Stochastic Dynamics and Statistical Fields

Pietro Faccioli

April 18, 2021

Contents

Chapter 1

Prologue

1.1 Introduction

These notes provide an introduction to modern theoretical physics methods to describe the dynamics complex systems. In the following, we shall conventionally refer to as *complex* a physical system which displays two main features: (i) it is composed by a large number of degrees of freedom and (ii) its dynamics is shaped by the presence of strong correlations and by thermal or quantum fluctuations. According such a working definition, complex systems cannot be described by regarding their elementary constituents as nearly independent (or weakly interacting) particles (or fields). Indeed, the interplay between correlations and fluctuations gives rise to critical phenomena, collective excitations, cooperativity etc.. For this reason, the theories developed to describe their dynamics make use of concepts and results of equilibrium and non-equilibrium statistical physics and are often formulated in terms of the effective (i.e non-fundamental) degrees of freedom, in both particle or field representation.

While the theories and methods we shall discuss are applicable to a generic complex system, these notes are conceived with in mind prototypical applications to soft condensed matter systems, and in particular to biopolymers in solution, such as proteins and nucleid acids. A brief summary of the basic structural and chemical properties of these macromolecules is provided in the appendix to these notes.

The present treatment of the dynamics of complex systems is based on four main concepts and tools: (i) the notion of systematic Effective Theory (ET), (ii) the path integral formalism, (iii) the concept of stochastic process and (iv) the notion quantum and statistical fields.

Generally speaking, effective theories (ETs)–for an excellent pedagogical introduction see [16]– are systematic low-resolution approximations of arbitrary more microscopic theories. Unlike other coarse-graining approaches inspired by heuristic or phenomenological arguments, ETs can be rigorously derived starting from the underlying microscopic theory within the framework of Renormalization Group (RG) theory. Under certain well-defined conditions to be discussed in the next section, ETs are guaranteed to approximate to any given finite degree of accuracy the long-distance (or long-time) dynamics.

Functional Calculus has become a standard mathematical language in which modern quantum and statistical field theories are formulated, even in non-relativistic regimes [2, 3, 4, 5]. This is because the path integral formalism offers several advantages with respect to the traditional operator-based one. For example, it allows to express the quantum dynamics in terms of purely classical concepts, providing a clean and physically sound framework to explore the transition between quantum and classical regime of physical systems. It also elucidates the role of fluctuations in disrupting the deterministic character of classical mechanics, and displays the analogies and the differences between the effect of thermal and quantum fluctuations. Furthermore, many powerful computational methods routinely used to investigate quantum or classical many body systems are based on exploiting the connection between the classical canonical partition function and the imaginary time quantum mechanical path integral. Finally, path integrals provide the mathematical framework to develop a number of useful approximations.

Stochastic processes are key to describe phenomena in very different fields of science, economy, and even sociology. The main reason is that a stochastic behavior emerges any time our description of the dynamics is restricted to a sub-set of degrees of freedom, which are nonetheless coupled to many others, which we do not explicitly specify. This situation is often encountered in statistical physics, when coupling the system of interest to a reservoir.

Finally, the field representation (as opposed to the particle representation) enables one attain a coarser grained representation of the state of a system (based on global descriptors such as the local density), rather than having to specify the position of each particle. Clearly, if the density field varies over length scales which are sufficiently large, then a field representation can be extremely more "economical" than the particle. It may take fewer variables to describe the field configuration than the $6N$ phase-space coordinates.

In this Introduction, we first review the standard Feynman path integral representation of Quantum Mechanics and the connection between classical statistical mechanics and imaginary time quantum mechanics. Then, we provide an illustrative example of a systematic ET, which illustrates the basic ideas behind the renormalization group based coarse graining.

1.2 Path integral Representation of Quantum Mechanics

The equation of motion of quantum states associated to non-relativistic point-particles is the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle. \tag{1.1}$$

Note that we have denoted explicitly the time dependence of the quantum state $|\Psi\rangle$. This differential equation is linear in time, hence can be formally solved by specifying the initial state, by introducing the time evolution operator:

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t}\,|\Psi(0)\rangle. \tag{1.2}$$

In particular, if the state is represented by projecting onto the complete basis set formed by the position eigenstates $|x\rangle$, we find:

$$\psi(x,t) = \langle x|\Psi(t)\rangle = \int dy\langle x|e^{-\frac{i}{\hbar}\hat{H}t}\,|y\rangle\langle y|\Psi(0)\rangle = \int dy\; K(x,t;y,0)\;\psi(y,0), \tag{1.3}$$

where we have used the resolution of the identity

$$1 = \int dy |y\rangle\langle y|, \tag{1.4}$$

and we have introduced Feynman propagator:

$$K(x,t;y) \equiv \langle x|e^{-\frac{i}{\hbar}\hat{H}t}\ |y\rangle. \tag{1.5}$$

Hence, solving a quantum system amounts to specifying the initial state and evaluating the propagator.

