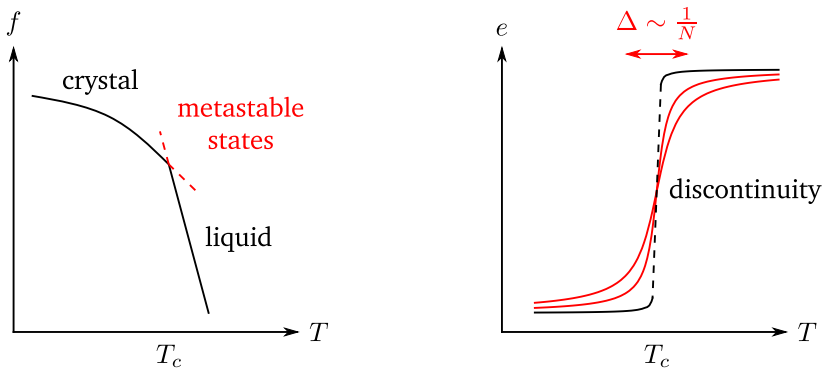

Figure 12: typical evolution of the free and internal energies around a phase transition

These discontinuities characterize first order phase transitions; they hint at sudden changes in the state of the system. One may argue that these discontinuities are unphysical. To be precise, they appear if we consider an infinite number of degrees of freedom. With a finite number of them, the functions are continuous, but variations happen on scales $\Delta \sim 1/N$; so with $N \sim 10^{23}$ they can be considered discontinuous for all practical purposes.

6.1.2 Second order phase transition

A transition is said to be of second order if the first derivatives of f are continuous while one of its second derivatives is singular. A good example of a second order phase transition is the ferromagnetic-paramagnetic phase transition.

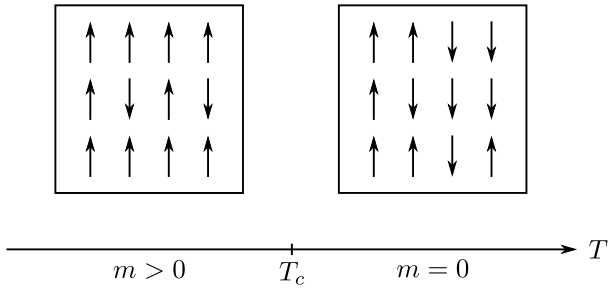

Figure 13: ferromagnetic-paramagnetic phase transition

This transition is well described by the ISING model. Let us consider a collection of interacting spins S_i which can take values ± 1 and are located on the sites of a cubic lattice indexed by i. To a given configuration $\{S_i\}$ we associate the following energy:

$$E(\{S_i\}) = -\sum_{\langle i,j \rangle} J S_i S_j - h \sum_i S_i \tag{6.4}$$

where the notation $\langle i, j \rangle$ stands for sites i and j which are closest neighbors on the cubic lattice. Let us also define the magnetization as:

$$m = \frac{1}{N} \sum_i S_i \tag{6.5}$$

its typical evolution during the phase transition is shown in figure 14.

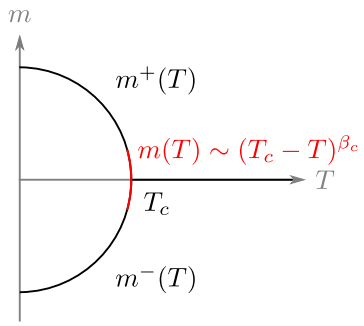

Figure 14: typical evolution of the magnetization around the ferromagnetic-paramagnetic phase transition

At the phase transition:

$$m(T) \sim (T_c - T)^{\beta_c} \tag{6.6}$$

where $\beta_c < 1$ is a critical exponent (one has to be careful not to confuse it with $\beta = 1/k_b T$). According to classical thermodynamics we have:

$$m = -\frac{1}{\beta}\frac{\partial}{\partial h}(\beta f) \tag{6.7}$$

so m is a first derivative of f, therefore it is continuous at a second order transition; but second derivatives are divergent, for instance:

$$\frac{dm}{dT} \quad \text{and} \quad \left.\frac{dm}{dh}\right|_{h=0} \tag{6.8}$$

are divergent quantities.

There are also higher order phase transitions, but they are harder to study experimentally. For third order transitions for instance, one has to fine-tune two parameters to get near the transition point, instead of one for second order.

6.2 Order parameter

An order parameter is an observable allowing to distinguish between the phases. As a simple example, the magnetization is the natural order parameter for the ferromagnetic-paramagnetic transition. Let us give another example; for the liquid-crystal transition, a good order parameter is the density field. Indeed, in FOURIER space:

$$\langle \tilde{\rho}(\boldsymbol{k}) \rangle = \tilde{\rho}_0 \delta(\boldsymbol{k}) \quad \text{in the liquid phase} \tag{6.9}$$

$$= \sum_{\boldsymbol{l} \in \text{ Bravais lattice}} \tilde{\rho}_{\boldsymbol{l}} \delta(\boldsymbol{k} - \boldsymbol{l}) \quad \text{in the crystal phase} \tag{6.10}$$

As a matter of fact, finding the good order parameter can sometimes be tough.

6.3 Symmetry breaking

A good example of symmetry breaking is once again given by the ferromagnetic-paramagnetic phase transition. Considering the ISING model already introduced in equation (6.4), with $h = 0$:

$$E(\{S_i\}) = -J \sum_{\langle i,j \rangle} S_i S_j \tag{6.11}$$

it is clear that this energy is invariant under a global spin flip, or $\mathbb{Z}_2$ symmetry:

$$\{S_1, ..., S_N\} \to \{-S_1, ..., -S_N\} \tag{6.12}$$

Nevertheless, after the transition to the ferromagnetic phase, this symmetry is broken: as we see in figure 13, in the case $T < T_c$, the spins have a privileged direction and the magnetization m is no longer equal to 0 as we would expect from symmetry invariance in equation (6.5). We can picture ourselves this symmetry breaking with the following cartoon:

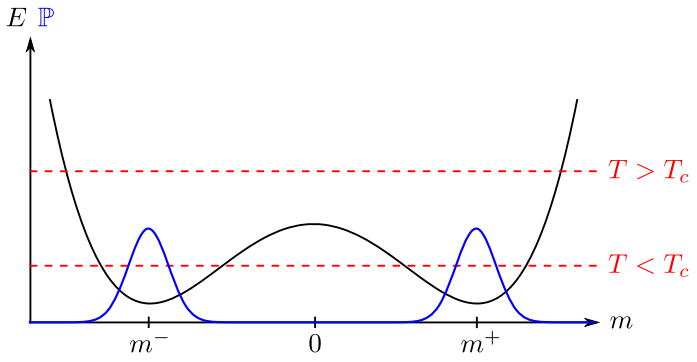

Figure 15: shape of the energy landscape of the ISING model without external field

For $T > T_c$, the system lives in a region where the magnetization can take arbitrary values that average out to 0 during observation time. But when $T < T_c$, the thermal fluctuations don't allow for the system to leave regions around the ground states where $m \neq 0$ (at least in a reasonable amount of time). In the process of going from above to below T_c, the system may choose between these regions arbitrarily, or according to uncontrolled perturbations. Notice that even if the value of m breaks it when going at low temperatures, the two ground states $\{+1, ..., +1\}$ and $\{-1, ..., -1\}$ are still related by the $\mathbb{Z}_2$ symmetry of the model; this is a general feature of symmetry breaking.

6.4 Long-range order

6.4.1 Example of the ferromagnetic and paramagnetic phases

In the ferromagnetic phase (that is $T < T_c$), if we measure the state of the system in one region, it can give us a decent amount of information about the state of the system far away from that region; while in the paramagnetic phase ($T > T_c$) there is no such long range order. The natural quantities to describe this characteristic are the correlations. Indeed, we can define long-range order by the condition:

$$\lim_{|x-y|\to\infty} \langle O_x O_y \rangle \neq 0 \tag{6.13}$$

where O_x can be any local observable at point x, for instance the value S_x of the spin at site x. In addition, if there is no long-range correlation, we have:

$$\langle S_x S_y \rangle = \langle S_x \rangle \langle S_y \rangle = m^2 = 0 \tag{6.14}$$

6.4.2 Example of the liquid and crystal phases

If we take a liquid in a recipient, deforming one side of the recipient does not necessarily affect the opposite side; while for a crystal, the deformation is carried all the way along in the structure, as depicted in figure 16.

Figure 16: the stiffness of liquid and crystal phases are different

In some sense, the stiffness is characterized by the order parameter. One can say that collective behavior makes the elements of the system act like an army, but without a general to guide them.

6.5 Universality and large scales properties

In the example of the ferromagnetic-paramagnetic phase transition, long range order starts to appear as soon as T approaches T_c. To be more precise, one can show that when approaching T_c from above, the correlation function behaves according to:

$$\langle S_x S_y \rangle \sim \frac{1}{|x-y|^{d-2+\eta}} \exp\left(-\frac{|x-y|}{\xi}\right), \quad T \to T_c^+ \tag{6.15}$$

here d is the dimension and ξ the correlation length; when we approach T_c, it behaves as:

$$\xi \sim \frac{1}{(T-T_c)^\nu}, \quad T \to T_c^+ \tag{6.16}$$

if $|x-y| \ll \xi$ (while being much bigger than the microscopic length a), what remains is:

$$\langle S_x S_y \rangle \sim \frac{1}{|x-y|^{d-2+\eta}}, \quad T \to T_c^+ \tag{6.17}$$

These kind of power laws are omnipresent in the study of systems close to their critical point. The exponents like ν and η appearing in them are called critical exponents. The precedent equations mean that when approaching T_c we obtain correlated regions of linear size ξ, inside which the correlations follow a power law, which is a much softer behavior than the decreasing exponential at higher temperature. The correlated regions are generally fractals of dimension η, which is called the anomalous dimension at critical point accordingly, as depicted on the figure:

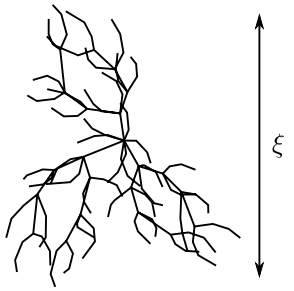

Figure 17: fractal region inside which the observables are correlated at critical point

Such structure can emerge even if it seems energetically inefficient, because at T_c, energy is balanced by entropy. The precedent discussion is not specific to the ferromagnetic-paramagnetic phase transition and S_x can be replaced

by any observable $\mathcal{O}_x$ in (6.15) and (6.17), depending on the system. The critical exponents ν and η may change. On large scales compared to the microscopic length a, the behaviors of fluctuations of the observables are universal; they do not depend on details in the microscopic properties such as the type of atom or interaction, but only on the kind of symmetry breaking. So in other words, systems that are in the same symmetry class share the same critical exponents. In this picture, the liquid-gas and ferromagnetic-paramagnetic transitions are found to be equivalent.

6.6 Techniques to study phase transitions and collective behavior

In what follows, we will introduce some of the main techniques that can be used to study phase transitions and collective behavior, which are the mean-field theory, the renormalization group theory and low-energy excitations. We will start by discussing the foundation of mean-field theory and how it can be used in computer science. Then we will introduce the (non)perturbative renormalization group theories. Finally, about low-energy excitations, we will discuss how interacting excitations destroy long-range order when the temperature increase and talk about the theory of topological excitations.

7 Mean-field theory

7.1 Mean field theory for the ISING model of ferromagnets

Let us consider once again the ISING model, with an energy given by (6.4):

$$E(\{S_i\}) = -\sum_{\langle i,j \rangle} J S_i S_j \tag{7.1}$$

where we consider $h = 0$. This energy can be rewritten to highlight that each spin is basically interacting with the external field created by its neighbors:

$$E(\{S_i\}) = -\sum_i h_i S_i \tag{7.2}$$

the instantaneous field h_i being of course:

$$h_i = \sum_{\langle i,j \rangle} J S_j \tag{7.3}$$

In principle, the local external field h_i is a fluctuating quantity. The basis of mean-field theory is to neglect its local fluctuations and to replace h_i by its spatial average $\langle h_i \rangle$. This approximation is better when each spin interacts -weakly- with a lot of others, because then h_i already looks like a spatial average. It is exact in infinite dimensions[7].

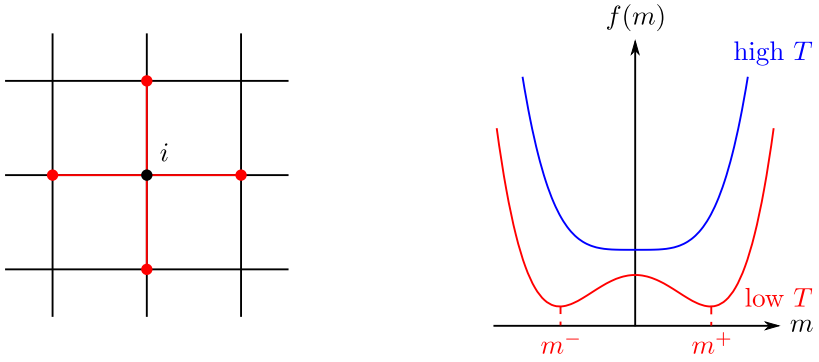

Figure 18: spin interacting with its nearest neighbors (left); free energy as a function of m close to $m = 0$ (right)

[7] in infinite dimensions, the local external field h_i takes into account the infinite number of spins surrounding the spin S_i, so it already integrates enough fluctuations to be equal to $\langle h_i \rangle$ without approximation

Using equation (7.2) with $h_i = \langle h_i \rangle$ we compute the magnetization $m_i = \langle S_i \rangle$; according to Boltzmann:

$$m_i = \sum_{S_i=\pm 1} S_i \exp(\beta \langle h_i \rangle S_i) \Big/ \sum_{S_i=\pm 1} \exp(\beta \langle h_i \rangle S_i) = \tanh(\beta \langle h_i \rangle) \tag{7.4}$$

then from (7.3) it is clear that $\langle h_i \rangle = Jzm_i$, where z is the connectivity of the lattice, that is to say the number of neighbors each spin has. Since m_i is actually supposed to be constant on the lattice, we obtain the following self-consistent relation:

$$\boxed{m = \tanh(\beta J z m)} \tag{7.5}$$

7.2 LANDAU theory

Here is a short introduction to LANDAU's theory for the ferromagnetic-paramagnetic phase transition. It starts by noticing that the order parameter m becomes $-m$ under the $\mathbb{Z}_2$ symmetry of the system. Then, we expand the free energy close to T_c^-:

$$f(m) = f(0) + f'(0)m + f''(0)\frac{m^2}{2} + f^{(3)}(0)\frac{m^3}{3!} + f^{(4)}(0)\frac{m^4}{4!} + \mathcal{O}(m^5) \tag{7.6}$$

this is legitimate since m is small close to T_c. Since under the $\mathbb{Z}_2$ symmetry, the free energy remains unchanged while m changes sign, we have $f'(0) = f^{(3)}(0) = 0$. Only the even derivative remain; and they may depend on T, it was implicit above :

$$f(m) = f_T(0) + f''_T(0)\frac{m^2}{2} + f^{(4)}_T(0)\frac{m^4}{4!} + \mathcal{O}(m^6) \tag{7.7}$$

For all temperatures, $f^{(4)}_T(0) > 0$; were it negative, the free energy would be unbounded from below, and the system would be driven to an inifinite value of magnetization, which is not physical. At high temperatures, $f''_T(0) > 0$, because the free energy has only one minimum, and at low temperatures $f''_T(0) < 0$, as depicted on figure 18 (right).

Now, let us hypothesize the following behaviors for $f''_T(0)$ and $f^{(4)}_T(0)$ at the phase transition:

$$f''_T(0) \sim \kappa(T - T_c) + \mathcal{O}((T - T_c)^2) \quad \text{and} \quad f^{(4)}_T(0) = 3!c + \mathcal{O}(T - T_c) \tag{7.8}$$

the second one simply means that $f^{(4)}_T(0)$ is independent of T. The free energy then rewrites:

$$f(m) = f(0) + \frac{\kappa}{2}(T - T_c)m^2 + \frac{c}{4}m^4 \tag{7.9}$$

and we find its minimum by solving:

$$\kappa(T - T_c)m + cm^3 = 0 \tag{7.10}$$

it gives $m = 0$ if $T > T_c$, and if $T < T_c$:

$$m = \left(\frac{\kappa}{c}\right)^{1/2} |T - T_c|^{1/2} \tag{7.11}$$

hence this very simple model allows to predict the value $\beta_c = 1/2$ for the critical exponent defined in equation (6.6). This prediction relies on finding a good order parameter, which is not always that easy as we already pointed out; here it is the scalar m, but in general it could be a more complicated object such as a tensor. We then exploited the symmetry. In the end, the only hypotheses we made is that f can be expanded and that its coefficients are regular.