For sake of simplicity and without loss of generality, we present the derivation for a quantum particle in one dimension. The generalization to multi-dimensional systems with local scalar potential is straightforward

Let us now split the time interval t into N equally sized incremental time steps:

$$\Delta t = t/N \tag{1.6}$$

Next, we introduce the so-called Trotter decomposition of the of the evolution operator:

$$e^{-\frac{i}{\hbar}\hat{H}t} = \left(e^{-\frac{i}{\hbar}\hat{H}\Delta t}\right)^N = e^{-\frac{i}{\hbar}\hat{H}\Delta t}\dots e^{-\frac{i}{\hbar}\hat{H}\Delta t} \qquad \text{(N times)} \tag{1.7}$$

We plug in this expression into the definition of the propagator and insert a complete set of position eigenstates between each of the term of this product:

$$\begin{aligned} K(x_f,t;x_i,0) \quad &= \quad \left(\prod_{l=1}^{N-1}\int dy_l\right)\ \langle x_f|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|y_{N-1}\rangle\langle y_{N-1}|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|y_{N-1}\rangle\dots \\ &\quad \dots\langle y_1|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|x_i\rangle \end{aligned} \tag{1.8}$$

In the large N limit (i.e. for small Δt the matrix elements of incremental evolution propagators can be analytically evaluated. Let us restrict our discussion here to Hamiltonians which are defined as operational functions of the momentum $\hat{p}$ and position $\hat{x}$ operators in the form[1].

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(\hat{x}) \tag{1.9}$$

For sufficiently small Δt the exponents in the matrix elements

$$\langle y_{k+1}|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|y_k\rangle \tag{1.10}$$

can be expanded to leading order in Δt. To enables us to split the exponent as follows

$$e^{-\frac{i}{\hbar}\hat{H}\Delta t} = e^{-\frac{i}{\hbar}(\hat{T}+\hat{U})\Delta t} = e^{-\frac{i}{\hbar}\hat{T}\Delta t}\ e^{-\frac{i}{\hbar}\hat{U}\Delta t} + \mathcal{O}(\Delta t^2) \tag{1.11}$$

---

[1] Some extra care is needed in the case of particles derivatively-coupled to vector potentials. The interested reader can find a detailed discussion of this problem e.g. in [5]

We find:

$$\begin{aligned}
\langle y_{k+1}|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|y_k\rangle &\simeq \langle y_{k+1}|e^{-\frac{i}{\hbar}\hat{T}\Delta t}\ e^{-\frac{i}{\hbar}\hat{U}\Delta t}|y_k\rangle \\
&= e^{-\frac{i}{\hbar}U(y_k)\Delta t}\langle y_{k+1}|e^{-\frac{i}{\hbar}\hat{T}\Delta t}|y_k\rangle \\
&= e^{-\frac{i}{\hbar}U(y_k)\Delta t}\int \frac{dp}{2\pi}\langle y_{k+1}|p\rangle\langle p|e^{-\frac{i}{\hbar}\hat{T}\Delta t}|y_k\rangle \\
&= e^{-\frac{i}{\hbar}U(y_k)\Delta t}\int \frac{dp}{2\pi}\ e^{\frac{i}{\hbar}p(y_{k+1}-y_k)}\ e^{-\frac{i}{\hbar}\frac{p^2}{2m}\Delta t} \\
&= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{1/2}\ e^{\frac{i}{\hbar}\frac{m}{2\Delta t}(y_{k+1}-y_k)^2}\ e^{-\frac{i}{\hbar}U(y_k)\Delta t}. \qquad (1.12)
\end{aligned}$$

Collecting all of such terms yields the expression

$$K(x_f,t;x_i) \simeq \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{N/2}\prod_{l=1}^{N}\int dy_l\ e^{\frac{i}{\hbar}\sum_k \Delta t\left(\frac{m}{2\Delta t^2}(y_{k+1}-y_k)^2-U(y_k)\right)}, \qquad (1.13)$$

where we have adopted the notation $y_0 = x_i$ and $y_N = x_f$. Note that this expression is exact in the so-called continuous limit, in which $N \to \infty$ while the product $\Delta t\ N$ is held fixed. This limit is symbolically represented by introducing the functional integral notation:

$$\left(\frac{m}{2\pi i\hbar\Delta t}\right)^{N/2}\prod_{l=1}^{N}\int dy_l \to \int \mathcal{D}y \qquad (1.14)$$

At the same time, in the continuous limit one can replace sums and discrete differences with integrals and derivatives.

$$\sum_k \Delta t \to \int_0^t dt \qquad (1.15)$$

$$\frac{(y_{k+1}-y_k)^2}{\Delta t^2} \to \dot{y}^2. \qquad (1.16)$$

From a mathematical stand-point, this step is quite tricky, because generic paths y_k are, strictly speaking, non-differentiable objects. In fact, the replacement (1.16) should be regarded as a notation, rather than a proper mathematical statement.

In the new notation, the exact expression for the propagator becomes:

$$K(x_f,t;x_i,0) = \int \mathcal{D}y\ e^{\frac{i}{\hbar}S[y]}, \qquad (1.17)$$

where $S[y]$ is a functional which coincides the classical action of the system:

$$S[y] = \int_0^t dt'\left(\frac{m}{2}\dot{y}^2(t') - U[y(t')]\right). \qquad (1.18)$$

which is evaluated along the trajectory $y(t')$, satisfying the boundary conditions $y(0) = x_i$ and $y(t) = x_f$.

Eq. (1.17) is the famous Feynman path integral representation of the propagator. We emphasize that, in such an expression, operators and states in Hilbert spaces have been removed and the quantum mechanics is described entirely in terms of complex numbers. One of the main advantages of this mathematical representation is that the classical limit of quantum mechanics can be recovered by the stationary phase approximation. To this end we note that the complex phase term $e^{\frac{i}{\hbar}S[x]}$ is fast oscillating. This fact tends to suppress the contribution to the total path integral of all paths x for which $S[x] \gg \hbar$, due to destructive interference. In contrast, paths in the functional vicinity of the stationary point $\delta S[x] = 0$ coherently add. Thus in the small $\hbar$ limit, the path integral can be approximated by a very small bundle of trajectories focused around the solution $\bar{x}(t)$ of the Euler-Lagrange equation of motion:

$$\delta S[\bar{y}] = 0 \to \left[\frac{d}{dt}\frac{\partial}{\partial \dot{y}} - \frac{\partial}{\partial y}\right] \mathcal{L}(y, \dot{y})\Big|_{y(t)\to\bar{y}(t)} = 0 \tag{1.19}$$

We close this section by noting that this correspondence between classical and quantum mechanics is possible only for bosonic degrees of freedom. For fermions, it turns out to be impossible to represent a propagator in terms of c-numbers while preserving Pauli principle. In this case a new algebra (called after Grassmann) has to be introduced in order to describe a set of numbers with anti-commuting perators. An important consequence is that no classical limit for fermions can be recovered in the limit $\hbar \to 0$.

1.2.1 Connection between Imaginary Time Quantum Dynamics and Classical Thermodynamics

In this section, we discuss how the path integral formalism allows to establish important mathematical duality between the formalism to describe the dynamics of quantum particles and that describing the thermodynamics in the canonical ensemble. These connections are very practical tools: indeed, the same techniques developed to compute quantum mechanical propagators can be exported to evaluate thermal averages. At the same time, standard approaches for classical thermodynamics (such as Monte Carlo averages) can be used to compute quantum mechanical observables (Quantum Monte Carlo).

To establish this mapping we perform the analytic continuation of the Feynman path integral to the imaginary time axis (the so-called Wick rotation):

$$t \to -i\tau, \tag{1.20}$$

where τ is a real variable. On the imaginary time axis, the exponent appearing in the path integral becomes:

$$\frac{i}{\hbar}S[y] \equiv \frac{i}{\hbar}\int_0^t dt\left(\frac{m}{2}\dot{y}^2 - U(y)\right) \to -\frac{1}{\hbar}S_E[y] \equiv -\frac{1}{\hbar}\int_0^t dt\left(\frac{m}{2}\dot{y}^2 + U(y)\right), \tag{1.21}$$

which contains the so-called Euclidean action, $S_E[y] = \int_0^t dt\left(\frac{m}{2}\dot{y}^2 + U(y)\right)$. The imaginary time path integral is therefore written as follows:

$$K(x', t|x) \to \Delta(\tau) \equiv K(x', -it|x) = \int \mathcal{D}y\ e^{-\frac{1}{\hbar}S_E[y]}. \tag{1.22}$$

It is convenient to return to the discretized form of the imaginary time path integral:

$$\Delta(\tau) \simeq \left(\frac{m}{2\pi i\hbar}\right)^{N/2} \prod_{l=1}^{N} \int dy_l \; e^{-\frac{1}{\hbar}\sum_k \Delta t\left(\frac{m}{2\Delta t^2}(y_{k+1}-y_k)^2+U(y_k)\right)}. \tag{1.23}$$

We note that this expression is formally equivalent to the classical canonical partition function of a system of N particles interacting through an effective "Hamiltonian"

$$H_{eff} = \sum_k \left(\frac{m}{2\Delta t}(y_{k+1}-y_k)^2 + \Delta t U(y_k)\right) \tag{1.24}$$

at an "effective" temperature $\hbar$. Hence, the discretized Euclidean path integral describing the imaginary time dynamics of a quantum particle is formally analog to the partition function of a classical virtual "polymer", with harmonic springs connecting N beads, one for each different discretization time step. This analogy is very practical, as it allows to evaluate quantum path integrals using stochastic methods originally developed to perform classical canonical averages, such as the Metropolis algorithm

1.3 Systematic Effective Theory

Consider a physical system characterized by a wide gap between characteristic length (or time) scales. For example, we imagine some short-distance physics to become evident at some scale λ and we are interested in systematically describing phenomena which occur at much larger scales, e.g. for distances $\gg \lambda$. Borrowing the terminology form high-energy physics, we shall refer to the physics at scales $\gtrsim \lambda$ as the Ultra-Violet (UV) sector and to the physics at the soft-scale $\gg \lambda$ as Infra-Red (IR) sector.

The basic idea behind the Effective Theory (ET) formalism is very familiar: Any experimental probe with wavelength λ (or frequency ν) is insensitive to the details of the physics at length-scales $\ll \lambda$. As a consequence of this fact, as long as one is interested on the IR sector, all the hard-frequency details of a microscopic theory are irrelevant, and can be accurately mimicked by a set of effective parameters each multiplying specific operators. The structure of these operators can be deduced solely by symmetry arguments and dimensional analysis. The effective parameters multiplying the operators (called the coupling constants) are not predicted by the RT, but have to be computed from the underlying microscopic theory, or directly extracted from experimental data. Clearly, ETs are only applicable to physical systems which display a gap in characteristic scales, i.e. for which it is possible to separate IR from UV scales.