7.3 BETHE approximation

7.3.1 General expression

Let us consider once more the ISING model:

$$E(\{S_i\}) = -\sum_{\langle i,j \rangle} J_{ij} S_i S_j \tag{7.12}$$

where as before $S_i = \pm 1$; but this time we allow the coupling J to depend on the considered pair of spins. The BETHE approximation consists in replacing the square lattice by a tree, as depicted here:

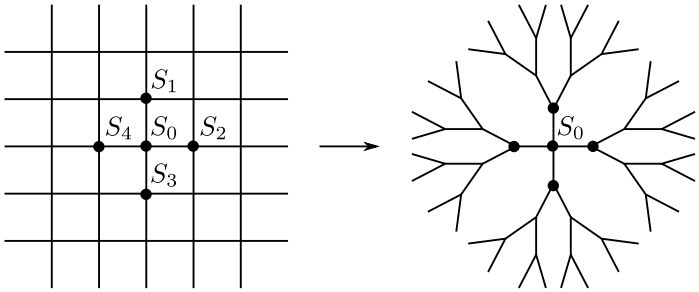

Figure 19: replacing the lattice around a point by trees rooted at that point

The main point of this approximation is that now, the spins connected to S_0 are not connected to each other; for instance there is no path going from S_1 to S_2. So the spin S_0 don't influence itself indirectly via a link between them. Now, let us define the probability $\mathbb{P}_{i\to j}$ as the one associated to the spin S_i when the link $i \leftrightarrow j$ is removed. We will obtain a closed set of equations for these probabilities. Let us start by computing $\mathbb{P}_{0\to 1}(S_0)$, the probability that the spin 0 has the value S_0 when we cut its link with the spin located at 1; according to BOLTZMANN:

$$\mathbb{P}_{0\to 1}(S_0) = \mathcal{N} \times \sum_{\{S_i, \{S_{i-}\}; i=2,3,4\}} \exp\left(-\beta \sum_{j=2,3,4} J_{0j} S_0 S_j - \beta H_{\text{rest}}[\{S_j, \{S_{j-}\}; j=2,3,4\}] \right) \tag{7.13}$$

where:

$$\frac{1}{\mathcal{N}} = \sum_{S_0, \{S_i, \{S_{i-}\}; i=2,3,4\}} \exp\left(-\beta \sum_{j=2,3,4} J_{0j} S_0 S_j - \beta H_{\text{rest}}[\{S_j, \{S_{j-}\}; j=2,3,4\}] \right) \tag{7.14}$$

here, $\{S_{i-}\}$ is a short notation for the tree of all the spins which are rooted at i, so $\{S_i, \{S_{i-}\}; i=2,3,4\}$ denotes the spins S_2, S_3, S_4 and all the spins in the trees starting from those three; and H_{rest} is the residual Hamiltonian containing all the links in the trees $\{S_{2-}\}$, $\{S_{3-}\}$ and $\{S_{4-}\}$. The $\mathcal{N}$ is here to normalize and contains a sum on S_0. Now, we can simplify a little bit this expression. The first thing to notice is that since the trees $\{S_{2-}\}$, $\{S_{3-}\}$ and $\{S_{4-}\}$ don't interact with each other, the Hamiltonian H_{rest} can be written as follows:

$$H_{\text{rest}}[\{S_j, \{S_{j-}\}; j=2,3,4\}] = \sum_{j=2,3,4} H_{j-}[\{S_j, \{S_{j-}\}] \tag{7.15}$$

as previously, H_{j-} is a notation for the Hamiltonian which concerns only the tree $\{S_{j-}\}$.

The expression of equation (7.13) can then be simplified a little factorizing the exponentials accordingly:

$$\mathbb{P}_{0\to 1}(S_0) = \mathcal{N} \sum_{S_2,S_3,S_4} \exp\left(-\beta \sum_{j=2,3,4} J_{0j} S_0 S_j\right) \prod_{i=2,3,4} \sum_{\{S_{i-}\}} \exp(-\beta H_{i-}[\{S_i, \{S_{i-}\}]) \tag{7.16}$$

note that the three terms of the product are different in the general case, but if the J_{ij} coefficients of equation (7.12) are chosen to be equal, the three trees are equivalent and the simplification can go further. The trick is then to extract the normalizations of the individual trees from $\mathcal{N}$, it is done replacing $\mathcal{N}$ by:

$$\mathcal{N}' = \mathcal{N} \prod_{i=2,3,4} \left(\sum_{S_i,\{S_{i-}\}} \exp(-\beta H_{i-}[\{S_i, \{S_{i-}\}]) \right) \tag{7.17}$$

then we remark:

$$\sum_{\{S_{i-}\}} \exp(-\beta H_{i-}[\{S_i, \{S_{i-}\}]) \Bigg/ \sum_{S_i,\{S_{i-}\}} \exp(-\beta H_{i-}[\{S_i, \{S_{i-}\}]) = \mathbb{P}_{i\to 0}(S_i) \tag{7.18}$$

and we obtain:

$$\boxed{\mathbb{P}_{0\to 1}(S_0) = \mathcal{N}' \prod_{i=2,3,4} \sum_{S_i} \exp(-\beta J_{0i} S_0 S_i) \mathbb{P}_{i\to 0}(S_i)} \tag{7.19}$$

where $1/\mathcal{N}'$ is the same thing with an additional sum on S_0, as was $1/\mathcal{N}$ in equation (7.14). As we claimed, we obtained a closed set of equations for the probabilities $\mathbb{P}_{i\to j}$; these equations are called BETHE equations, they can be written for all the sites of the lattice. This approximation is much more sophisticated than the "pure" Mean Field Theory we presented earlier. In principle here, the J_{ij} can vary and make these equations difficult to solve; however, there are numerical methods which can do it quite easily, and in some cases, even with random J_{ij}, an exact analytic treatment is possible. After we have solved these equations, the probability of the spin at 0 to be equal to S_0 writes as:

$$\mathbb{P}(S_0) = \mathcal{N} \prod_{i=1,\dots,4} \sum_{S_i} \exp(-\beta J_{0i} S_0 S_i) \mathbb{P}_{i\to 0}(S_i) \tag{7.20}$$

here $\mathcal{N}$ is not the same as in equation (7.14), but it is obtained the same way. These arguments could be generalized to any kind of variable other than the spins S_i; as long as these variables are connected by a lattice, this lattice can be approximated by a tree. It can be generalized to any kind of lattice as well, in any dimension. The main idea of the approximation is to neglect the correlations we are not able to compute between the spins $S_1, ..., S_4$ connected to S_0. Also, there are ways to control this approximation with correction terms (see KIKUCHI's approximation).

7.3.2 Application of the BETHE approximation to the ferromagnetic ISING model

Now, let us apply this approximation to the ferromagnetic ISING model, where $J_{ij} = -J$ for any i and j which are nearest neighbors on the lattice. Basically, we want to solve the set of BETHE equations (7.19) in that particular case. A big simplification comes from the fact that the system is uniform : all the spins are statistically equivalent and the probability $\mathbb{P}_{i\to j}(S_i)$ does not depend on $i \leftrightarrow j$. In other word, we are looking for a uniform solution $\bar{\mathbb{P}}(S_i)$ of the equations. We start by introducing a useful trick using that $S_i = \pm 1$:

$$\bar{\mathbb{P}}(S_i) = \frac{\exp(\beta h S_i)}{2\cosh(\beta h)} \tag{7.21}$$

we can write this since it satisfies $\bar{\mathbb{P}}(+1) + \bar{\mathbb{P}}(-1) = 1$; it comes down to write $\bar{\mathbb{P}}(+1)$ or $\bar{\mathbb{P}}(-1)$ as a function of h, which is interpreted as the effective field acting on S_i. Thus, equations (7.19) rewrite:

$$\bar{\mathbb{P}}(S_0) = \mathcal{N} \prod_{i=2,3,4} \sum_{S_i} \exp(\beta J S_0 S_i) \frac{\exp(\beta h S_i)}{2\cosh(\beta h)} \tag{7.22}$$

then, there are two other useful identities using the fact that $S_i = \pm 1$, which are:

$$\exp(\beta h S_i) = \cosh(\beta h) + S_i \sinh(\beta h) \tag{7.23}$$

and:

$$\exp(\beta J S_0 S_i) = \cosh(\beta J) + S_0 S_i \sinh(\beta J) \tag{7.24}$$

since they are obviously true for all the values that S_0 and S_i can take. Using them, we obtain:

$$\bar{\mathbb{P}}(S_0) = \mathcal{N}\left(\frac{1}{2}\sum_{S_1}(\cosh(\beta J) + S_0 S_1 \sinh(\beta J))(1 + S_1 \tanh(\beta h))\right)^{2d-1} \tag{7.25}$$

where d is the dimension of the space, it could be different than $d = 2$. Summing over $S_1 = \pm 1$ and after same algebra:

$$\begin{aligned}
\bar{\mathbb{P}}(S_0) &= \mathcal{N}(\cosh(\beta J) + S_0 \sinh(\beta J)\tanh(\beta h))^{2d-1} = \mathcal{N}'(1 + S_0 \tanh\beta J \tanh\beta h)^{2d-1} && (7.26)\\
&= \mathcal{N}''(\cosh(\tanh^{-1}(\tanh\beta J \tanh\beta h)) + S_0 \sinh(\tanh^{-1}(\tanh\beta J \tanh\beta h)))^{2d-1} && (7.27)\\
&= \mathcal{N}'' \exp((2d-1)\tanh^{-1}(\tanh\beta J \tanh\beta h) S_0) && (7.28)
\end{aligned}$$

here, we absorbed S_0-independent terms into redefinitions of the normalization constant $\mathcal{N}$; tracking these redifinitions, one can show that $\mathcal{N}'' = 1/2\cosh\beta h$, so that in the end we obtain the following auto coherence relation.

$$\boxed{\beta h = (2d-1)\tanh^{-1}(\tanh(\beta J)\tanh(\beta h))} \tag{7.29}$$

We obtained this simple prediction using BETHE approximation for the ferromagnetic ISING model, and we will see that it is much better than the prediction from basic MFT. This equation can be solved graphically as shown on the following figure:

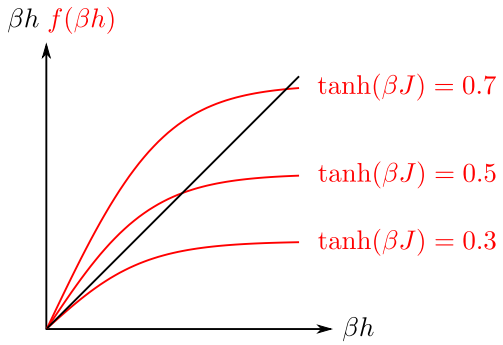

Figure 20: graphical resolution of equation (7.29), plotted with $d = 2$

On the graphical resolution of figure 20, it is clear that there is a critical value $\beta_c J$ under which $\beta h = 0$ is the only solution. To find this critical value, notice that $f(\beta h) = (2d-1)\tanh^{-1}(\tanh(\beta J)\tanh(\beta h))$ is a concave function. Its slope at the origin is $(2d-1)\tanh(\beta h)$, and it is easy to convince oneself that βh has a non trivial solution if and only if this slope is bigger than 1. This gives:

$$\boxed{(2d-1)\tanh(\beta_c J) = 1} \tag{7.30}$$

7.3.3 Some remarks on the BETHE approximation

At the qualitative level, the BETHE approximation predicts the same critical exponents as basic MFT when $d \to +\infty$. But at the quantitative level, it is much better. For instance, in the ferromagnetic ISING model, with $J = 1/2d$ (this amounts to a particular choice of units for the temperature, since what matters ultimately is always βJ) and $d = 3$, the standard MFT gives $\beta_c = 1$ after solving (7.5) as we did for (7.29). The BETHE approximation directly gives $\beta_c = 1.22$, which, although not perfect, is closer to the true value $\beta_c = 1.32$. The BETHE approximation is very good in any dimension, especially high ones; it becomes exact when the dimension goes to infinity and in one dimension.

In infinite dimensions the BETHE equations actually tend to Mean Field equations (this can be shown as exercise). For instance we can see in equation (7.30) that β_c becomes 1 when $d \to +\infty$, with the precedent choice of $J = 1/2d$.

When $d = 1$ the approximation is exact for the simple reason that the one dimensional lattice is already a tree. Contrary to the previously introduced MFT, which is wrong about that, the BETHE approximation does not predict a phase transition in one dimension; in fact, to be precise, it predicts that it would happen at zero temperature. This is easily shown using equation (7.30), because if $d = 1$, with $J = 1/2d$, it yields $\tanh(\beta_c/2) = 1$, which implies $\beta_c = +\infty$.

The BETHE approximation is of course exact on tree lattices and it is very good on lattices which are locally trees, for which loops are rare. If the lengths of the loops are big enough on average, the approximation becomes fantastic.

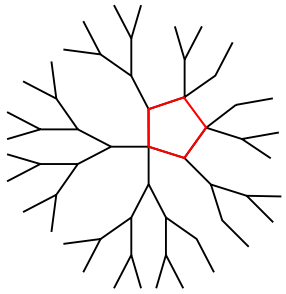

Figure 21: lattice which is locally tree

From this consideration on the role of loops in the approximation, we can understand why it works well in high d. Indeed, for a cubic lattice in dimension d, the loops only make for a tiny fraction of the possible paths; if we imagine we travel along the lattice, at each vertex we would have $2d$ edges to choose from and no chance to come back from where we started. This combines with the fact the coupling $1/2d$ becomes relatively small in higher dimension.

The loops give a criterion whether the approximation will be good or not; and the more dimensions, the less loops. So as pedagogical as it was to introduce it from a two dimensional square lattice, the BETHE approximation is not at its best. With loops of length four everywhere, it is close to be the worst case. As we mentioned earlier though, there are other fields of science where we can find collective behaviors and phase transitions; for instance, in biology and computer science, we find a lot of lattices with little loops where the approximation is used.

Let us introduce a problem of computer science where we can find a phase transition : the problem of inference. Imagine we have a signal $\{X_i^*; i = 1, ..., N\}$, called the ground truth, which is transmitted and corrupted by a noise. We receive the signal Y, which is a function of X^* and of that noise, and we want to recover the ground truth. There is a phase transition, because if the noise is higher than a threshold, the ground truth cannot be recovered and conversely. In the following part, we shall study this problem as an application of the BETHE approximation.

7.3.4 Application of the BETHE approximation to problem of inference

Link between this problem and statistical physics: let us consider a signal $\{X_i^*; i = 1, \dots, N\}$, called the ground truth, which is corrupted by a noise before we measure Y^μ. The index μ is here to take into account that we can measure multiple times the signal stemming from the same ground truth; we would obtain different outcomes, since the noise is random, but it would increase our knowledge of the ground truth. The situation is pictured here.

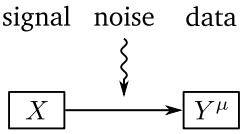

Figure 22: the problem of inference

For instance we can take:

$$Y_i = \lambda X_i^* + \xi_i \quad , \quad Y_{ij} = \lambda X_i^* X_j^* + \xi_{ij} \quad \text{or whatever} \tag{7.31}$$

where the ξ_i or ξ_{ij} would be Gaussian independent and identically distributed random variables, with $i \leq j$, and the λ is called the Signal to Noise Ratio (SNR). A possible approach for this problem is principal component analysis. Let us give another example to illustrate this approach; which has to do with the correlation between stock and financial markets. When the financial market evolves, the stock market follows it, moving along the principal mode. There are fluctuations and other modes of course, that one can try to understand to become somewhat richer. The point here it that we can use what we know *a priori* about the financial market to decipher the stock market.

In other parlance, this is a Bayesian approach. Let us call $\tilde{\mathbb{P}}(X^*)$ the probability *a priori*, before any measurement, that the ground truth is X^*. The determination of this probability depends on the nature of the signal; for instance with the precedent illustration, the probability *a priori* when we study the stock was given by the financial market. As another example, if the signal is a text in French, french words are more likely to be part of the ground truth.

From X^* we then get one or several signals Y^μ. If we know the noise well enough, we are able to compute $\mathbb{P}(Y|X)$, the probability that the Y we received stemmed from any signal X. In the end, we want to know what is the most probable X, given all the information contained in the copies Y^μ we received. Such probabilities of individual X are given by BAYES formula:

$$\mathbb{P}(X|\{Y^\mu\}) = \frac{\mathbb{P}(X, \{Y^\mu\})}{\mathbb{P}(\{Y^\nu\})} = \prod_\mu \frac{\mathbb{P}(Y^\mu|X)\tilde{\mathbb{P}}(X)}{\mathbb{P}(\{Y^\nu\})} \tag{7.32}$$

this can easily be rewritten:

$$\mathbb{P}(X|\{Y^\mu\}) = \frac{1}{Z} \exp\left(\ln \tilde{\mathbb{P}}(X) + \sum_\mu \ln \mathbb{P}(Y^\mu|X)\right) \tag{7.33}$$

where Z is a normalization constant, allowing the sum of the probabilities for all X to be one. In this expression, everything is known or can be guessed. In computer science, the number of variables N can be big; for instance these variables can be the color values of the pixels in an image, which are several millions. We talk about High Dimensional statistics, this is a limit equivalent to thermodynamics. In many cases, without enough information on the probability of *a priori*, we consider the X_i independent and make:

$$\tilde{\mathbb{P}}(X) = \prod_i \tilde{\mathbb{P}}(X_i) \tag{7.34}$$

and thus:

$$\mathbb{P}(X|\{Y^\mu\}) = \frac{1}{Z}\exp\left(\sum_i \ln\tilde{\mathbb{P}}(X_i) + \sum_\mu \ln\mathbb{P}(Y^\mu|X_1,\dots,X_N)\right) = \frac{1}{Z}\exp(-\beta H(X)) \tag{7.35}$$

with $\beta = 1$ and $H(X)$ can be seen as the sum of an external potential term and of interaction terms:

$$H(X) = \sum_i V(X_i) + \sum_\mu V_\mu(\{X\}) \tag{7.36}$$

Under these form, it becomes easier to understand how to apply what we know about thermodynamics. In the problem class, we are going to solve an actual problem of inference and see how statistical physics is related to it.

What is known about phase transitions and inference: one can introduce the following analog of the magnetization for this problem of inference:

$$m = \left\langle \sum_{i=1}^{N} \frac{X_i X_i^*}{N} \right\rangle \tag{7.37}$$

here, X is the signal we deciphered from the Y^μ we received and X^* is the ground truth we were looking for. If the X and X^* are too different, then $m \approx 0$, but if they are close to each other $m \approx 1$. As we mentioned earlier, there is a phase transition for a critical value of the Signal to Noise Ratio; if $\lambda < \lambda_c$ then $m = 0$ whatever we try and we say that the inference is impossible.