To illustrate in a practical example how rigorous ET are constructed let us consider a standard problem of classical electrostatics: how to determine the electrostatic potential produced by some arbitrary charge distribution $\rho(\mathbf{r})$. This is controlled by the first Maxwell's equation:

$$\nabla^2\phi(\mathbf{x}) = 4\pi\rho(\mathbf{x}) \tag{1.25}$$

A standard way to solve this problem would of course be to consider the exact solution

$$\phi(\mathbf{x}) = \int d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{x}-\mathbf{r}|} \tag{1.26}$$

We emphasize that the solution obtained this way would hold exact at *all distances* (at least those at which classical electrodynamics is a viable theory).

We now imagine that in the physical system we are investigating some crucial simplifications occur. Namely, (i) the charge distribution $\rho(\mathbf{r})$ is appreciably different from 0 only within a region of finite size $\sim \sigma$ (for simplicity we set the origin of the reference somewhere in the middle of the charge distribution) and (ii) the electrostatic phenomena we are studying take place far from the charge, i.e. at distances $|\mathbf{x}| \gg \sigma$. The standard theory of classical electromagnetism provides a technique for computing the solution of Eq. (1.25) under such circumstances: the multi-pole expansion. This can be readily obtained by expanding the denominator in Eq. (1.26) for $|\mathbf{r}-\mathbf{x}| \ll |\mathbf{x}|$. The first few terms in this expansion are the dipole and quardupole term, respecitvely:

$$\phi(\mathbf{x}) = \frac{Q}{|\mathbf{x}|} + \frac{\mathbf{P}\cdot\hat{\mathbf{x}}}{|\mathbf{x}|^2} + \ldots \tag{1.27}$$

where

$$\begin{aligned} Q &= \int d^3\mathbf{r}\ \mathbf{r}\ \rho(\mathbf{r}) &(1.28)\\ \mathbf{P} &= \int d^3\mathbf{r}\mathbf{r}\ \rho(\mathbf{r}) &(1.29)\end{aligned}$$

Corrections beyond the dipole term can be added in order to reach any desired finite accuracy.

We want to revisit this familiar problem in a modern perspective and from a slightly different standpoint. Rather than stating that we are interested in the physics at distances much larger than the charge size distribution, we imagine that this system is experimentally investigated using a probe characterized by wavelength λ. This sets a limit to the smallest length scale (or, equivalently, on the largest momenta $p_{UV} \sim 2\pi/\lambda$) that are resolved.

We are interested in developing a theory for this system which takes into account only of the available experimental information and (equally importantly) does not make any assumption about the physics at scales shorter than λ, for which we have no experimental information. In particular, this implies accounting for the fact that we have only limited knowledge about the structure of the charge distribution $\rho(\mathbf{x})$.

Our first step consists in evaluating the Fourier transform of the Maxwell's equation:

$$-\mathbf{k}^2\tilde{\phi}(\mathbf{k}) = 4\pi\ \tilde{\rho}(\mathbf{k}) \qquad \Rightarrow \qquad \tilde{\phi}(\mathbf{k}) = -\frac{4\pi\tilde{\rho}(\mathbf{k})}{\mathbf{k}^2} \tag{1.30}$$

A key important point to emphasize is that this expression formally contains information about Fourier components at arbitrary moment. However, we are assuming that we are probing the charge density distribution with a probe of wavelength

λ, i.e. up to some finite momentum $p_{UV} = \frac{2\pi}{\lambda}$. Assuming that Eq. (1.30) holds true even for momenta larger than this momentum scale corresponds to introducing unphysical information into our calculation, thus undesired model dependence.

Based on this considerations, we want to exclude from the calculation all modes with hard momenta from our theoretical models. In practice this is done in practice by introducing a UV-cutoff which quenches all hard modes. For example, this is implemented by a Gaussian cut-off:

$$\tilde{\phi}(\mathbf{k}) = -\frac{4\pi\tilde{\rho}(\mathbf{k})}{\mathbf{k}^2} \to \tilde{\phi}_\lambda(\mathbf{k}) = -\frac{4\pi\tilde{\rho}(\mathbf{k})}{\mathbf{k}^2}\; e^{-\frac{\mathbf{k^2}\lambda^2}{2}} \tag{1.31}$$

The position dependent solution is obtained from the inverse Fourier Transform of $\tilde{\phi}_\lambda(\mathbf{x})$:

$$\phi_\lambda(\mathbf{x}) = -\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{4\pi}{\mathbf{k}^2}\;\rho_\lambda(\mathbf{k}), \tag{1.32}$$

where

$$\tilde{\rho}_\lambda(\mathbf{k}) = e^{-\frac{\mathbf{k^2}\lambda^2}{2}}\,\tilde{\rho}(\mathbf{k}) \tag{1.33}$$

According to the properties of the Fourier Transform, Eq. (1.32) leads to the convolution of two inverse Fourier transforms:

$$\phi(\mathbf{x}) = F^{-1}\left[\frac{1}{\mathbf{k}^2}\right] \star F^{-1}\left[\tilde{\rho}_\lambda(\mathbf{k})\right] \tag{1.34}$$

We find:

$$F^{-1}\left[-\frac{1}{\mathbf{k}^2}\right] \;\equiv\; = \frac{1}{|\mathbf{x}|} \tag{1.35}$$

$$F^{-1}\left[\tilde{\rho}_\lambda(\mathbf{k})\right] \;\equiv\; \rho_\lambda(\mathbf{x}) \tag{1.36}$$

We recall that we are interested in describing the electrostatics in the IR sector, i.e. at scales much softer than then UV cut-off. Then, we can approximate $\tilde{\rho}_\lambda(\mathbf{k})$ with a Taylor series for small momenta:

$$\tilde{\rho}_\lambda(\mathbf{k}) \;=\; \tilde{\rho}_\lambda(\mathbf{k})|_{k\to 0} + \nabla\tilde{\rho}_\lambda(\mathbf{k})|_{k\to 0}\cdot\mathbf{k} + \dots \tag{1.37}$$

Now we note:

$$\tilde{\rho}_\lambda(\mathbf{k})|_{k\to 0} \;=\; \int d^3\mathbf{r}\rho_\lambda(\mathbf{r}) \equiv Q_\lambda \tag{1.38}$$

$$\nabla\tilde{\rho}_\lambda(\mathbf{k})|_{k\to 0} \;=\; i\int d^3\mathbf{r}\;\mathbf{r}\;\rho_\lambda(\mathbf{r}) \equiv i\mathbf{P}_\lambda \tag{1.39}$$

Thus

$$\tilde{\rho}_\lambda(\mathbf{k}) \;=\; Q_\lambda + i\mathbf{P}_\lambda\cdot\mathbf{k} + \dots \tag{1.40}$$

After computing the inverse Fourier Transform we obtain

$$\rho_\lambda(\mathbf{r}) = Q_\lambda\delta_\lambda(\mathbf{r}) + \mathbf{P}_\lambda\cdot\nabla\delta_\lambda(\mathbf{r}) \tag{1.41}$$

We note that, for the specific choice of smooth UV cut-off we are adopting, $\delta_\lambda(\mathbf{r})$ denotes a Gaussian function of variance λ. For a generic cut-off scheme, $\delta_\lambda(\mathbf{r})$ denotes an arbitrary representation of a Dirac-delta, smeared to the scale λ.

Finally we integrate by parts:

$$\begin{aligned} \phi_\lambda(\mathbf{x}) &= \int d^3\mathbf{r} \frac{1}{|\mathbf{x}-\mathbf{r}|} \left(Q_\lambda \delta_\lambda(\mathbf{r}) + \mathbf{P}_\lambda \cdot \nabla \delta_\lambda(\mathbf{r})\right) \\ &= \int d^3\mathbf{r}\ \delta_\lambda(\mathbf{r})\ \left(Q_\lambda - \mathbf{P}_\lambda \cdot \nabla_\mathbf{r}\right) \frac{1}{|\mathbf{x}-\mathbf{r}|} \end{aligned} \tag{1.42}$$

Thus, performing the integral over $d^3\mathbf{r}$ we recover the standard multipole expansion

$$\phi_\lambda(\mathbf{x}) \simeq \frac{Q_\lambda}{|\mathbf{x}|} + \frac{\mathbf{P}_\lambda \cdot \hat{\mathbf{x}}}{|\mathbf{x}|^2} + \ldots \tag{1.43}$$

Some comments are in order. First, we emphasize that in this derivation we have made explicit at each step that we are building a low-energy approximation which is valid only up to some length scale λ and we made no statement about the physics above such a scale. For example P_λ is the total electric dipole of the distribution, measured with a probe with wavelength λ. A variation of this scale: $\lambda \to \lambda' = \lambda + \delta\lambda$ would result in a small change in the effective coefficients $Q_\lambda, P_\lambda, \ldots$. This fact is referred by saying that the effective coefficients "run with the renormalization scale".

In the regime of validity of the expansion underlying an ET, predictions can be made to *any desired accuracy*, at the price of introducing more and more effective parameters. These have to be computed from the underlying microscopic theory or fitted directly to experiment. The need of adding more effective parameters to increase the theoretical accuracy is referred to by saying that the effective theory is *non-renormalizable*.

Part I

Dynamics in Open Systems

Chapter 2

Hamiltonian Dynamics

The theoretical description of the physics of complex systems is based on both equilibrium and non-equilibrium statistical physics.

Examples of a fundamental process which require an out of equilibrium description are conformational reactions, such as e.g. *protein or RNA folding*. Usually proteins are produced from RNA in coil configurations, and later reach their thermodynamically stable native state by a sequence of conformational changes. The folded molecule can be artificially driven to denature configurations by means of denaturants (such as temperature or urea concentration). After the denaturing conditions have been removed, the proteins spontaneously refold to their original native state. Clearly, the coil states obtained by means of denaturants are not in the global minimum of the free energy corresponding to the absence of denaturants. As a consequence, the protein re-folding reaction is an out-of-equilibrium process and the calculation of the folding rate clearly requires a non-equilibrium formalism. At the same time, in ordinary physiological conditions, a small fraction of proteins in the ensemble undergoes spontaneous unfolding and refolding transitions as a consequence of thermal fluctuations. The general theoretical result which allows to relate the dynamics of spontaneous unfolding and refolding and re-folding starting from out-of-equilibrium initial conditions is the so-called Onsager regression hypothesis (to be discussed below)

For the analysis of the dynamics of many phenomena occurring in complex systems a classical approach is satisfactory. For example, in describing structural rearrangements of macromolecules, quantum effects are in general expected to be small for most atomic species. On the other hand, the description of chemical reactions or charge transfer processes obviously requires a fully quantum approach.