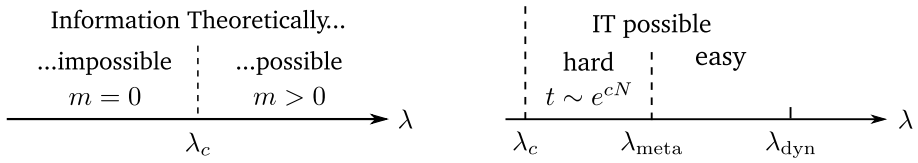

Figure 23: (left) phase transition in inference; (right) first order phase transition

In this context, the BETHE approximation is called belief propagation. A lot of work has been done between the 80's and the year 2000 to understand the phase diagrams of inference problems. As in physics, there are cases where the phase transition between possible and impossible inference is second order, and other cases where it is first order. Now, scientists are studying the dynamics of these transitions. When the transition is second order, the critical λ_c can be approached easily, in the sense that algorithms trying to decipher the signal can do it for λ close to λ_c. When the transition is first order though, the inference is made harder by meta stable equilibrium: there is a dynamical value λ_{dyn}, which is algorithm-dependent, under which metastablility can appear. Then, even though the inference is still possible, it can take an exponential amount of computation time. Finding algorithms with the lower λ_{dyn} possible is an open topic of research.

Other occurrences of phase transitions in computer science

Compressed sensing and denoising: for instance to recover an image corrupted by noise. Using belief propagation, CANDES, TAO and DONOHO showed in 2006 that if noise is not too strong, the image can be recovered. As another example, sometimes in physics one need to introduce a seeds to understand how metastable states break, for instance clouds would not form without the presence of dust in the air. This idea of nucleation seeds can also be found in computer science, in the work of KRZAKALA, MEZARD, SAUSSET, SUN and ZDEBOROVA among others.

Error correction: as used for instance to recover audio signals.

8 Introduction to the Renormalization Group. Real space RG.

Let us summarize what we have learned so far about the applicability of Mean Field Theory as a function of the dimension. First, MFT is exact in infinite dimensions. It is qualitatively correct above a critical dimension d_u, in the sense that it gets critical exponents and other properties of the phase transition right; for the ISING model, $d_u = 4$. When $d = 1$ for the ISING model, there is no phase transition, as captured correctly by the BETHE approximation. Between a lower critical dimension d_l and d_u, MFT is not correct; at best it gives a rough description of the phase transition. This is why new methods are needed for dimensions between lower critical d_l and upper critical d_u.

Historically, when Mean Field Theory was compared to experiments and simulations, it looked really great. But when ONSAGER obtained the exact solution for the two dimensional ISING model, MFT revealed its imperfections. Many other examples highlighting the limits of MFT followed. In the 70's, a new method emerged, called the Renormalization Group, which was developed in parallel for high energy and statistical physics. It was a revolution. This is the reason why the study of the Renormalization Group earned Kenneth WILSON a NOBEL prize in 1982.

8.1 Coarse-graining and RG transformation

The idea of RG is to link different scales of the theory. To illustrate how it works, let us consider the two dimensional ISING model in a square lattice with a linear distance between sites l:

$$H = -J \sum_{\langle i,j \rangle} S_i S_j \tag{8.1}$$

As pictured in the figure, we perform the transformation of the lattice which consists in grouping spins with their neighbors and removing one out of two, creating a dual lattice with spins S'_i and linear distance between sites $\sqrt{2}l$.

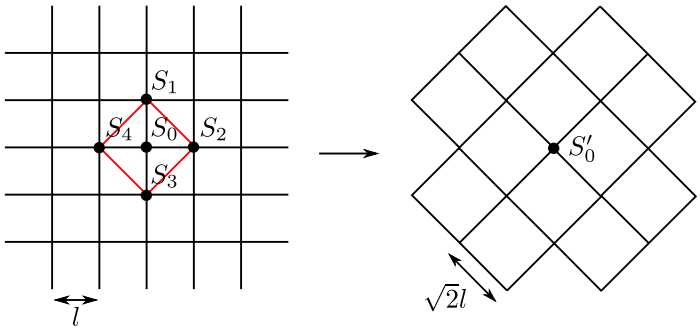

Figure 24: transformation obtained by grouping the spins with their nearest neighbors and removing one out of two

Where for S'_0 we choose for instance:

$$S'_0 = \text{sign}\left[\sum_{i=0,\dots 4} S_i\right] \tag{8.2}$$

in a sense, this is the average of the adjacent and interacting spins of the initial lattice, which belongs to a new lattice containing only "coarse-grained" spins. More generally we take:

$$S'_{\bar{x}} = \text{sign}\left[\sum_{x \in B_{\bar{x}}} S_x\right] \tag{8.3}$$

here, $\bar{x}$ is a site in the new lattice and $B_{\bar{x}}$ the "ball" around $\bar{x}$ in the initial, microscopic lattice. Usually, we say that S_x is the microscopic description while $S_{\bar{x}}$ is the coarse-grain description.
This transformation not only changes the lattice microscopic length scale l, but it affects the Hamiltonian $H[\{S_x\}]$. Let us show it by looking at the partition function of the system; initially:

$$Z = \sum_{\{S_x\}} \exp(-\beta H[\{S_x\}]) \tag{8.4}$$

in the sum, we can insert the quantity:

$$\sum_{\{S'_{\bar{x}}\}} \prod_{\bar{x}} \delta \left(S'_{\bar{x}}, \text{sign} \left[\sum_{x \in B_{\bar{x}}} S_x \right] \right) \tag{8.5}$$

which is equal to 1; it is the number of lattices we can define from a given lattice using the procedure described by figure 24 and equation (8.3). Doing so and inverting the sums allows us to obtain the new Hamiltonian:

$$Z = \sum_{\{S'_{\bar{x}}\}} \sum_{\{S_x\}} \exp(-\beta H[\{S_x\}]) \prod_{\bar{x}} \delta \left(S'_{\bar{x}}, \text{sign} \left[\sum_{x \in B_{\bar{x}}} S_x \right] \right) = \sum_{\{S'_{\bar{x}}\}} \exp(-\beta H'[\{S'_{\bar{x}}\}]) \tag{8.6}$$

Now let us formulate some remarks. First, we notice that the new Hamiltonian is still describes an ISING model, in the sense that $S'_{\bar{x}} = \pm 1$. Also, this new model still lives on a square lattice, with a change of scale $l \to \sqrt{2}l$. There is the apparition of next-to-nearest neighbor interactions under this transformation though:

$$H'[\{S'_{\bar{x}}\}] = -\sum_{\langle i,j \rangle} J_1(J) S'_i S'_j - \sum_{\langle i,j,k,l \rangle} J_2(J) S'_i S'_j S'_k S'_l \tag{8.7}$$

If we iterate this transformation, called the renormalization group transformation, we will create all the couplings allowed by the symmetry; for instance here, there will be no term with three spins, as it would violate the $\mathbb{Z}_2$ symmetry. An important question to address then is how the couplings transform under the renormalization group.

$$\{J_1, J_2, \ldots\} \longrightarrow \{J'_1, J'_2, \ldots\} \tag{8.8}$$

To answer it, is better to perform continuous renormalization group transformation $l \to (1+\varepsilon)l$ with ε very small, and study what is called the renormalization flow. To do this, we will work in FOURIER space, as we will see later. Finally, note that the "renormalization group" is not a group because it has not inverse; instead, it is a semi-group. It is not important because in general, we want to escape the microscopic scale and only go towards higher ones.

8.2 RG flow and fixed points

Let us now consider the space of all possible Hamiltonians respecting the symmetry of the system. We can represent any of these models by a set of couplings which may depend on the temperature, as shown in the following figure.

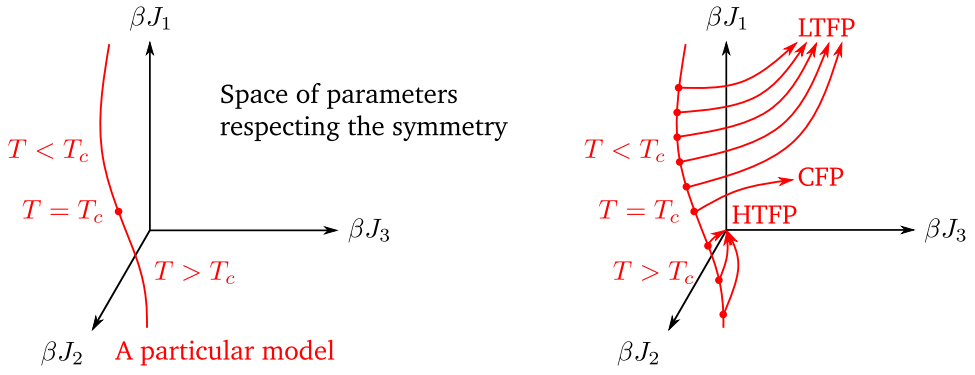

Figure 25: (left) representation of a model in the space of couplings; (right) RG flow of this model

In the representation of figure 25 (left), two different models, or Hamiltonians, correspond to two different curves. Note that we consider models having temperature as the only control parameters, but there can be others of course. Starting from a particular point of a given model and iterating many renormalization group transformations, we always arrive to a fixed point. This is a fundamental hypothesis, which is verified in all reasonably physical system.

- If $T > T_c$, the correlation length $\xi(T)$ is finite. After many Renormalization Group transformations $l \to (1+\varepsilon)l$, we arrive at a scale which is larger than $\xi(T)$ and a decoupling occurs. Therefore, the system flows to a High-Temperature Fixed Point (HTFP) corresponding to a non-interacting model for which all the couplings are 0.

- If $T < T_c$, at small scale, in the example of the ISING model, the spins are almost all in the same direction. Then, when we increase the scale using RG transformations, there are more parallel spins. In the end, the system is ordered at large scales and we do not see the fluctuations anymore as if the couplings were strong. Therefore, the model flows towards a Low-Temperature Fixed Point (LTFP), which is strongly interacting.

- If $T = T_c$, the correlation length becomes infinite and the model flows towards the Critical Fixed Point (CFP)

Our precedent considerations on the Renormalization Group flow are summarized in figure 25 (right). There is some kind of universal behavior under the Renormalization Group: all microscopic models which have the same kind of phase transition in terms of symmetry breaking will flow to the same Critical Fixed Point for $T = T_c$. The reason behind universality is that all such systems have their critical temperature on the same critical manifold, which is an hypersurface in the space of parameters having all of its points flowing towards the CFP. Here, we assume there is an unique fixed point on this manifold, but in principle there can be others, yielding multiple classes of models. Characterizing the CFP gives the critical exponents and the large scale behavior of all the corresponding models.

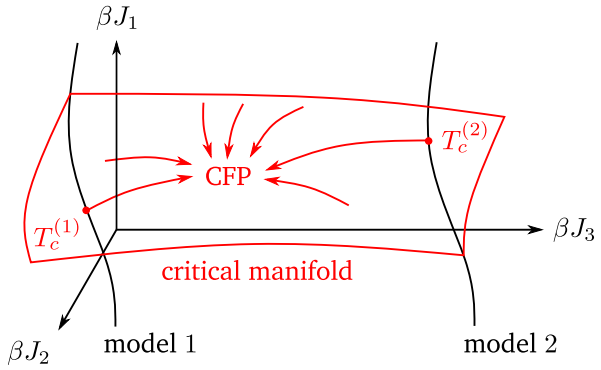

Figure 26: critical manifold shared by two models belonging to the same class

8.3 The β-function

Let us now study how the couplings change under RG transformations:

$$J^{(i)}(l) \to J^{(i)}((1+\varepsilon)l) \tag{8.9}$$

If ε is small enough, then:

$$J^{(i)}((1+\varepsilon)l) - J^{(i)}(l) = \beta^{(i)}(\{J(l)\})\varepsilon + \mathcal{O}(\varepsilon^2) \tag{8.10}$$

and by TAYLOR expansion:

$$\frac{dJ^{(i)}(l)}{dl} \times \varepsilon l = \beta^{(i)}(\{J(l)\})\varepsilon + \mathcal{O}(\varepsilon^2) \tag{8.11}$$

therefore when $\varepsilon \to 0$:

$$\boxed{l \times \frac{dJ^{(i)}(l)}{dl} = \beta^{(i)}(\{J(l)\})} \tag{8.12}$$

The critical fixed point is defined by a set of couplings $J_*^{(1)}, J_*^{(2)}, \dots$ such that for all i, we have $\beta^{(i)}(\{J_*^{(k)}\}_k) = 0$. Close to the critical temperature, the correlation length ξ, behaving as $(T - T_c)^{-\nu}$ with $\nu > 0$, is much bigger than the length scale a of the microscopic model. At intermediate scales $a \ll l \ll \xi$, the microscopic details don't matter, but in a sense the scale is not yet large enough to see that ξ is not infinite, and we cannot realize we are not at T_c.

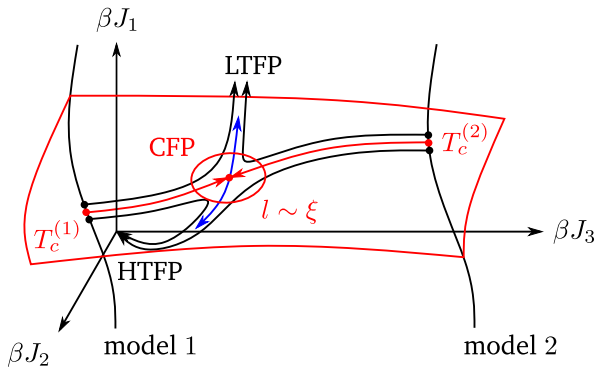

Figure 27: RG flow of the model close to the critical temperature

So when $T \approx T_c$, the model starts flowing like the one at T_c, until it arrives at length scales of the order of ξ, as shown in the precedent figure. The closer T is to T_c, the closer the model will approach the CFP. Let us develop the RG equations (8.12); close to the CFP:

$$J^{(i)}(l) = J_*^{(i)} + \delta J^{(i)}(l) \tag{8.13}$$

equation (8.12) for the perturbation $\delta J^{(i)}(l)$ around the CFP defined by the $J_*^{(i)}$ is:

$$l \times \frac{d\delta J^{(i)}(l)}{dl} = \sum_k \frac{d\beta^{(i)}(\{J\})}{dJ^{(k)}} \times \delta J^{(k)}(l) + \cdots \tag{8.14}$$

which can be reformulated as:

$$\boxed{l \times \frac{d\delta J^{(i)}(l)}{dl} = \sum_k M_{ik} \delta J^{(k)}(l)} \tag{8.15}$$

Now, let the $\{w_\alpha\}$ be the left eigenvectors of M, associated to the eigenvalues $\{\lambda_\alpha\}$, and let us define the $\delta J^{(\alpha)}(l)$ in the basis of these eigenvectors:

$$\delta J^{(\alpha)}(l) := \sum_i \delta J^{(i)}(l) w_\alpha^i \tag{8.16}$$

inserting this into (8.15):

$$\sum_i w_\alpha^i l \times \frac{d\delta J^{(i)}(l)}{dl} = \sum_i w_\alpha^i \sum_k M_{ik} \delta J^{(k)}(l) \tag{8.17}$$

we obtain for the $\delta J^{(\alpha)}(l)$:

$$l \times \frac{d\delta J^{(\alpha)}(l)}{dl} = \lambda_\alpha \delta J^{(\alpha)}(l) \tag{8.18}$$

and it is solved directly as:

$$\boxed{J^{(\alpha)}(l) = l^{\lambda_\alpha} \delta J^{(\alpha)}(1)} \tag{8.19}$$

In this equation, we see that if $\lambda_\alpha > 0$ then $\delta J^{(\alpha)}$ increases; in this case the perturbation is unstable and we talk about a relevant coupling. If $\lambda_\alpha < 0$ then $\delta J^{(\alpha)}$ decreases, the direction is stable, we talk about irrelevant coupling.

In zero field, there is only one $\lambda_{\bar{\alpha}} > 0$. We can ask how "long" does it takes following the flow to go significantly away from the CFP. Let us say that $T = T_c + \varepsilon$ and that initially $\delta J^{(\alpha)}(1) \sim \varepsilon$, so that $\delta J^{(\bar{\alpha})}(l) = l^{\lambda_{\bar{\alpha}}}\varepsilon$. Then, we have flowed significantly away from the CFP when $\delta J^{(\bar{\alpha})}(l) \sim \mathcal{O}(1)$, that is to say $l \sim \varepsilon^{-1/\lambda_{\bar{\alpha}}}$. In this case, we argued previously that $l \sim \xi$; and since $\xi \sim (T - T_c)^{-\nu}$ we can identify:

$$\boxed{\nu = \frac{1}{\lambda_{\bar{\alpha}}}} \tag{8.20}$$

To summarize, from the β-functions $\beta^{(i)}$, we obtain the couplings of the critical theory $J_*^{(i)}$ such that $\beta^{(i)}(\{J_*\}) = 0$. Then, the derivatives of the β-functions with respects to the couplings and evaluated at the CFP, which enter in the definition of the previous matrix M, yield the critical exponents (ν, etc.). The matrix M is a property of the CFP. Here, we argued with only one relevant direction, but for instance in the ISING model with a magnetic field, there would be two relevant directions λ_1 and λ_2; from them and scaling relations, one would recover all the exponents.

8.4 Real space RG

Let us apply the decimation procedure of figure 24 to the one dimensional ISING model, which is easier to solve.

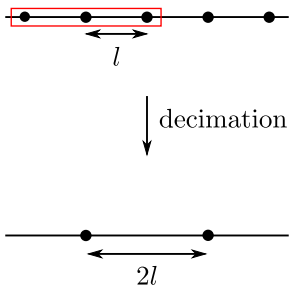

Figure 28: same transformation as in figure 24 but for one dimension

Separating the spins into odd and even, the partition function can be decomposed as follows:

$$Z = \sum_{\{S_i\}} \exp\left(-\beta H[\{S_i\}]\right) = \sum_{\{S_i, i \text{ even}\}} \sum_{\{S_i, i \text{ odd}\}} \exp\left(-\beta H[\{S_i\}]\right) \tag{8.21}$$

$$= \sum_{\{S_i, i \text{ even}\}} \left[\left(\sum_{S_1=\pm 1} \exp(\beta J S_0 S_1 + \beta J S_1 S_2) \right) \times \left(\sum_{S_3=\pm 1} \cdots \right) \right] \tag{8.22}$$

then

$$\exp(\beta J S_0 S_1) \exp(\beta J S_1 S_2) = (\cosh(\beta J) + S_0 S_1 \sinh(\beta J))(\cosh(\beta J) + S_1 S_2 \sinh(\beta J)) \tag{8.23}$$

we have:

$$\sum_{S_1=\pm 1} \exp(\beta J S_0 S_1) \exp(\beta J S_1 S_2) = \mathcal{N} \exp(S_0 S_2 \tanh^{-1}((\tanh(\beta J))^2)) \tag{8.24}$$

and finally:

$$Z \propto \sum_{\{S_i, i \text{ even}\}} \exp(\beta J'(J)(S_0 S_2 + S_2 S_4 + \cdots)) \tag{8.25}$$

here, $J'(J)$ is given by:

$$\beta J' = \tanh^{-1}\left((\tanh(\beta J))^2\right) \tag{8.26}$$

In one dimension, the RG transformation does not generate new couplings and the model can be solved exactly.

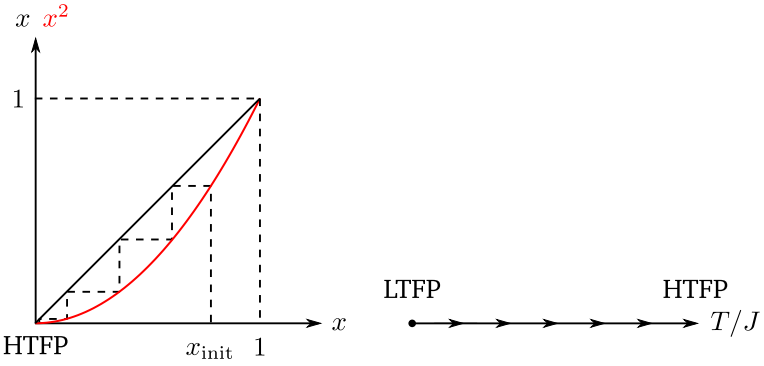

Figure 29: (left) resolution of $x' = x^2$ with $x = \tanh(\beta J)$; (right) flow to the HTFP

If the temperature is non zero, β is finite, and $x_{\text{init}} < 1$, where $x = \tanh(\beta J)$. As shown in the precedent figure, the model flows to a HTFP defined by $x = 0$, that is to say $\beta J_{\text{eff}} = 0$. So in the RG description of the ISING model, there is no finite-temperature transition in one dimension. For dimensions bigger than one, the RG transformations generate all the couplings allowed by the symmetry as we have seen previously. One solution to ease the study is to make a truncation and consider only a finite set of couplings; this is what we will do in the problem class.

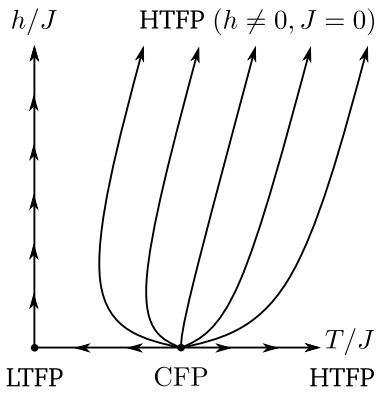

Figure 30: RG flow of the two dimensional ISING model in the truncated $h - J$ plane

The procedure we showed here is called the real space RG, because we performed a transformation directly on the lattice, which lives in the real space, in opposition to the FOURIER space for instance. It has some advantages. First, it is really simple and intuitive, and it is a good way to introduce the RG flow. Moreover, it is not limited to simple cases; it can be applied to the study of non-perturbative fixed points and disordered systems for which the J_{ij} are random, and in fact it is the only method that allows to do flows of functions. But this procedure is not improvable in a controlled way, and it does not allow one to obtain the upper critical dimension for which the MFT is right.

As we said in the introduction, the RG allows one to establish a connection between lengthscales; this is extremely useful in Quantum Field Theory, where the couplings, depend on the energy scale at which they are probed.

9 Non-perturbative and functional Renormalization Group

9.1 Tools

9.1.1 Continuum model

Let us consider a continuum model whose state is described by the function φ; its partition function writes:

$$Z = \int \mathcal{D}[\varphi] \exp(-S[\varphi]) \tag{9.1}$$

where $x \in \mathbb{R}^d$; the action $S[\varphi]$ is:

$$S[\varphi] = \int d^d x \left(\frac{(\nabla \varphi)^2}{2} + V[\varphi(x)] \right) \tag{9.2}$$

here, $V[\varphi]$ depends on the T:

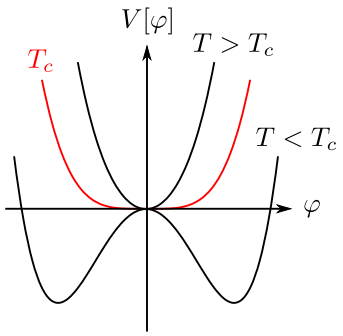

Figure 31: shape of $V[\varphi]$ as a function of T

This potential enjoys the $\mathbb{Z}_2$ symmetry since $V(\varphi) = V(-\varphi)$; it also induces a symmetry breaking under critical T_c. The CFP of this model will be similar to the one of the ISING model which has the same kind of symmetry breaking.

9.1.2 Free-energy $\Gamma[M]$ and LEGENDRE transform

Now, we define:

$$W[H] = \ln \left(\int \mathcal{D}[\varphi] \exp \left(-S[\varphi] + \int d^d x H(x) \varphi(x) \right) \right) \tag{9.3}$$

for $H(x)$ any external field. The LEGENDRE transform $\Gamma[M]$ is then defined as follows:

$$\Gamma[M] \coloneqq -W[H] + \int d^d x H(x) M(x) \tag{9.4}$$

here, $H(x)$ is a functional of $M(x)$:

$$\frac{\delta W}{\delta H}[H(x)] = M(x) \tag{9.5}$$

and vice versa under the exchange of $H(x)$ and $M(x)$. Putting it differently:

$$\frac{\delta W}{\delta H(x)} = \langle \varphi(x) \rangle = M(x) \tag{9.6}$$

The quantity $\Gamma(x)$ is called the free energy and $M(x)$ is the magnetization in analogy with what we presented previously. Since it is positive definite, Γ is convex. One can also use the general LEGENDRE-FENCHEL transform.

9.1.3 Properties of $\Gamma[M]$

Let us first compute:

$$\frac{\delta\Gamma}{\delta M(x)} = -\int d^d y \frac{\delta W}{\delta H(y)} \frac{\delta H(y)}{\delta M(x)} + \int d^d y \frac{\delta H(y)}{\delta M(x)} M(y) + H(x) \tag{9.7}$$

using identity (9.6), we replace the first differential of the first integral by $M(y)$; then, we observe that the two integral cancel each other, leaving only:

$$\boxed{\frac{\delta\Gamma[M]}{\delta M(x)} = H(x)} \tag{9.8}$$

A second derivative yields the Hessian:

$$\frac{\delta^2\Gamma[M]}{\delta M(x)\delta M(y)} = \frac{\delta H(x)}{\delta M(y)} \tag{9.9}$$

Let us give precision about the right hand side of this precedent equation; using that $M(x) = \langle\varphi(x)\rangle$ from (9.6):

$$\frac{\delta M(y)}{\delta H(x)} = \frac{\delta}{\delta H(x)} \left(\int \mathcal{D}[\varphi]\varphi(y) \exp\left(-S[\varphi] + \int d^d z H(z)\varphi(z)\right) \Big/ \int \mathcal{D}[\varphi] \exp\left(-S[\varphi] + \int d^d z H(z)\varphi(z)\right)\right) \tag{9.10}$$

so:

$$\frac{\delta M(y)}{\delta H(x)} = \langle\varphi(x)\varphi(y)\rangle - \langle\varphi(x)\rangle\langle\varphi(y)\rangle = \langle\varphi(x)\varphi(y)\rangle_c \tag{9.11}$$

the subscript c stands for connected correlator. Then, $\delta M(x)/\delta H(y)$ is the inverse of $\delta H(y)/\delta M(x)$; for x, x':

$$\int d^d y \frac{\delta H(x)}{\delta M(y)} \frac{\delta M(y)}{\delta H(x')} = \frac{\delta H(x)}{\delta H(x')} = \delta^{(d)}(x - x') \tag{9.12}$$

as a consequence, we obtain:

$$\frac{\delta^2\Gamma[M]}{\delta M(x)\delta M(y)} = \langle\varphi(x)\varphi(y)\rangle_c^{-1} \tag{9.13}$$

So Γ is interesting for two things: it provides a relation between $H(x)$ and $M(x)$ and the correlation function.

9.2 Non-perturbative Renormalization Group (Polchinsky and Wetterich)

9.2.1 Free-energy

Let us start from:

$$Z_k = \int \mathcal{D}[\varphi] \exp\left(-S[\varphi] - \frac{1}{2}\int d^d x d^d y \varphi(x) R_k(x-y)\varphi(y)\right) \tag{9.14}$$

here, R_k is a regulator, allowing to have coarse grain on length-scales $l \sim 1/k$; indeed the term $\varphi(x)R_k(x-y)\varphi(y)$ is the Fourier transform of $|\hat{\varphi}(q)|^2\hat{R}_k(q)$ where $\hat{R}_k$ behaves as depicted in the following figure:

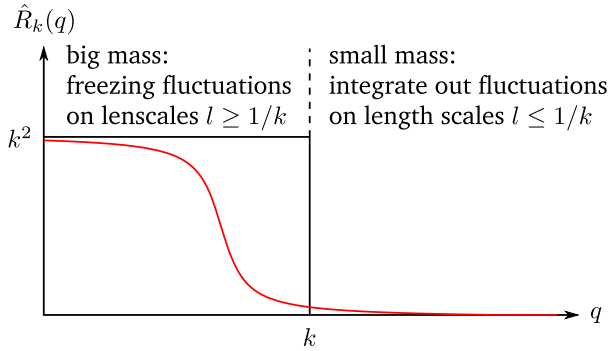

Figure 32: behavior of the Fourier transformed regulator $\hat{R}_k$

The fluctuations on small length-scales have no effect: they are free. The "big mass" part of the regulator corresponds to frozen fluctuations on length-scales $l \geq 1/k$, whereas on small length-scales $l \leq 1/k$ fluctuations are integrated out. The idea behind non-perturbative renormalization group is to not integrate the frozen fluctuations.

Let us give the LEGENDRE transform $\Gamma_k^L[M]$ which is associated to Z_k; from the definitions of equations (9.3) and (9.4), the term proportional to $H(x)M(x)$ being inserted in the path integral because it does not depend on φ:

$$\Gamma_k^L[M] = -\ln\left(\int \mathcal{D}[\varphi]\exp\left(-S[\varphi] + \int d^dx H(x)(\varphi(x) - M(x)) - \frac{1}{2}\int d^dx d^dy \varphi(x) R_k(x-y)\varphi(y)\right)\right) \tag{9.15}$$

now we study the quantity:

$$\boxed{\Gamma_k[M] = \Gamma_k^L[M] - \frac{1}{2}\int d^dx d^dy M(x) R_k(x-y) M(y)} \tag{9.16}$$

which is the free-energy functional as a function of $M(x)$ on length-scales $\sim 1/k$. The idea here is that by adding R_k to the partition function, we changed the problem, and we have to compensate it after having done the LEGENDRE transform. In short, we integrate out the microscopic fluctuations up to $1/k$ using R_k, then we make a correction.

9.2.2 Properties of $\Gamma_k[M]$

The macroscopic side of the theory corresponds to large length-scales, so $k = 0$. As we can see on figure 32, this implies $R_k(q) = 0$. So $\Gamma_{k=0}[M]$ is the true free-energy of the system, because in this case R_k vanishes from Z.

In the microscopic side, corresponding to $k \to \infty$, we show that $\Gamma_k[M] = S[M]$. In this case, we look at infinitesimal fluctuations, none of them being integrated out. Let us start from the identity (9.6) applied to $\Gamma_k^L[M]$ and the definition of $\Gamma_k[M]$:

$$H(x) = \frac{\delta\Gamma_k^L[M]}{\delta M(x)} = \frac{\delta\Gamma_k[M]}{\delta M(x)} + \int d^dy R_k(x-y)M(y) \tag{9.17}$$

we can develop $\Gamma_k[M]$ from its definition, inserting the term proportional to $M(x)R_k(x-y)M(y)$ in the integral as previously:

$$\begin{aligned}
\Gamma_k[M] &= -\ln\left(\int \mathcal{D}[\varphi]\exp\left(-S[\varphi] - \frac{1}{2}\int d^dx d^dy \varphi(x) R_k(x-y)\varphi(y)\right.\right. &(9.18)\\
&+ \left.\left.\frac{1}{2}\int d^dx d^dy M(x)R_k(x-y)M(y) + \int d^dx H(x)(\varphi(x) - M(x))\right)\right) &(9.19)
\end{aligned}$$

and we have:

$$\int d^dx H(x)(\varphi(x) - M(x)) = \int d^dx \left(\frac{\delta\Gamma_k[M]}{\delta M(x)} + \int d^dy R_k(x-y)M(y)\right)(\varphi(x) - M(x)) \tag{9.20}$$

now, let us combine all the terms containing R_k in the precedent expression of $\Gamma_k[M]$, equation (9.18):

$$-\frac{1}{2}\varphi R_k \varphi + \frac{1}{2} M R_k M + \varphi R_k M - M R_k M = -\frac{1}{2}(\varphi - M)R_k(\varphi - M) \tag{9.21}$$

so that:

$$\begin{aligned}
\Gamma_k[M] &= -\ln\left(\int \mathcal{D}[\varphi]\exp\left(-S[\varphi] + \int d^dx \frac{\delta\Gamma_K[M]}{\delta M(X)}(\varphi(x) - M(x))\right.\right. &(9.22)\\
&- \left.\left.\frac{1}{2}\int d^dx d^dy (\varphi(x) - M(x))R_k(x-y)(\varphi(y) - M(y))\right)\right) &(9.23)
\end{aligned}$$

Using FOURIER:

$$\int d^dx d^dy(\varphi(x)-M(x))R_k(x-y)(\varphi(y)-M(y)) = \int \frac{d^dq}{(2\pi)^d}|\hat{\varphi}(q)-\hat{M}(q)|^2\hat{R}_k(q) \tag{9.24}$$

Looking back at figure 32, we see that on the macroscopic length-scales $q \leq k$ we have a large $\hat{R}_k(q)$ which makes it costly for $\hat{\varphi}(q)$ to differ from $\hat{M}(q)$ in the precedent path integral; while on the microscopic side $q \geq k$ it is the inverse. When $k \to \infty$, we have $R_k(q) \sim k^2$ and the LEGENDRE transform $\Gamma_k[M]$ of equation (9.22) becomes:

$$\Gamma_K[M] = -\ln\left(\int \mathcal{D}[\varphi]\exp\left(-S[\varphi]+\int d^dx\frac{\delta\Gamma[M]}{\delta M(x)}(\varphi(x)-M(x))-k^2\int d^dq|\hat{\varphi}(q)-\hat{M}(q)|^2\right)\right) \tag{9.25}$$

there, it is clear that when $k \to \infty$, the term which is proportional to it dominates. So we can use a saddle-point argument to state that the functional integral is dominated by $\varphi_{SD}(x) = M(x)$; and as we announced earlier:

$$\boxed{\Gamma_{k\to\infty}[M] = S[M]} \tag{9.26}$$

Note that this result can be proven in a more rigorous way, mathematically speaking; if we apply the saddle-point corrections, they give constants which disappear with a renormalization of $\mathcal{D}[\varphi]$. Now that we know the limits of $\Gamma_{k\to\infty}[M]$ in both macroscopic and microscopic sides, we look for the flow equation which connects the two.

9.3 Flow equation

To obtain the flow equations, we start back from equation (9.18) and differentiate it with respects to k:

$$\begin{aligned}\partial_k\Gamma_k[M] &= \partial_k\left(-\ln\left(\int\mathcal{D}[\varphi]\exp\left(-S[\varphi]-\frac{1}{2}\int d^dxd^dy\varphi(x)R_k(x-y)\varphi(y)\right.\right.\right. &(9.27)\\ &+ \left.\left.\left.\frac{1}{2}\int d^dxd^dyM(x)R_k(x-y)M(y)+\int d^dxH(x)(\varphi(x)-M(x))\right)\right)\right) &(9.28)\end{aligned}$$

In the precedent expression, $S[\varphi]$ does not depend on k and is eliminated. But from the relation (9.17), which links $H(x)$ and $\Gamma_K^L[M]$ we observe that H depends implicitly on k and should not be excluded as well. Then:

$$\begin{aligned}\partial_k\Gamma_k[M] &= \frac{1}{2}\int d^dxd^dy\partial_kR_k(x-y)\left\langle\varphi(x)\varphi(y)\right\rangle-\frac{1}{2}\int d^dxd^dyM(x)\partial_kR_k(x-y)M(y) &(9.29)\\ &- \int d^dx\partial_kH(x)\left\langle\varphi(x)-M(x)\right\rangle &(9.30)\end{aligned}$$

where we can replace $M(x)$ by $\langle\varphi(x)\rangle$ using (9.6), so that the term proportional to $H(x)$ vanishes and we get:

$$\partial_k\Gamma_k[M] = \frac{1}{2}\int d^dxd^dy\partial_kR_k(x-y)(\langle\varphi(x)\varphi(y)\rangle-\langle\varphi(x)\rangle\langle\varphi(y)\rangle) \tag{9.31}$$

from (9.13) and (9.16):

$$\langle\varphi(x)\varphi(y)\rangle_c = \left(\frac{\delta^2\Gamma_k^L[M]}{\delta M(x)\delta M(y)}\right)^{-1} = \left(\frac{\delta^2\Gamma_K[M]}{\delta M(x)\delta M(y)}+R_k(x-y)\right)^{-1} \tag{9.32}$$

so finally we have:

$$\boxed{\partial_k\Gamma_k[M] = \frac{1}{2}\int d^dxd^dy\partial_kR_k(x-y)\left(\frac{\delta^2\Gamma_K[M]}{\delta M(x)\delta M(y)}+R_k(x-y)\right)^{-1}} \tag{9.33}$$

this is the exact renormalization group flow equation on $\Gamma_k[M]$, found by WETTERICH in 1993.