In this chapter, we shall review and extend some aspects of equilibrium and out-of-equilibrium of classical statistical mechanics.

2.1 Statistical Mechanics of Hamiltonian Systems

2.1.1 Dynamics of Phase-Space Distributions

Let us begin our discussion by reviewing the standard notion of statistical *ensemble*. Let us consider a system made by N particles, and let $\Gamma(t) = (Q(t), P(t))$ denote the (time-dependent) phase-space coordinate, i.e. $Q = (q_1, \ldots, q_N)$ and

$P = (p_1, \ldots, p_N)$. The statistical ensemble is defined as the infinite set of macroscopically equivalent realizations of the same physical system, each one corresponding to a different initial condition (microstate), $\Gamma(0)$. The ensemble contains an infinite number of copies of the given system which span the phase space specified by the macroscopic constraints. Hence, the number of copies in a particular microstate $\Gamma(t)$ at time t may be assumed to change smoothly, as a function of Γ. Under such conditions, it is possible to define a *phase-space density distribution* $f(\Gamma, t)$, which assigns a probability density to each point in phase space:

$$\delta n(t) = f(Q, P, t)\delta P \delta Q \tag{2.1}$$

where $\delta n(t)$ is the fraction of copies which at time t have phase-space coordinates in the elementary volume $dPdQ$ located at the point (Q, P).

Let us consider the trajectories which at time t enter a small rectangular volume element. The fraction of the ensemble which enters the first face normal to the q_1 axis is

$$f(q_1, \ldots, p_N)\ \dot{q}_1(q_1, \ldots, p_N)\delta q_2 \ldots \delta q_N \delta P. \tag{2.2}$$

Similarly, the fraction of trajectories leaving through the second face parallel to the q_1 axis is

$$\begin{aligned} & f(q_1 + \delta q_1, \ldots, p_N,\ t)\ \dot{q}_1(q_1 + \delta q_1, \ldots, q_N)\delta q_2 \ldots \delta q_N \delta P \\ & = \left(f(q_1, \ldots, p_N,\ t) + \frac{\partial f}{\partial q_1}\delta q_1 \right) \ \left(q_1(q_1, \ldots) + \frac{\partial \dot{q}_1}{\partial q_1} \right) \ \delta P \delta Q \end{aligned} \tag{2.3}$$

Hence, the change in density $n(t)$ attributed to the flux in the q_1 direction is:

$$\frac{d}{dt}\delta n(t)_{q_1} = -\left(\dot{q}_1 \frac{\partial f}{\partial q_1} + f\ \frac{\partial \dot{q}_1}{\partial q_1} \right) \delta P \delta Q \tag{2.4}$$

Summing up all coordinate and momentum directions and dividing out the elementary phase-space volume we obtain:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\delta n(P, Q, t)}{dQdP} \right) = \left.\frac{\partial f}{\partial t}\right|_{P,Q} &= -\sum_{i=1}^{N} \left[f\ \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) + \dot{q}_i \frac{\partial f}{\partial q_i} + \dot{p}_i \frac{\partial f}{\partial p_i} \right] \\ &\equiv -f \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} - \dot{\Gamma} \cdot \frac{\partial f}{\partial \Gamma} \end{aligned} \tag{2.5}$$

On the other hand, from the usual differentiation chain rule, allowing for explicit dependence on time t we have

$$\frac{df}{dt} = \frac{\partial}{\partial t}\ f + \dot{\Gamma}\frac{\partial f}{\partial \Gamma} \tag{2.6}$$

Combining Eq.(2.5) with Eq. (2.6) we arrive to the Liouville's Eq:

$$\frac{d}{dt}f = -f\frac{\partial}{\partial \Gamma}\dot{\Gamma} \equiv -f\Lambda(\Gamma) \tag{2.7}$$

Note that, remarkably, the validity of this Eq. relies only on the assumption that microstates cannot be created or destroyed, while it does not require that the equations of motion of the system are generated by an Hamiltonian. However, if the system is Hamiltonian, then one can immediately show that $\Lambda(\Gamma) = 0$:

$$\Lambda(\Gamma) = \sum_i \frac{\partial}{\partial q_i} \cdot \frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i} \cdot \frac{\partial H}{\partial q_i} = 0. \tag{2.8}$$

Hence, for an Hamiltonian system, the Liouville's Eq. takes the simpler form:

$$\frac{df}{dt} = 0 \tag{2.9}$$

2.1.2 Evolution operator formalism in classical dynamics

Let us now restrict to Hamiltonian system and introduce the (f-)Liouvillian operator iL_f defined as:

$$iL_f \; f \equiv \left(\frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} + \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} \right) f \tag{2.10}$$