As we showed previously, the boundary conditions for this equation are given by $\Gamma_{k\to\infty}[M] = S[M]$ and the fact that $\Gamma_{k=0}[M]$ corresponds to the macroscopic theory. Historically, the first renormalization group flow equation to be found used the bare energy instead of Γ; but it was harder to do approximations with it so it was less useful. The WETTERICH equation cannot be solved exactly, approximations are still needed, but it gives a nice theoretical framework. In principle, when $T > T_c$ then $\Gamma_{k=0}[M]$ gives the High Temperature Fixed Point, and similarly when $T = T_c$ and $T < T_c$. A commonly used approximation is for instance the Local Potential Approximation, where we postulate that $\Gamma_k[M]$ takes the form:

$$\Gamma_k[M] = \int d^d x \frac{Z_k}{2} (\nabla M)^2 + V_k[M] \tag{9.34}$$

where, when $k \to \infty$:

$$\Gamma_k[M] \to S[M], \quad Z_k \to 1 \quad \text{and} \quad V_k[M] \to V[M] \tag{9.35}$$

then, we can translate the renormalization group flow equations on Z_k and V_k; this will be done in the problem class. As depicted in the following figure, these equations allow us to make V_k flow from microscopic $V_{k\to\infty}$ to macroscopic $V_{k=0}$, from which we can extract thermodynamical information. For instance the minimum of $V_{k=0}[M]$ gives an approximation of the free energy, but we can also obtain $M_\pm$ or the critical temperature T_c, etc.

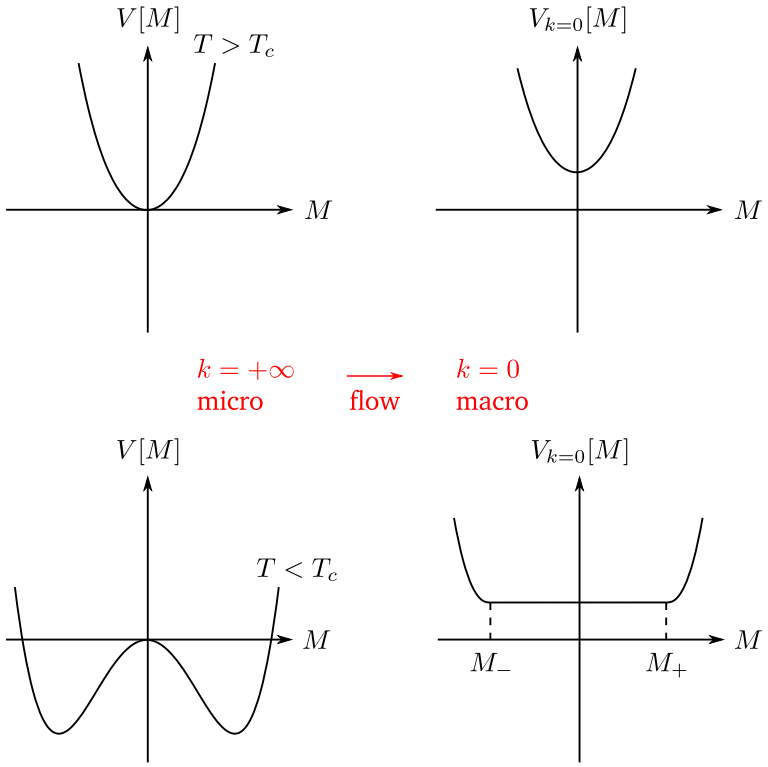

Figure 33: possible flow of V_k from $k = +\infty$ to $k = 0$ for $T > T_c$ and $T < T_c$

Previously, we saw that Mean Field Theory is incorrect between lower and upper critical dimensions d_l and d_u. The Local Potential Approximation for the functional renormalization group can allow to get around this problem.

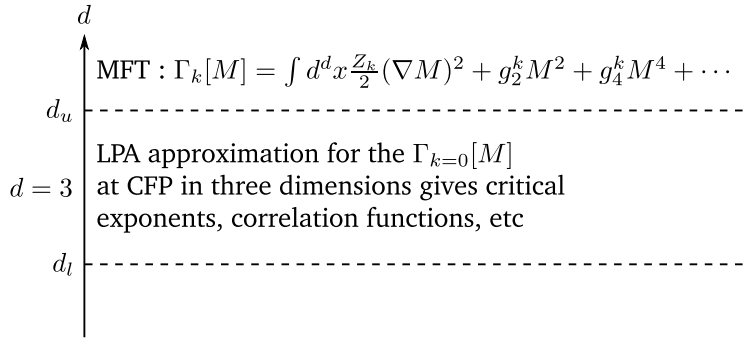

Figure 34: applicability of the Mean Field Theory and Local Potential Approximation

Above the upper critical dimension, and at the CFP, the MFT is a Gaussian theory. But it does not allow to easily get information in three dimensions. Historically, the first method to do this appeared with the perturbative renormalization group, in a paper from WILSON and FISHER in 1972. Since we know the CFP for $d > d_u$, the idea was to expand it perturbatively in $\varepsilon = d_u - d$, as a ε-expansion. It had the advantage to be a very deductive and controlled way of constructing the CFP, without thinking much about physics. But there are theories for which $d_u = 6$ and $\varepsilon = 3$ is not small anymore. In fact, for many interesting problems the critical point is non-perturbative.

The non-perturbative renormalization group appeared later, with WETTERICH in 1993. It is less controlled, since H depends on the assumptions we do on the class of $\Gamma_k[M]$, based on physical reasoning and guessing. But it allows to study theories in which the CFP is non-perturbative, and theories for which we need flows of functions.

Non-perturbative renormalization group has been used many times with success. For instance it allowed precise computation of the critical exponents of simple theories, like ϕ^4; the LPA reproduced the ε-expansion to the first order and gave quite good estimations of the critical exponents for $d = 3$. There are other examples from quantum chromodynamics, condensed matter and statistical physics. For instance it worked on the statistical physics problem of disordered systems, where couplings depend on space or are random, and which could not be solved with perturbative renormalization group. It also worked on problems related to out of equilibrium systems.

10 Phase transitions in low dimension

In this section, we consider the case where the dimension is too low for Mean Field Theory to apply.

10.1 Phase transition and broken symmetry

In the following we will consider short-range interactions; in fact most of what we will do fails at long-range. The simplest example is once again the ISING model on $\Lambda \subset \mathbb{Z}^d$, with periodic boundary conditions. For a given configuration $\mathcal{C}$, the energy is:

$$E(\mathcal{C}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \tag{10.1}$$

If $h = 0$, this model enjoys spin reversal symmetry; when going from the configuration $\mathcal{C}$ to the configuration $\bar{\mathcal{C}}$ by doing $\sigma_i \to -\sigma_i$ the energy does not change. Under this transformation, the magnetization is flipped:

$$\langle \sigma_1 \rangle = \frac{1}{Z} \sum_{\mathcal{C}} \sigma_1(\mathcal{C}) \exp(-\beta E(\mathcal{C})) \to \frac{1}{Z} \sum_{\bar{\mathcal{C}}} \sigma_1(\bar{\mathcal{C}}) \exp(-\beta E(\bar{\mathcal{C}})) = - \langle \sigma_1 \rangle \tag{10.2}$$

so $\langle \sigma_1 \rangle = 0$; this is valid even when $|\Lambda| \to \infty$, that is to say for a large system, in other words we have:

$$\boxed{\lim_{|\Lambda| \to \infty} \lim_{h \to 0} \langle \sigma_1 \rangle = 0} \tag{10.3}$$

This double limit is thus not a good way to detect a phase transition, since it always be null whatever the temperature. But these two limits do not commute, and we can put them in the other way:

$$\lim_{h \to 0^+} \lim_{|\Lambda| \to \infty} \langle \sigma_1 \rangle = 0 \quad \text{in the disordered phase} \tag{10.4}$$

$$= m_{\text{sp}} > 0 \quad \text{in the ordered one} \tag{10.5}$$

where m_{sp} is a spontaneous magnetization, associated with a spontaneous symmetry breaking. In this context, the magnetization is an order parameter. There are two typical scenarii depending on the order of the phase transition.

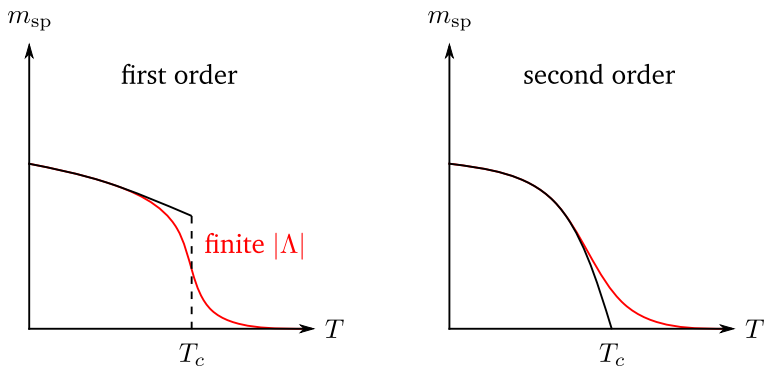

Figure 35: evolution of the spontaneous magnetization with the temperature

The jump in the case of a first order phase transition is associated to a latent heat. In the case of an infinite system, the free energy has a singular behavior even if the partition function is a sum of analytical terms, because there are infinitely many of such terms. In fact, in numerical simulations, we don't see such sharp variations; and as we see in the precedent figure it can be tough to distinguish between first and second order transitions. In the physical world, the symmetry is spontaneously broken thanks to uncontrolled fluctuations in the external magnetic field or at the boundary. Numerically, we can break the symmetry by adding a small h or asymmetric boundary conditions.

10.2 Competition between order and disorder

The terms order and disorder respectively refer to energy and entropy in this section. Let us start back from the energy of the ISING model (10.1) with $h = 0$:

$$E(\mathcal{C}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \tag{10.6}$$

for N spins. Canonical partition function is:

$$Z = \sum_{\mathcal{C}} \exp(-\beta E(\mathcal{C})) \tag{10.7}$$

many configurations share same energy:

$$Z = \sum_{E_\alpha} W(E_\alpha) \exp(-\beta E_\alpha) \tag{10.8}$$

where $W(E_\alpha)$ is the degeneracy of the states of energy E_α, the number of configurations with energy E_α. The degeneracy of the ground state E_0 is $W(E_0) = 2$ because of the spin flip symmetry. The degeneracy of the first excited state E_1 is $W(E_1) = 2N$. This combinatorial problem becomes more complicated at higher energy levels. This degeneracy defines the micro canonical entropy at fixed energy:

$$S_\alpha = k_B \ln W(E_\alpha) \tag{10.9}$$

so that the partition function rewrites:

$$Z = \sum_{E_\alpha} \exp(-\beta(E_\alpha - TS_\alpha)) \tag{10.10}$$

In the large N limit, we can use the fact that E and S are extensive, so they are proportional to the number N of spins: $E_\alpha \sim N e_\alpha$ and $S_\alpha \sim N s_\alpha$ with e_α and s_α independent of N. Then:

$$Z \underset{N\to\infty}{\sim} \sum_{E_\alpha} \exp(-\beta N(e_\alpha - T s_\alpha)) \tag{10.11}$$

and by a saddle point approximation:

$$Z \underset{N\to\infty}{\sim} \exp(-\beta N \min_\alpha(e_\alpha - T s_\alpha)) \tag{10.12}$$

only the leading term is kept in the sum. In this equation, we clearly see a competition between e_α and s_α. At low temperatures, $e_\alpha - T s_\alpha \approx e_\alpha$ and only the energy term plays a role, making for an ordered phase. At higher temperatures, the entropy term which is linked to the degeneracy can be higher, yielding a disordered phase. Note that this does not mean we can't have configurations with high energy and high entropy at the same time. For short, energy is associated with order while entropy is associated with disorder, and the two compete with each other.

Before $\exp(-\beta N \min_\alpha(e_\alpha - T s_\alpha))$ there is a prefactor in the precedent equation, but it depends on the microscopic details and it is often like N^μ. Since we ultimately look at the free energy $f_N = -1/(\beta N) \ln Z$, it does not play a role.

For the saddle point approximation to be valid, the number of configurations which dominate has to be small, in other words the energy distribution has to be sharp; and this is the fact for short-range interactions. There are examples of long-range interactions, for instance in $\sigma_i \sigma_j |i - j - 2|^{-\alpha}$, for which$E$ and S are not even extensive.

The existence of a phase transition depends on the dimension d of the model, and on its type of symmetry; whether it is a discrete symmetry as in the case of the ISING model, or a continuous one generated by a LIE group, as for the XY model, HEISENBERG model or $O(n)$ model. As we already mentioned, there exists in general a dimension d_l, called the lower critical dimension, under which there are no spontaneous symmetry breaking and phase transition.

10.3 Discrete symmetry groups

Let us start with the example of the ISING model, which enjoys $\mathbb{Z}_2$ as a symmetry group, with two simple elements wich are the identity $\pm \mapsto \pm$ and the spin flip $\pm \mapsto \mp$. The extension of this model to $\mathbb{Z}_p$ is called the POTTS model.

10.3.1 Scaling argument

First, we consider a one dimensional system with L sites and two spins up at the boundary.

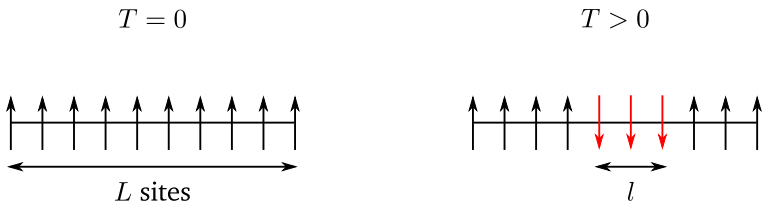

Figure 36: ISING model, with L sites and spins up at the boundary

At zero temperature, all the spins are aligned with the direction dictated by the boundaries. When the temperature goes finite, we expect the formation of clusters of spins up and clusters of spin down. Let us consider a minimal excitation as depicted in the precedent figure and compute the variation in energy and entropy associated with it. This excitation costs an energy $\Delta E = 2J$: only the borders of the domain cost energy. It is associated to a gain in entropy $\Delta S = k_B \ln(L - l)$: there are $L - l$ possibilities to fit the excitation of length l inside the chain of length L. If $L \gg l$ the free energy varies by:

$$\Delta F = \Delta E - T\Delta S = 2J - k_B T \ln L \tag{10.13}$$

so for large L and finite temperature, $\Delta F < 0$. In one dimension, entropy wins over energy and there is no symmetry broken phase when the temperature is non zero. Here this is due to the short range nature of the interaction. There can be phase transitions in one dimension with long range interactions, for instance MFT predicts a phase transition.

Let us now consider the two dimensional case, once again with fixed spins up at the boundary. At zero temperature, all the spins are aligned with the boundaries as previously. At finite temperature, we consider an excitation which is a cluster of spins pointing towards the opposite direction. The boundary of this cluster can be well approximated by a random walk. Let n be the length of this boundary, which is approximately the number of spins it contains.

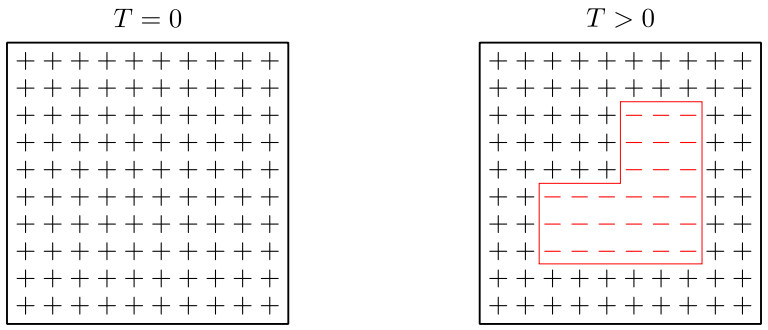

Figure 37: two dimensional ISING model, at zero and non zero temperature

This excitation costs an energy $\Delta E = 2nJ$ and it is associated to a gain in entropy which is hard to compute in a precise way, proportional to the logarithm of the number of random walks of n steps making a loop. Note that these loops are called bridges by probabilists. For an unconstrained random walk in two dimensions, there are four possibilities at each step, so with n steps it makes for 4^n walks. But in our case, the walker cannot turn back on a step and it has to make a loop, so the number of possible walks will scale like μ^n with some $\mu \approx 3$ which can be computed. It gives $\Delta S \approx n k_B \ln \mu$ and:

$$\Delta F = n(2J - k_B T \ln \mu) \tag{10.14}$$

So if T is small enough, $\Delta F > 0$ and the order can stay; but if T is larger, the entropy wins, the defects proliferate and the order is destroyed. Therefore, there must exists a critical temperature $T_c \propto J/k_B$ such that for $T < T_c$ energy wins and for $T > T_c$ entropy wins. This shows that the lower critical dimension of the ISING model is equal to one. In fact, this is not only true for the ISING model but for all models with short range interactions and discrete symmetry. These kind of arguments cannot tell us about the order of the transition; only whether it exists or not.

10.3.2 Exact solution in $d = 1$: transfer matrix approach

In the case of the one dimensional ISING model, the exact solution can easily be computed using what is called the transfer matrix formalism. Let us introducing it, starting from the Hamiltonian:

$$H = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - \frac{h}{2} \sum_{i=1}^{N} \sigma_i - \frac{h}{2} \sum_{i=1}^{N} \sigma_{i+1} \tag{10.15}$$

we decomposed it in a way that will be convenient later. We assume periodic boundary conditions $\sigma_{N+1} = \sigma_1$. The partition function can be written:

$$Z_N = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \langle \sigma_1 | \, T \, | \sigma_2 \rangle \cdots \langle \sigma_N | \, T \, | \sigma_1 \rangle \tag{10.16}$$

here, T is the following matrix:

$$T = \begin{pmatrix} \exp(\beta(J+h)) & \exp(-\beta J) \\ \exp(-\beta J) & \exp(\beta(J-h)) \end{pmatrix} \tag{10.17}$$

and is called the transfer matrix. The partition function can be reexpressed as:

$$Z_N = \mathrm{Tr}(T^N) = \lambda_+^N + \lambda_-^N \tag{10.18}$$

where λ_+ and λ_- are the eigenvalues of the transfer matrix, with $\lambda_+ < \lambda_-$ so that when $N \to \infty$ we have $Z_N \approx \lambda_+^N$ and the free energy becomes:

$$f_\infty = \lim_{N \to \infty} \left(-\frac{k_B T}{N} \ln Z_N \right) = -k_B T \ln \lambda_+ \tag{10.19}$$

The mean magnetization being:

$$\lim_{N \to \infty} \langle \sigma_1 \rangle = \frac{\partial f_\infty}{\partial h} = m_h \tag{10.20}$$

it can be computed explicitly, and:

$$\lim_{h \to 0} \langle \sigma_1 \rangle = 0 \qquad \text{if } T > 0 \tag{10.21}$$

$$= \mathrm{sign}(h) \qquad \text{if } T = 0 \tag{10.22}$$

so in a sense, there is a phase transition for the ISING model in one dimension, but it happens at zero temperature.

10.3.3 Exact solution in $d = 2$ (ONSAGER)

In the case of the two dimensional ISING model, the exact solution is somewhat harder to compute, but it has been found by ONSAGER that:

$$\lim_{h\to 0^+}\lim_{N\to\infty}\langle\sigma_1\rangle = 0 \qquad \text{if } T > T_c \tag{10.23}$$

$$= \left(1 - \left(\sinh\left(\frac{2J}{T}\right)\right)^{-4}\right)^{1/8} \qquad \text{if } T < T_c \tag{10.24}$$

where T_c is the solution of $\sinh(2J/T_c) = 1$. This proves that Mean Field Theory is wrong in low dimensions.

10.4 Continuous symmetry groups and MERMIN-WAGNER theorem

Here, we give an introduction; more technical details will be seen in the tutorial.

10.4.1 Heuristics

One of the most simple example of a continuous model is the $O(n)$ model, defined in any dimension d by this Hamiltonian, with the normalization $\boldsymbol{S}_i^2 = 1$:

$$H = -J\sum_{\langle i,j\rangle}\boldsymbol{S}_i\cdot\boldsymbol{S}_j \tag{10.25}$$

here, n is the number of components of each spin. As its name indicates, this model is invariant under continuous $O(n)$ transformations, which are rotations. Note that n can be different from d, for instance we consider spins with three components in a two dimensional lattice; in fact, it is sometimes useful to study the $n \to \infty$ case. When $d = 2$, the $O(n)$ model is called the XY model. We will consider the two dimensional case, with spins up at the boundary.

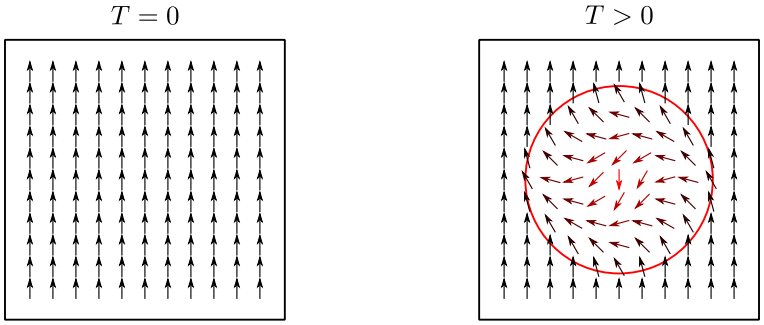

Figure 38: two dimensional $O(n)$ model, at zero and non zero temperature

At zero temperature, the spins are aligned with the boundary as previously. At finite temperature, we consider an excitation which is a spherical droplet and that we call Γ. The spins gradually change orientation from the boundary of Γ to reach anti alignment at its center, as depicted in the precedent figure. Let us call $\Delta\theta_{ij} = \theta_i - \theta_j$ the angle between two adjacent spins, a the lattice spacing and l the radius of Γ; if $l \gg a$, we have $\Delta\theta_{i,j} \approx a/l \ll 1$, and this excitation costs an energy:

$$\Delta E = -J\sum_{\langle i,j\rangle\in\Gamma}(\cos\Delta\theta_{ij} - 1) \approx J\sum_{\langle i,j\rangle\in\Gamma}\frac{1}{2}\left(\frac{a}{l}\right)^2 \sim \frac{Ja^2l^{d-2}}{2} \tag{10.26}$$

since the number of terms in the sum scales as l^d, the volume of the d-dimensional sphere. Note that in the two precedent cases with discrete symmetry, we had $\Delta E \propto l^{d-1}$, if we identify l with n in the two dimensional case. The gain in entropy associated with this excitation can also be estimated with a random walk, yielding $\Delta S \sim l \ln \nu$ with some ν. Finally we obtain:

$$\boxed{\Delta F = \frac{Ja^2 l^{d-2}}{2} - k_B T l \ln \nu} \tag{10.27}$$

We see on that precedent equation that the dimension $d = 2$ plays the same role as $d = 1$ in the discrete case. Indeed, The MERMIN-WAGNER theorem states that entropy always wins for $d \leq 2$ in models with continuous symmetry and short-range interactions; thus there is no long-range order and no phase transitions in this case. This theorem is also known as the COLEMAN theorem in the context of high energy physics. As we will see in the next section for case of the XY model, this does not exclude the existence of a quasi long-range order, as defined by KOSTERLITZ and THOULESS. For now, let us give some details about the approach of MERMIN and WAGNER, and the study of the stability of the system around an ordered phase, using small θ deviations and an effective model of perturbations.

10.4.2 MERMIN-WAGNER theorem

Let us go back to the Hamiltonian of the $O(n)$ model, in d dimensions:

$$H = -J \sum_{\langle i,j \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \tag{10.28}$$

we assume the existence of a phase with a spontaneous magnetization $m_{sp} = 1$, and we denote $S_{i,j}$ the components of $\boldsymbol{S}_i$, with the same normalization $\boldsymbol{S}_i^2 = 1$ as before, so that we can parameterize $\boldsymbol{S}_i$ using an $n-1$ components vector $\boldsymbol{\sigma}_i$, usually we take:

$$S_{i,1} = (1 - \boldsymbol{\sigma}_i^2)^{1/2}, \quad S_{i,2} = \sigma_{i,1}, \cdots, S_{i,n} = \sigma_{i,n-1} \tag{10.29}$$

and we assume that $|\sigma_{i,j}| \ll 1$ in the ordered phase. Note that there exists non-linear σ_i models, which don't make this assumption and are more general than what we will do. Here, we perform a linear expansion of $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$:

$$\boldsymbol{S}_i \cdot \boldsymbol{S}_j = \left(1 - \sum_\alpha \sigma_{i,\alpha}^2\right)^{1/2} \left(1 - \sum_\alpha \sigma_{j,\alpha}^2\right)^{1/2} + \sum_\alpha \sigma_{i,\alpha}\sigma_{j,\alpha} = 1 - \frac{1}{2}\sum_\alpha (\sigma_{i,\alpha} - \sigma_{j,\alpha})^2 + \mathcal{O}(\sigma^4) \tag{10.30}$$

then:

$$H = H_0 + \frac{J}{2} \sum_{\langle i,j \rangle} \sum_\alpha (\sigma_{i,\alpha} - \sigma_{j,\alpha})^2 + \mathcal{O}(\sigma^4) \tag{10.31}$$

which is an effective Hamiltonian for the $\boldsymbol{\sigma}_i$, which reveals the existence of GOLDSTONE modes. These are modes which preserve $(\sigma_{i,\alpha} - \sigma_{j,\alpha})^2$ and thus leave the energy unchanged. Since there are $n-1$ of such modes, there is a partial symmetry breaking, from $O(n)$ to $O(n-1)$. To better understand the GOLDSTONE modes, let us compare the situation to what we obtained previously with a discrete symmetry. When $T < T_c$, the effective potential was shaped like a double well, and any perturbation from the minimum costed energy. We can picture a GOLDSTONE mode for $n = 2$ using the so-called "mexican hat" potential, whose minimum is reached by a continuous family of points linked by a $O(n-1)$ symmetry. The GOLDSTONE modes are the generators of this family of points, thus one can move along them without leaving the minimum of the potential, so without spending energy. In the continuous limit, the terms in $\sum (\sigma_{i,\alpha} - \sigma_{j,\alpha})^2$ would yield $(\nabla\sigma)^2$, which is the Hamiltonian of a massless scalar particle.

For $r \gg 1$ it can be shown that:

$$\langle (\sigma_{i,\alpha} - \sigma_{i+r,\alpha})^2 \rangle \sim r \quad \text{if } d = 1 \tag{10.32}$$

$$\sim \ln r \quad \text{if } d = 2 \tag{10.33}$$

$$\sim \text{cst} \quad \text{if } d \geq 3 \tag{10.34}$$

so we cannot use the $\sigma_i \ll 1$ approximation for $d \leq 2$. The ordered state is not stable in this case; because if it were the $\sigma_i \ll 1$ would be valid. In the following, we will talk about the emergence of quasi long-range order in $d = 2$.

11 Topological defects

In this section, we give an introduction to the theory of topological defects, which earned KOSTERLITZ, THOULESS and HALDANE the NOBEL prize in 2016. Let us remind what we have seen so far. In the case of a discrete symmetry and short range interactions, for instance in the ISING model, there was no phase transition in one dimension because of the proliferation of defects above $T = 0$; there was a transition in two dimension though, since then this proliferation of defects only happened above a critical T_c. We also talked about the MERMIN-WAGNER theorem, stating that there is no long-range order in $d \leq 2$ in the case of a continuous symmetry and short-range interactions. Now, we introduce the notion of topological defects, which only arise in the case of continuous symmetries. As defined by James P. SETHNA, a topological defect is a tear that we cannot repair in the order parameter field.

Let us start from the XY model with $d = 2$:

$$H = -J \sum_{\langle i,j \rangle} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \tag{11.1}$$

where i, j label the sites of a square lattice and $\boldsymbol{S}_i$ is a unitary vector in two dimensions which can be parameterized as $\boldsymbol{S}_i = (\cos\theta_i, \sin\theta_i)$. Even though the MERMIN-WAGNER theorem tells us that there is no long-range order, there is still a phase transition as we will see now.

11.1 Existence of a pase transition

In order to prove the existence of a phase transition, we compute correlator $\langle \boldsymbol{S}_i \cdot \boldsymbol{S}_j \rangle$ for high and low temperatures.

11.1.1 High temperature phase

If $T \gg J$ then $\beta J \ll 1$. For $\beta J = 0$, we can show easily that $\langle \boldsymbol{S}_i \boldsymbol{S}_j \rangle = \delta_{i,j}$. From there, we can do a low-βJ expansion, and a tedious calculation would give that for $|x_i - x_j| \to \infty$:

$$\langle \boldsymbol{S}_i \cdot \boldsymbol{S}_j \rangle \sim \exp\left(-\frac{|x_i - x_j|}{\xi(\beta J)} \right) \tag{11.2}$$

the thing to notice here is that this correlation decreases exponentially.

11.1.2 Low-temperature phase

For $T = 0$ then $\boldsymbol{S}_i = \bar{\boldsymbol{S}}$ for all i. From there, if $T \ll 1$, we can safely assume that $|\theta_i - \theta_j| \ll 1$ so that the Hamiltonian can be expanded as:

$$H = E_0 + J \sum_{\langle i,j \rangle} \frac{1}{2} (\theta_i - \theta_j)^2 + \cdots \tag{11.3}$$

we set $i = 0$ and $j = x$ without loss of generality and we compute $\langle \boldsymbol{S}_0 \cdot \boldsymbol{S}_x \rangle = \langle \cos(\theta_0 - \theta_x) \rangle$:

$$\langle \boldsymbol{S}_0 \cdot \boldsymbol{S}_x \rangle = \int \prod_i d\theta_i \cos(\theta_0 - \theta_x) \exp(-H_1) \Big/ \int \prod_i d\theta_i \exp(-H_1) \tag{11.4}$$

here, H_1 is given by:

$$H_1 = \frac{J}{2} \sum_{\langle i,j \rangle} (\theta_i - \theta_j)^2 \tag{11.5}$$

Let us stress that this expansion holds for $\Delta\theta$ small, so for i, j neighbors, which is the case in the sums over $\langle i, j \rangle$. To calculate this, we use:

$$\cos(\theta_0 - \theta_x) = \frac{1}{2} (\exp(i(\theta_0 - \theta_x)) + \exp(-i(\theta_0 - \theta_x))) \tag{11.6}$$

and we introduce the variable $Y_{0x} = \theta_0 - \theta_x$. Since the θ_i are Gaussian variables, Y_{0x} is Gaussian too with $\langle Y_{0x}\rangle = 0$ and $\langle Y_{0x}^2\rangle = \sigma_{0x}^2$. We now need to compute σ_{0x}^2, because once it's done we obtain:

$$\begin{aligned}\langle \cos Y_{0x}\rangle &= \frac{1}{2}(\langle \exp(iY_{0x})\rangle + \langle \exp(-iY_{0x})\rangle) &(11.7)\\ &= \frac{1}{2}\left(\exp\left(-\frac{1}{2}\sigma_{0x}^2\right) + \exp\left(-\frac{1}{2}\sigma_{0x}^2\right)\right) &(11.8)\\ &= \exp\left(-\frac{1}{2}\sigma_{0x}^2\right) &(11.9)\end{aligned}$$

To compute σ_{0x}^2, we can draw an analogy between our current problem and the one of 2D crystals that we shall study in a problem class:

$$E_{\text{2D crystal}} = \frac{K}{2}\sum_{\langle i,j\rangle}((U_i^{(1)} - U_j^{(1)})^2 + (U_i^{(2)} - U_j^{(2)})^2) \tag{11.10}$$

where (1) and (2) are uncorrelated. The analogy between the two is obtained with:

$$U_i^{(1)} \longleftrightarrow \theta_i \quad \text{and} \quad K \longleftrightarrow J \tag{11.11}$$

using a result from the problem class, we can express σ_{0x}^2 as the following integral:

$$\sigma_{0x}^2 = \frac{2T}{J}\int_{-\pi}^{\pi}\frac{dq_x}{2\pi}\int_{-\pi}^{\pi}\frac{dq_y}{2\pi}\frac{1-\exp(-i\boldsymbol{q}\cdot\boldsymbol{x})}{4-2\cos q_x - 2\cos q_y} \tag{11.12}$$

for $x \to \infty$, we then get:

$$\sigma_{0x}^2 \sim \frac{T}{\pi J}\ln|x| \tag{11.13}$$

so that finally, we obtain:

$$\langle \boldsymbol{S}_0 \cdot \boldsymbol{S}_x\rangle = |x|^{-T/2\pi J} \tag{11.14}$$

now, the correlation decreases as a power-law, which is softer than exponentially.

11.1.3 Conclusions

In what precedes, we saw that for high temperatures the correlation function goes down exponentially with the distance between sites, while for low temperatures it goes down much softer, as a power law. There must be a critical temperature T_c which separates between the to regimes; in other words there must be a phase transition.

We are not saying that the low temperature phase enjoys long range order; this would violate the MERMIN-WAGNER theorem. Remind that we defined long range order by the condition:

$$\lim_{|x-y|\to\infty}\langle \boldsymbol{S}_0 \cdot \boldsymbol{S}_x\rangle \neq 0 \tag{11.15}$$

and this is clearly not the case here, the correlation may decrease slowly, it eventually goes to zero when $x \to \infty$. Instead, we say that the low temperature phase has Quasi Long-Range Order (QLRO). Since the exponent in the power law of the correlation is proportional to the temperature, the transition of the system is very soft. At $T = 0$, the spins are all aligned. When $0 < T < T_c$, fluctuations arise in the θ field, in the form of waves of high length. Thee spin waves are good indicators of the Quasi Long-Range Order.

11.2 Topological defects

Let us now introduce KOSTERLIZ and THOULESS theory of topological defects. For that, we start considering the continuum limit of (11.5), which is:

$$H_1 = \frac{J}{2}\int d^2\boldsymbol{r}(\boldsymbol{\nabla}\theta(\boldsymbol{r}))^2 \tag{11.16}$$

then, we define a defect as a non constant field θ which is stable, or at least meta stable, in other words it must satisfy $\delta H/\delta\theta = 0$. This imposes:

$$\boxed{\Delta\theta = 0} \tag{11.17}$$

with $\theta(\boldsymbol{r}) \neq \theta_0$, and this implies:

$$\boxed{\oint d\theta = 2\pi K} \tag{11.18}$$

for some $K \in \mathbb{Z}$. Note that θ is not physical, contrary to $\boldsymbol{S} = (\cos\theta, \sin\theta)$; there are multiple choices of $\theta(\boldsymbol{r})$ which can lead to the same physical situation. The precedent equation $\Delta\theta = 0$ is the subject of harmonic theory. If $\theta(\boldsymbol{r})$ has no singularity, LIOUVILLE theorem states that it is a constant; this is the case where $K = 0$. Therefore, non-trivial solutions must have singular points, or defects, which cannot be compensated with smooth functions. Then, STOKES theorem tells us that the integrals of $d\theta$ on different contours containing the same singularity are all equal to $2\pi K$. This makes K a characteristic of this defect, we often call it its charge. The general solution of the equation $\Delta\theta = 0$ can be written:

$$\theta(\boldsymbol{r}) = K\varphi + \theta_0 \tag{11.19}$$

here φ is the polar coordinate of the point $\boldsymbol{r}$, and θ is the angle of the spin located at $\boldsymbol{r}$. Let us show that this is a solution. In polar coordinates:

$$\boldsymbol{\nabla}\theta(\boldsymbol{r}) = \left(\frac{1}{r}\frac{\partial}{\partial\varphi}\boldsymbol{u}_\varphi + \frac{\partial}{\partial r}\boldsymbol{u}_r\right)\theta(\boldsymbol{r}) = \frac{K}{r}\boldsymbol{u}_\varphi \tag{11.20}$$

indeed, $\Delta\theta = \boldsymbol{\nabla}\boldsymbol{\nabla}\theta = 0$, and:

$$\oint d\theta = \oint K d\varphi = 2\pi K \tag{11.21}$$

for any contour around origin. Note that for $K \neq 0$ this solution has a singularity at 0 which goes like $1/r$.