Note that the f-Liouvillian is a function of the initial phase-space coordinate Γ. Using the Liouvillian operator, the equation of motion of the distribution function f is formally equivalent to a Schrödinger equation

$$\frac{\partial f}{\partial t} = -iL_f f \tag{2.11}$$

We can formally solve the equation of motion in terms of the evolution operator:

$$f(\Gamma, t) = e^{-iL_f t} f(\Gamma, 0). \tag{2.12}$$

Let us now derive the equation of motion for the en arbitrary function $B(\Gamma)$ of the phase-space coordinate Γ of the system. In general, we shall refer to such functions as to phase-variables. Note that, by definition, phase variables depend on time only implicitly, through the phase-space coordinate Γ. Hence, using the chain rule for time differentiation, we find:

$$\dot{B}(\Gamma) = \dot{\Gamma} \cdot \frac{\partial}{\partial \Gamma} B \equiv iL_p B(\Gamma), \tag{2.13}$$

where the operator iL_p is referred to as the phase-variable or p-Liouvillian operator. As before, we can give a formal solution of the equation of motion of any phase variable as

$$B(t) = e^{iL_p t} \; B(0). \tag{2.14}$$

In general, it is possible to show that the f-Liouvillian is the adjoint of the p-Liouvillian operator. Such a property follows directly from the identity

$$\int d\Gamma f(0) iL_p B(\Gamma) = - \int d\Gamma B(\Gamma) iL_f f(0), \tag{2.15}$$

which can be easily proven by integrating by parts the LHS and discarding surface terms. In addition, we observe that the p-Liouvillian and the f-Liouvillian differ for the compression factor $\Lambda(\Gamma)$. Hence, for Hamiltonian systems, for which such a term is identically zero, one has

$$iL_p = iL_f \equiv iL \qquad \text{(for Hamiltonian systems)} \tag{2.16}$$

and the Liouvillians are Hermitian operators.

In complete analogy with quantum mechanics, it is possible to give two equivalent representations of the dynamics of average phase variables. In the Heisemberg representation, one chooses the initial vlaue of the phase from an initial distribution $f(\Gamma) = f(\Gamma, 0)$ and consider the evolution of the phase variable $B(\Gamma(t))$:

$$\langle B(t)\rangle_H = \int d\Gamma B(t) f(\Gamma) = \int d\Gamma (e^{iL_p} B(0)) f(\Gamma) \tag{2.17}$$

Alternatively, it is possible to consider first the evolution of the density function from an initial distribution and the average the phase variable over the final distribution, i.e. the Schrödinger representation:

$$\langle B(t)\rangle_S = \int d\Gamma B(\Gamma) f(\Gamma, t) = \int d\Gamma B(\Gamma) e^{-iL_f} f(\Gamma, 0) \tag{2.18}$$

It is possible to show that the two representation are equivalent, i.e. that

$$\langle B(t)\rangle_S = \langle B(t)\rangle_H. \tag{2.19}$$

Of course, it is also possible to introduce alternative representations of the dynamics in which the time dependence is shared by both the phase variable and density function. This is the analog of the quantum mechanical interaction representation.

2.1.3 Dyson Representation of the Propagator

The Dyson representation is a self-consistent Eq. which involves the propagator of the system. To derive it, we begin by defining the *resolvent* of the operator e^{-At} which is defined by its .

$$G_A(s) \equiv \int_0^\infty dt e^{-At}\, e^{-st} = \frac{1}{A+s} \tag{2.20}$$

We now apply the general operator Eq.

$$(A+B)^{-1} = A^{-1} - A^{-1}B(A+B)^{-1} \tag{2.21}$$

Which leads to

$$(s+A+B)^{-1} = (s+A)^{-1} - (s+A)^{-1}B(s+A+B)^{-1} \tag{2.22}$$

Substituting the Laplace integrals for $(A+s)^{-1}$ and $(A+B+s)^{-1}$ this into this Eq. we find

$$\int_0^\infty dt e^{-st} e^{-(A+B)t} = \int_0^\infty dt e^{-st} \left(e^{-At} - \int_0^t dt_1 e^{-At_1} B e^{-(A+B)(t-t_1)} \right) \tag{2.23}$$

Since this Eq. must hold for every time t the equality extends to the integrands and leads to:

$$e^{-(A+B)t} = e^{-At} - \int_0^t dt_1 e^{-At_1} B e^{-(A+B)(t-t_1)} \tag{2.24}$$

Similarly, we can define an analog Eq. for the backwards propagator

$$e^{(A+B)t} = e^{At} + \int_0^t dt_1 e^{At_1} B e^{(A+B)(t-t_1)} \tag{2.25}$$

This is the Dyson representation, which relates the propagator defined by the A and B operators.

2.2 Equilibrium Time Correlation Functions

We define an equilibrium distribution function f_0 as one which satisfies the Liouville's equation in the form:

$$\frac{\partial}{\partial t} f_0 = -iL_f f_0 = 0. \tag{2.26}$$

This condition implies that the equilibrium average of any phase variable is a stationary quantity:

$$\begin{aligned} \frac{d}{dt}\langle B(t)\rangle_0 &= \frac{d}{dt}\int d\Gamma f_0(\Gamma) e^{iL_p t} B(\Gamma,0) \\ &= \int d\Gamma f_0(\Gamma) iL_p e^{iL_p t} B(\Gamma,0) \\ &= -\int d\Gamma (iL_f f_0(\Gamma)) e^{iL_p t} B(\Gamma,0) = 0 \end{aligned} \tag{2.27}$$

where we have used the relationship between p- and f- Liouvillian operators.