Now, let us depict some solutions. First, for $K = 0$, we have $\theta(\boldsymbol{r}) = \theta_0$:

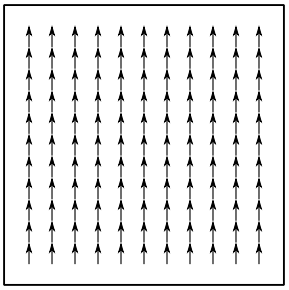

Figure 39: $K = 0$ and $\theta_0 = \pi/2$

this is the least interesting case.

It becomes interesting when $K = 1$, because of the apparition structures that we might call vortices of charge 1.

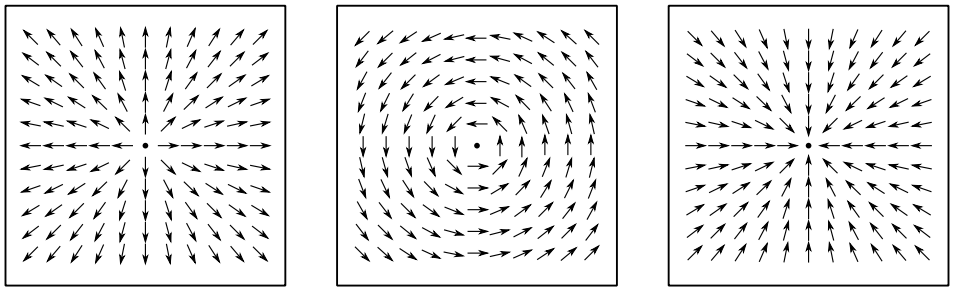

Figure 40: representation of the $\boldsymbol{S}$ field for $K = 1$ and θ_0=0 (left); $\theta_0 = \pi/2$ (center); $\theta_0 = \pi$ (right)

Singularities at 0 are clearly visible in the precedent figure. This defect cannot be killed by any transformation which is smooth at 0. The energy of a vortex of charge K is given by, using (11.16) and (11.20):

$$H = \frac{J}{2}\int d^2\boldsymbol{r}(\boldsymbol{\nabla}\theta(\boldsymbol{r}))^2 = \frac{J}{2}\int_a^L 2\pi r dr \left(\frac{K}{r}\right)^2 = J\pi K^2 \ln\left(\frac{L}{a}\right) \tag{11.22}$$

where a is the distance between neighboring points on the lattice and L the size of the system.

These vortices have some interesting properties. Only one of them is enough to destroy Long Range Order, since it obviously prevents nearby spins from pointing in the same direction. A finite density of vortices even kills Quasi Long Range Order. To understand this, one can observe that a field line would be sent in an other direction by each vortex it passes close to. Or one can try to imagine how badly a propagating spin wave would be affected by them.

As we already mentioned, vortices are topologically stable. It means that there is no continuous transformation of the field allowing to change their charge K. Moreover, vortices with the same charge are topologically equivalent. For instance, it is possible to deform the θ field from $\theta(\boldsymbol{r}) = \varphi$ to $\theta(\boldsymbol{r}) = \varphi + \pi/2$ in a smooth manner. This further means that there is an energy-free path connecting them. As we already said: K is the characteristic of the defects.

Finally, these topological defects are energetically stable. To understand this, let us look back at the precedent figure and determine how much energy it would cost to destroy one of the vortices, for instance the first one. Imagine we split the system in two parts around a vertical axis passing through the vortex. It would then be possible to continuously transform both parts so that all of the spins are pointing up, and this would cost no energy at all. But then, there would remain all the down pointing spins in the lower half of the frontier. Flipping all of them would cost an energy scaling like the size L of the system. According to ARRHENIUS law, this eventually happens with a characteristic $\tau \propto \exp(L/T)$, which can be extremely long.

11.3 KOSTERLITZ-THOULESS phase transition

We will use the same argument as previously in order to prove the existence of phase transitions and determine T_c. For reminders, we compared the energy cost of an excitation to the gain of entropy it represents. This, by the way, is called PEIERL argument. The energy of a vortex of charge K is $\Delta E = J\pi K^2 \ln(L/a)$. It is associated to a gain in entropy $\Delta S = k_B \ln((L/a)^2)$, using BOLTZMANN definition, since there are $(L/a)^2$ sites. So with the formation of a vortex the free energy varies by:

$$\Delta F = (J\pi K^2 - 2Tk_B)\ln\left(\frac{L}{a}\right) \tag{11.23}$$

From the precedent equation (11.23), we deduce the critical temperature $T_c = J\pi K^2/2k_B$. If $T > T_c$, entropy wins, defects proliferate to reach a finite density and QLRO is lost. On the contrary, if $T < T_c$ there are no free vortices. Note that we used the term "free vortices" here. In fact, below T_c, there are vortices forming bound pairs. To see that, one can compute the energy of a pair of vortices, which is not easy, to obtain:

$$H = A(K_1 + K_2)^2 \ln\left(\frac{L}{a}\right) + BK_1K_2 \ln\left(\frac{a}{r_{12}}\right) \tag{11.24}$$

with K_1, K_2 the charges of the two vortices r_{12} the distance between the two and A, B some constants. For two opposite charges K and $-K$, the energy reads:

$$E = -BK^2 \ln\left(\frac{a}{r_{12}}\right) \tag{11.25}$$

and we see that the interaction between the two vortices is attractive, since this energy decreases with the distance. From far away, or equivalently if they are close to each other, we can't distinguish a pair of opposite charge vortices.

We can describe the phase below T_c as a soup of bound pair of vortices. From $T = 0$, where they are not there, these pairs start to form when the temperature increase. Since their interaction is attractive, they may coalesce. The behavior of the system is essentially dynamical, with pairs forming and coalescing. The higher the temperature, the less bounded the pairs and the less frequently they coalesce, until we reach T_c above which they can be free.

Topological defects are of course not specific to the XY model. Their nature depends on symmetries of the model. As we did here with the charge K, it is in general possible to classify the defects of a model using homotopy theory. In the $O(n)$ model with d dimensions and $n > d$, there are no topological defects. We can see this for instance in figure 40, where if $n = 3$ there is a third direction towards which the spins can be smoothly rotated to kill the vortices. However, in the $O(3)$ HEISENBERG model in three dimensions there are defects which look like hedgehogs.

Let us conclude with some examples of topological defects:

- Superfluid helium is a phase living in three dimensions and associated with a $U(1)$ symmetry breaking. At $T = 0$ it does not move when we rotate its container; but at finite temperature, doing so creates vortices
- Liquid crystal is a two dimensional phase breaking a $SO(2)$ symmetry in which we observe topological defects.
- A fundamental property of crystalline solids is to dislocate under a shear stress; this creates a lot of defects. For instance, a copper rod can easily be bended and deformed by hand, but it is then extremely hard to make it go back to its initial state, unless we remove the defects by heating the rod in a process called annealing.

12 Dynamical emergence of long-range order

Let us consider a system possessing a second order phase transition at critical T_c that we couple to a thermostat. Then, we make the temperature of the thermostat vary from $T_i > T_c$ to T_f. We will consider two cases; in the first one $T_f = T_c + \varepsilon$ with ε small and we remain in equilibrium, in the second $T_f < T_c$ and the system exits equilibrium.

12.1 Critical equilibrium dynamics

First, we consider the case $T_f = T_c + \varepsilon$ and study the critical equilibrium dynamics. Close to the critical temperature, the dynamics of the system is extremely slow. If we pick a spin at random, it must flip after some time because the dynamics aboce T_c is ergodic, but this time can be long. This is linked to the fact that correlation lengths are large. We will find that the relaxation time is finite, so each intermediate state in the process of going from T_i to T_f can be considered at equilibrium. Therefore, the dynamics can be simulated using Monte Carlo methods. Also, as we get closer to T_c, we might start to observe the self-similarity which characterizes the critical phase, as discussed earlier.

We start from continuous action:

$$S[\varphi] = \int d^d\boldsymbol{x} \left(\frac{(\nabla\varphi(\boldsymbol{x}))^2}{2} + \frac{m\varphi^2(\boldsymbol{x})}{2} + \frac{g\varphi^4(\boldsymbol{x})}{4} \right) \tag{12.1}$$

and over-damped LANGEVIN equation:

$$\partial_t\varphi(\boldsymbol{x},t) = -\frac{\delta S}{\delta\varphi(\boldsymbol{x},t)} + \xi(\boldsymbol{x},t) \tag{12.2}$$

here, $\xi(\boldsymbol{x},t)$ is a thermal noise which satisfies $\langle\xi(\boldsymbol{x},t)\rangle = 0$ and $\langle\xi(\boldsymbol{x},t)\xi(\boldsymbol{x}',t')\rangle = 2T\delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')$ as always. Combining the two yields:

$$\partial_t\varphi(\boldsymbol{x},t) = \Delta\varphi(\boldsymbol{x},t) - m\varphi(\boldsymbol{x},t) - g\varphi^3(\boldsymbol{x},t) + \xi(\boldsymbol{x},t) \tag{12.3}$$

We want to study correlations:

$$C(\boldsymbol{x},t;\boldsymbol{x}',t') = \langle\varphi(\boldsymbol{x},t)\varphi(\boldsymbol{x}',t')\rangle \tag{12.4}$$

for that, we can use the MARTIN-SIGGIA-ROSE-DE DOMINICIS-JANSEN formalism, which yields good results. We can start at zeroth order in perturbation theory, that is to say with $g = 0$; and we can choose $m = a(T - T_c)$. In FOURIER space, the field $\varphi(\boldsymbol{x},t)$ becomes:

$$\hat{\varphi}(\boldsymbol{k},t) = \int d^d\boldsymbol{x} \exp(-i\boldsymbol{k}\cdot\boldsymbol{x})\varphi(x,t) \tag{12.5}$$

and the LANGEVIN equation:

$$\partial_t\hat{\varphi}(\boldsymbol{k},t) = -\boldsymbol{k}^2\hat{\varphi}(\boldsymbol{k},t) - m\hat{\varphi}(\boldsymbol{k},t) + \hat{\xi}(\boldsymbol{k},t) \tag{12.6}$$

Doing some computation that we don't show here, we would obtain the following:

$$\langle\hat{\xi}(\boldsymbol{k},t)\hat{\xi}(\boldsymbol{k}',t')\rangle = 2T(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}')\delta(t-t') \tag{12.7}$$

therefore the $\hat{\varphi}(\boldsymbol{k},t)$ are decoupled. This is an ORNSTEIN-UHLENBECK process, as we saw in the first problem class. Note that in real space, the $\varphi(\boldsymbol{x},t)$ are coupled at different $\boldsymbol{x}$ by the $\Delta\varphi(\boldsymbol{x},t)$. In the problem class, we studied $\dot{x}(t) = -\mu x(t) + \xi(t)$, with $\langle\xi(t)\xi(t')\rangle = 2T\delta(t-t')$, and we obtained $\langle x(t)x(t')\rangle = T/\mu\exp(-\mu|t-t'|)$. Let us use this result, with:

$$x(t) \leftrightarrow \hat{\varphi}(\boldsymbol{k},t), \quad \mu \leftrightarrow \boldsymbol{k}^2 + m \quad \text{and} \quad 2T \leftrightarrow 2T(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}') \tag{12.8}$$

it directly yields:

$$\langle\varphi(\boldsymbol{k},t)\varphi(\boldsymbol{k}',t')\rangle = \frac{T(2\pi)^d\delta^{(d)}(\boldsymbol{k}+\boldsymbol{k}')}{\boldsymbol{k}^2+m}\exp(-(\boldsymbol{k}^2+m)|t-t'|) \tag{12.9}$$

so the relaxation time is:

$$\tau_{\text{relax}} = \frac{1}{\boldsymbol{k}^2 + m} \tag{12.10}$$

this latter depends on k; the longer the wavelength the slower the relaxation, as for the waves made by throwing a rock into a lake. From equation (12.9), we go back to real space simply by reverse FOURIER transform:

$$\begin{aligned} C(\boldsymbol{x}, t; \boldsymbol{x}', t') &= \int \frac{d^d\boldsymbol{k}}{(2\pi)^d} \int \frac{d^d\boldsymbol{k}'}{(2\pi)^d} \exp(i\boldsymbol{k}\cdot\boldsymbol{x} + i\boldsymbol{k}'\cdot\boldsymbol{x}') \left\langle \varphi(\boldsymbol{k}, t)\varphi(\boldsymbol{k}', t')\right\rangle & (12.11)\\ &= T\int \frac{d^d\boldsymbol{k}}{(2\pi)^d} \exp(i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}'))\frac{\exp(-(\boldsymbol{k}^2+m)|t-t'|)}{\boldsymbol{k}^2+m} & (12.12)\\ &= Tm^{(d-2)/2}\int \frac{d^d\bar{\boldsymbol{k}}}{(2\pi)^d} \exp(im^{1/2}\bar{\boldsymbol{k}}\cdot(\boldsymbol{x}-\boldsymbol{x}'))\frac{\exp(-m(\bar{\boldsymbol{k}}^2+1)|t-t'|)}{\bar{\boldsymbol{k}}^2+1} & (12.13) \end{aligned}$$

using the natural change of variable $\boldsymbol{k} = m^{1/2}\bar{\boldsymbol{k}}$, where $m = a(T - T_c)$. We therefore obtain the correlation length and characteristic time:

$$\xi(T) = (a(T-T_c))^{-1/2} \quad \text{and} \quad \tau(T) = (a(T-T_c))^{-1} \tag{12.14}$$

so the C can be rewritten:

$$C(\boldsymbol{x}, t; \boldsymbol{x}', t') = a(T-T_c)^{(d-2)/2} f\left(\frac{\boldsymbol{x}-\boldsymbol{x}'}{\xi(T)}, \frac{|t-t'|}{\tau(T)}\right) \tag{12.15}$$

Note that the correlation length and relaxation time both diverge when the temperature approaches critical value. We also observe dynamical critical scaling. When the temperature varies, the scaling function f remains the same; only the prefactor and the characteristic scales ξ, τ are modified. The correlation length and characteristic time are linked by the time-length relation $\tau \sim \xi^z$ where $z = 2$ is called the dynamical exponent. We made our analysis with $g = 0$, but it remains valid for any g provided that $d > d_u$. If $d < d_u$, it might involve renormalization at some point.

The dynamical critical exponent z is new here. We can define α, another one, by $C(\boldsymbol{x}, t; \boldsymbol{x}', t') \sim |t-t'|^{-\alpha}$ at $T = T_c$. For a static exponent, such as the ν defined by $\xi(T) \sim (T-T_c)^{-\nu}$, the universality classes only depend on the symmetry broken by the ordered phase. In addition to that, the universality classes of a dynamic exponent such as z or α shall also depend on the conservation laws imposed during the dynamics. For instance, we can impose magnetization conservation by always flipping spins by pairs containing a spin up and a spin down, and vice versa. If we do so, one can show that the LANGEVIN equation (12.2) becomes the following:

$$\partial_t\varphi(\boldsymbol{x}, t) = \boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\left(\frac{\delta S}{\delta\varphi(\boldsymbol{x}, t)}\right) + \boldsymbol{\xi}(\boldsymbol{x}, t)\right) \tag{12.16}$$

with a thermal noise defined by $\langle\xi^\alpha(\boldsymbol{x}, t)\xi^\beta(\boldsymbol{x}', t')\rangle = 2T\delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')\delta_{\alpha\beta}$. This implies the conservation of the magnetization, which is defined by:

$$M = \frac{1}{V}\int d^d\boldsymbol{x}\varphi(\boldsymbol{x}, t) \tag{12.17}$$

Let us set $g = 0$ again and $T = T_c$, or equivalently $m = 0$. Previously we had:

$$\partial_t\hat{\varphi}(\boldsymbol{k}, t) = -\boldsymbol{k}^2\hat{\varphi}(\boldsymbol{k}, t) + \hat{\xi}(\boldsymbol{k}, t) \tag{12.18}$$

implying that $\tau_{\text{relax}} \sim 1/\boldsymbol{k}^2$ and $z = 2$. With the modified LANGEVIN equation:

$$\partial_t\hat{\varphi}(\boldsymbol{k}, t) = -\boldsymbol{k}^4\hat{\varphi}(\boldsymbol{k}, t) - i\boldsymbol{k}\cdot\hat{\boldsymbol{\xi}}(\boldsymbol{k}, t) \tag{12.19}$$

so we obtain $\tau_{\text{relax}} \sim 1/\boldsymbol{k}^4$. The correlation length is not affected by the choice of conservation law, as a static quantity. Therefore, imposing magnetization conservation we obtain $z = 4$. The system takes a longer time to relax. It is intuitive that with less constraints, the system can explore configurations and it breaks long-range order easier.

As a conclusion, the universality classes for static critical exponents, which only depend on the kind of symmetry which is broken, are divided into subclasses for dynamic critical exponents, depending on the conservation laws. This subclassification has been achieved by HALPERIN and HOHENBERG in 1977. They identified six subclasses, that they labeled with the letters from A to G. The subclass A corresponds to dynamics without conservation law, B to dynamics with conserved order parameter, C to dynamics with conserved energy, etc. Note that everything we have seen so far, from Mean Field Theory to the theory of topological defects, can be generalized to the dynamical case.

12.2 Quench across the phase transition

Let us now tackle the case $T_f < T_c$ when we vary the temperature of the thermostat from $T_i > T_c$ to T_f. The first thing to observe in that case is that the dynamics is necessarily out of equilibrium, no matter how slowly we adjust the temperature, because in general the relaxation time diverges when T approaches T_c, as in equation (12.14). In fact, after crossing T_c, the system organizes itself in domains of characteristic size $\xi(t)$ inside which spins are aligned, as depicted:

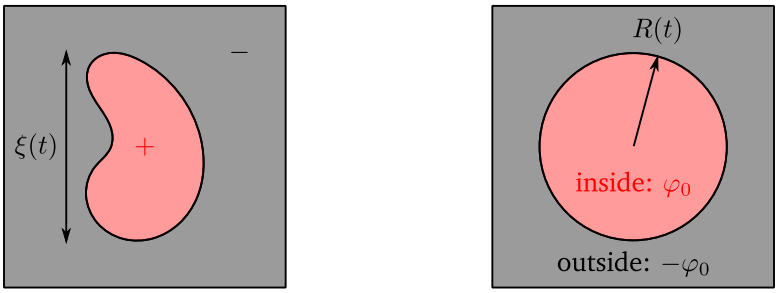

Figure 41: domains inside which the spins are aligned (left); we approximate these domains as spherical (right)

Inside the domains, the spins are at equilibrium; we shall study this equilibrium in a problem class. Still, the system as a whole is out of equilibrium since the domains move. At that point, there is no symmetry breaking. If we pick any spin, it must flip after some time; for instance if the domain it used to be in moves and is replaced by another.

12.2.1 Dynamics of domains

To simplify, we consider a spherical domain of radius $R(t)$, as shown in the right hand side of precedent figure. We also consider $R \gg 1$ and $t \gg 1$. Then, we start from an action similar to (12.1):

$$S[\varphi] = \int d^d\boldsymbol{x} \left(\frac{(\nabla\varphi(\boldsymbol{x}))^2}{2} + V[\varphi(\boldsymbol{x})] \right) \tag{12.20}$$

where the potential V is shaped like a symmetric double well, with two minima located at $\pm\varphi_0$. According to the LANGEVIN equation (12.2) without thermal noise, the equation of motion reads:

$$\partial_t\varphi(\boldsymbol{x},t) = \Delta\varphi(\boldsymbol{x},t) - V'[\varphi(\boldsymbol{x},t)] \tag{12.21}$$

since the problem is spherically symmetrical, we introduce spherical coordinates:

$$\partial_t\varphi(r,t) = \frac{d^2\varphi(r,t)}{dr^2} + \frac{2}{r}\frac{d\varphi(r,t)}{dr} - V'[\varphi(r,t)] \tag{12.22}$$

Now, we will try to find a solution in the form of a traveling wave, that is to say: $\varphi(r,t) = f(r - R(t))$, with a profile as the one shown in the following figure:

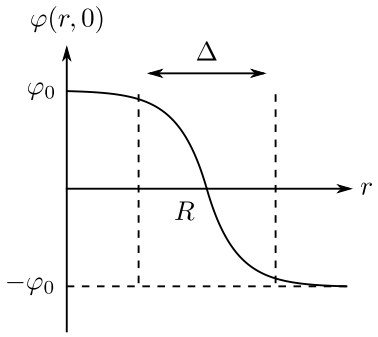

Figure 42: profile of φ inside the spherical domain

as depicted in figure 41, the field φ is equal to φ_0 inside the domain of radius R and to $-\varphi_0$ outside; a continuous connection between the two is made on a lengthscale Δ which can be small with respects to R. The first step to solve (12.22) with the traveling wave ansatz $\varphi(r,t) = f(r - R(t))$ is to find $R(t)$; and after that we will look for $f(r)$.

Since there is no thermal noise in (12.21), there is no second term in the Itō chain rule and we can insert the ansatz in equation (12.22) without worrying much:

$$-\dot{R}f' = f'' + \frac{2f'}{r} - V'[f] \tag{12.23}$$

multiplying both sides of this equation by f' and integrating over some length around $R(t)$:

$$-\int_{R(t)-l}^{R(t)+l} dr \left(\left(\dot{R} + \frac{2}{r} \right) (f')^2 \right) = \int_{R(t)-l}^{R(t)+l} dr (f''f' - V'[f]f') \tag{12.24}$$

we can choose l such that $R \gg l \gg 1$. The right hand side of this equation is straightforward:

$$\int_{R-l}^{R+l} dr \frac{d}{dr} \left(\frac{(f')^2}{2} - V[f] \right) = \left[\frac{(f')^2}{2} - V[f] \right]_{R-l}^{R+l} \tag{12.25}$$

The first term is zero, because the profile f is assumed to become quickly flat around R as we see on the precedent figure, so $f'(R \pm l) = 0$. The second term is also zero, because f is assumed to be antisymmetric while V is symmetric, so $V[f(R+l)] = V[f(R-l)]$. All this means that the left hand side also has to be zero; and since $R \gg l$ we can replace the $2/r$ by $2/R$:

$$-\int_{R(t)-l}^{R(t)+l} dr \left(\left(\dot{R} + \frac{2}{R} \right) (f')^2 \right) = 0 \tag{12.26}$$

finally, $\dot{R} + 2/R$ come out of the integral, and since $f' \neq 0$ close to $R(t)$ we get:

$$\boxed{\dot{R}(t) = -\frac{2}{R(t)}} \tag{12.27}$$

and in the end the solution is $R^2(t) = R^2(0) - 4t$. The important thing to notice here is that under our assumptions on the profile f, that are summarized by the precedent figure, the radius of a spherical domain is always decreasing.

But let us explain why these assumptions are coherent with what we found. Using the precedent equation (12.27), two terms of equation (12.23) vanish and the remaining ones read $f'' - V'[f] = 0$. This is exactly the same as the equation of motion of some massive mobile in the potential $-V$, which would be $\ddot{x} = -(-V'(x))$. We identify $f(r)$ with $x(t)$. The potential $-V$ is like a symmetric double hill. If we put the mobile close to the top of one hill, but slightly towards the other one, it will remain there for quite a long time, but at some point it will run away from equilibrium and quickly reach the other hill, on which it will climb for another quite long time. If we plot the position $x(t)$ of the mobile along the axis as a function of time, we obtain something very similar to figure 42, as expected.

We can now draw some conclusions. First, as we mentioned, the radius of a domain is always decreasing, and a spherical domain of initial size R_0 disappears after a time $t \sim R_0^2/4$. To put it differently, at time t all the domains of size $\xi(t) \sim 2\sqrt{t}$ have disappeared and the remaining ones are of size $l \geq \xi(t)$. In a more realistic context, the domains are not spherical; their frontiers evolve according to curvature driven dynamics, where their local speed is proportional to their curvature. These dynamics tend to make domains more spherical before making them shrink. This is all due to the fact that the interface costs energy. For the same perimeter, spheres enclose the greater area.

12.2.2 Second order phase transition

Finally, let us illustrate the emergence of long-range-order when quenching across the second order phase transition. Right after the temperature crosses its critical value, the out of equilibrium dynamics that we just described, with the shrinking domains, starts to take place. The smaller domains are annihilated first. After a time t, there remains only the ones with a characteristic size $\xi(t) \sim t^{1/z_0}$, where z_0 is the dynamical critical exponent out of equilibrium. When the characteristic size of the domains reaches the size of the system, one of them can win over the others and propagate over the whole system, establishing long range order. But the system is not frozen in this equilibrium. At any time, thermal fluctuations trigger the formation of small domains of opposite spin. Every once in a while, these domains can coalesce, grow bigger, and take over the whole system to establish an other equilibrium with opposite magnetization. This happens on timescales $\tau \sim \exp(JL^{d-1}/T)$, where L is the size of the system. We can summarize all of this in the following figure, which depicts the evolution of magnetization in time.

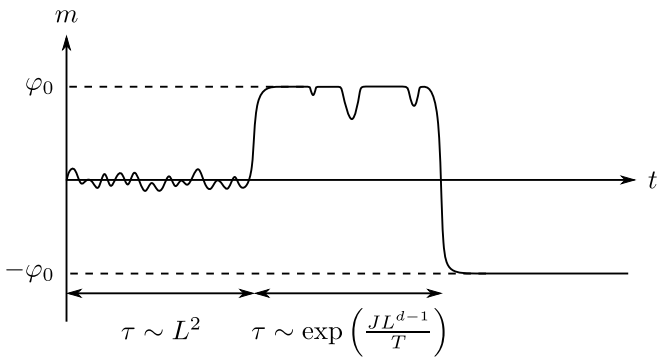

Figure 43: schematic evolution of the magnetization in time

These kind of dynamics have been studied in the context of some cosmology models, which considered that galaxies were the product of slowly growing topological defects induced by the Big Bang; but these models are not used anymore. Quenches across second order phase transition are also studied experimentally in cold atoms experiments.

13 Entropy in information theory

Information theory is a field of science which was pioneered by Claude SHANNON in his article published in 1948, *Mathematical theory of communication*. As its name indicates, it investigates communication. In some sense, we can define communication as an exchange of signals. In general, these signals have to go through several steps between emission from some information source and reception, as illustrated by the following figure:

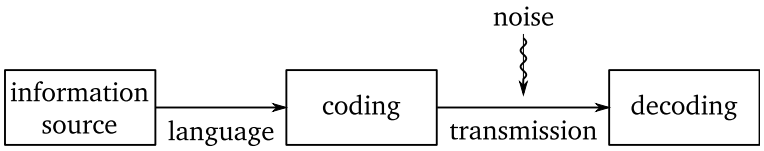

Figure 44: a signal's journey between information source and receptor

From the information source, the signal has to be coded in order to be transmitted via a given media, for instance via electromagnetic waves, where it can pick up some noise. After reception, it has to be denoised and decoded.

13.1 Choice, uncertainty, entropy

13.1.1 SHANNON's entropy

The first fundamental question asked by information theory is how to quantify the amount of uncertainty contained in a decoded message ? This uncertainty should be related to the gain of information we obtain from knowing the message directly from the information source. We would like to have a quantity measuring either one of those; and this is precisely what is done by SHANNON's entropy H. Let us formalize the intuition we have on this quantity by considering a random experiment with N possible outcomes of respective probabilities p_i, such that $p_1 + \cdots + p_N = 1$. The first condition we impose on the entropy $H(p_1, \ldots, p_N)$ it is that it should be a continuous function of the p_i. Then, if all the outcomes of the experiment are equiprobable, H must be a monotonic increasing function of N. Finally, if the events can be broken into successive events, then H should be breakable accordingly, as we will better illustrate with this example:

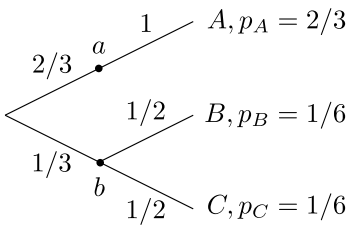

Figure 45: events broken into successive events

On the precedent figure, the events A, B and C are broken into successive events. The entropy H can then be decomposed in this way:

$$H(2/3, 1/6, 1/6) = H(2/3, 1/3) + \frac{2}{3}H(1) + \frac{1}{3}H(1/2, 1/2) \tag{13.1}$$

In his article, SHANNON proved that the only entropy H satisfying the three requirements is the following one; as we will demonstrate in problem class:

$$H(p_1, \ldots, p_N) = -K \sum_i p_i \ln p_i \tag{13.2}$$

with $K = 1/\ln 2$, which amounts to replacing $\ln$ by $\log_2$. Let us give some examples of application of this formula. First, if one of the outcomes is almost certain, the uncertainty is obviously zero:

$$H(1, 0, \ldots, 0) = -1 \log_2 1 - 0 \log_2 0 = 0 \tag{13.3}$$

Then, we consider two events with probability p and $1 - p$, we can directly write:

$$H(p, 1-p) = -p \log_2 p - (1-p) \log_2 (1-p) \tag{13.4}$$

and it is easy to show that this uncertainty is maximal when $p = 1/2$ as we would expect. Moreover, for that particular value, $H(1/2, 1/2) = \log_2 2 = 1$; thus, we say that knowing the outcome gives us one bit of information.

Finally, let us take two independent processes A and B, with events of probabilities p_i and q_j respectively. Since A and B are independent, the joint probability can be factorized, $p_{ij} = p_i q_j$, and the entropies can be summed:

$$H_{A,B} = -\sum_{i,j} p_i q_j (\log_2 p_i + \log_2 q_j) = -\sum_i p_i \sum_j q_j \log_2 p_i - \sum_i p_i \sum_j q_j \log_2 q_j = H_A + H_B \tag{13.5}$$

13.1.2 Generalized entropies

We can define other notions of entropy which can be useful in the case of two processes which are not independent. Let us take again our two processes A and B with events of probabilities p_i and q_j respectively. But now the joint probability cannot be factorized. The conditional entropy is then defined by:

$$H_{B|A} = \sum_i p_i \left(-\sum_j p_{j|i} \log_2 p_{j|i} \right) \tag{13.6}$$

here, we use the standard notation $p_{j|i} = p_{i,j}/p_i$. The conditional entropy satisfies $H_{A,B} = H_A + H_{B|A}$; as we check simply by writing:

$$H_A + H_{B|A} = -\sum_i p_i \log_2 p_i - \sum_{i,j} p_{ij} (\log_2 p_{ij} - \log_2 p_i) = H_{A,B} \tag{13.7}$$

The common information in the processes A and B, or mutual information, is defined by:

$$I_{A,B} = \sum_{i,j} p_{ij} \log_2 \left(\frac{p_{ij}}{p_i q_j} \right) \tag{13.8}$$

it satisfies $I_{A,B} = H_A + H_B - H_{A,B}$ or, using the previous identity satisfied by conditional entropy, $I_{A,B} = H_B - H_{B|A}$. This last identity is intuitive, it says that the information shared by the two processes is the total information of B to which we subtract the part which is not contained in A. Once again these identities can be checked by writing, for instance:

$$I_{A,B} = \sum_{i,j} p_{ij} \log_2 p_{ij} - \sum_{i,j} p_{ij} \log_2 p_i - \sum_{i,j} p_{ij} \log_2 q_j = -H_{A,B} + H_A + H_B \tag{13.9}$$

Let us give examples. If A and B are such that A implies B, in the sense that $p_{j|i} = 1$ for some $j = j^*(i)$ and $p_{j|i} = 0$ otherwise, then:

$$I_{A,B} = H_B + \sum_i p_i \left(p_{j|i} \log_2 p_{j|i} \right) = H_B \tag{13.10}$$

Similarly, if A and B are uncorrelated, we can check that $I_{A,B} = 0$ as it should.

13.2 Compression

Information theory also gives tools to compare the different languages in which the information can be expressed, as depicted in figure 44. For instance, imagine that the signal emerging from the information source is composed of characters α, β, γ and δ with probabilities $p_\alpha = 1/2, p_\beta = 1/4$ and $p_\gamma = p_\delta = 1/8$. A word in this language could be for instance $\alpha\alpha\gamma\beta\delta\alpha\beta$. It is convenient to code this using 1 and 0 in order to transmit it through some media. This code should be injective. For example if we used $\alpha \to 0$, $\beta \to 1$ and $\gamma \to 01$, we would not distinguish $\alpha\beta$ from γ. It is also convenient if the code is efficient, in the sense that the most used characters should be coded with less bits than the least used ones. In our case, such an efficient code is given by $\alpha \to 0, \beta \to 10, \gamma \to 110$ and $\delta \to 111$.

The entropy of the source is:

$$-\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{2}{8}\log_2\left(\frac{1}{8}\right) = \frac{7}{4} \tag{13.11}$$

and the average length of a coded character, which measure of the compression, is $1/2 + 2/4 + 2 \times 3/8 = 7/4$. We notice that the two are equal; and this is not by chance. With is optimal compression theorem, SHANNON proved that considering a string of M characters with $M \to \infty$, the shortest average length divided by M achievable by any code is equal to the entropy of the source. With that in mind, we see that the code for α, β, γ and δ is in fact optimal. This also means that if we know the entropy of a source, we can try to optimize the code for that source.

Here, entropy acquired a new meaning. We already interpreted it as a measure of the uncertainty of an event that is going to happen, or equivalently as a measure of the information we obtain from knowing which event happened. Now, we interpret it as a measure of the minimum amount of space needed to code a representation of this event.

The optimal compression theorem allows to characterize the most efficient languages. But it is not always the ones we are looking for; indeed, if the signal is corrupted by noise during transmission, having redundancies in the language makes it easier to be denoised. This is for example the case in natural languages, which have a lot of extra entropy to prevent mishearing. In fact, depending on the nature of the noise, some languages cannot even be used.

13.3 Back to physics

In order to see how this entropy of information theory resembles the one we know in physics, let us consider for the last time a system of N interacting spins $S_1, \dots S_N$ arranged in a square lattice. When the temperature goes to infinity, all these spins are independent and uncorrelated, so we can isolate them one by one:

$$H_{\{S_1,\dots S_N\}} = H_{\{S_1\}} + H_{\{S_2,\dots S_N\}|\{S_1\}} = 1 + H_{\{S_2,\dots S_N\}} = \dots = N \tag{13.12}$$

where $H_{\{S_1\}} = 1$ because this individual spin can only be up or down with probability $1/2$. This case, where all the spins are independent and uncorrelated, is the case of maximal entropy. To describe the system, we would need N bits of information, one for the state of every spin. In this case, the entropy is extensive. Even at finite, high temperature, this argument holds, provided we regroup the spins in clusters the size of the correlation length. This works because these clusters would be uncorrelated with each other. The entropy is also extensive in this case.

At $T = 0$, all the spins are aligned:

$$H_{\{S_1,\dots S_N\}} = H_{\{S_1\}} + H_{\{S_2,\dots S_N\}|\{S_1\}} = 1 \tag{13.13}$$

dividing by N, we would recover $S = 0$ in the large N limit. To describe the system, we would only need one bit of information, since the state of any spin tells the state of the others. So as we see, entropy in physics can be interpreted as the amount of memory needed to encode all the information contained in the equilibrium configuration. This is an interpretation which extends for out of equilibrium systems and for systems which are not exactly solvable.

This interpretation of entropy is extremely rich. For instance, it can allow us to understand compression algorithms using physical intuition; or it can highlights how biological systems process information in a very optimized way.