We define the equilibrium time correlation function of two phase variables A and B as

$$C_{AB}(t) \equiv \int d\Gamma \; f_0(\Gamma) \; B^*(\Gamma,0) e^{iL_p t} A(\Gamma,0) = \langle A(t) B^*(0)\rangle_0 \tag{2.28}$$

After some straightforward manipulation, one arrives at the intuitive relation:

$$C_{AB}(t) = \int d\Gamma f_0(\Gamma) A(t_1 + t) B^*(t_1), \tag{2.29}$$

where t_1 is some arbitrary instant. Note that this is true only if the average is performed over equilibrium distributions. Using the time-translation invariance of the distribution function, one can show that

$$C_{AB}(t) = \langle A(t) B^*(0)\rangle_0 = \langle A(0) B^*(-t)\rangle_0 \tag{2.30}$$

Finally, we observe that

$$C^*_{AB}(t) = \langle A^*(0) B(-t)\rangle_0 = C_{BA}(-t). \tag{2.31}$$

One can use the autocorrelation function of a phase variable A to define a norm in Liouville's space:

$$||A||^2 \equiv \int d\Gamma f_0(\Gamma) A(\Gamma) A^*(\Gamma) = \int d\Gamma f_0(\Gamma) |A(\Gamma)|^2 \geq 0 \tag{2.32}$$

If the Liouvillian operator is Hermitian, the evolution operator e^{iLt} is unitary and the norm is preserved

$$||A(t)||^2 = ||A(0)||^2. \tag{2.33}$$

2.3 Relating Equilibrium and Non-Equilibrium Fluctuations: The Regression Hypothesis

The goal of this section is to introduce Lars Onsager's regression hypothesis, a consequence of the fluctuation-dissipation theorem[1]. Here, we derive the regression hypothesis from the principles of statistical mechanics. The regression hypothesis states that **the regression of microscopic thermal fluctuations at equilibrium follows the macroscopic law of relaxation of small non-equilibrium disturbances** [2].

In order to understand this hypothesis, consider an observable A for a system at thermal equilibrium. Such property fluctuates in time with spontaneous microscopic fluctuations:

$$\delta A(t) \equiv A(t) - \langle A \rangle, \tag{2.34}$$

Here, is the instantaneous value of the observable and is the equilibrium ensemble average

$$\langle A \rangle = \frac{\mathrm{Tr}\ A(t)\ e^{-\beta H_0}}{\mathrm{Tr}\ e^{-\beta H_0}} \tag{2.35}$$

The average correlation between an instantaneous fluctuations of A at time t and an instantaneous fluctuation of A at time zero is described by the correlation function

$$C(t) = \langle \delta A(t) \delta A(0) \rangle = \langle A(t) A(0) \rangle - \langle A^2 \rangle \tag{2.36}$$

At short time $\delta A(t)$ and $\delta A(0)$ are completely correlated:

$$\lim_{t \to 0} C(t) \neq 0 \tag{2.37}$$

At large times $\delta A(t)$ and $\delta A(0)$ decorrelate and $C(t)$ vanishes, i.e.

$$\lim_{t \to \infty} C(t) = 0 \tag{2.38}$$

This is the regression of microscopic thermal fluctuations which is referred in the Onsager's hypothesis. Hence, Onsager's regression hypothesis can be formulated as follows

$$\frac{\overline{A}(t) - \langle A \rangle}{\overline{A}(0) - \langle A \rangle} = \frac{\langle \delta A(t) \delta A(0) \rangle}{\langle (\delta A(0))^2 \rangle} \tag{2.39}$$

---

[1] H. B Callen and T. A. Welton, Phys. Rev. 83, 34 (1951)

[2] L. Onsager, Phys. Rev.37, 405 (1931); 38, 2265 (1931))

In this Eq. $\overline{A}(t)$ describes the *macroscopic* relaxation of the observable towards its equilibrium value while the system evolves from an initial state that is not far from equilibrium to its final state in equilibrium with a thermal reservoir.

In order to derive Eq.(2.39) from the postulates of statistical mechanics, consider preparing the system in a state that is not far from equilibrium by applying to the system a weak perturbational field f (e.g., an electric field) that couples to the dynamical variable (e.g., the instantaneous total dipole moment). The external field is assumed to be so weak that the perturbation Hamiltonian ΔH, written as an expansion in powers of f, can be approximated to be first order in the field as follows,

$$\Delta H = -f A(0) \tag{2.40}$$

Assume that the perturbation was applied until the system is equilibrated according to the total Hamiltonian

$$H = H_0 + \Delta H \tag{2.41}$$

The macroscopic relaxation of the system is analyzed by switching off the external perturbational field and computing the evolution of the non-equilibrium ensemble average

$$\overline{A}(t) = \frac{\text{Tr } A(t) e^{-\beta(H_0+\Delta H)}}{\text{Tr } e^{-\beta(H_0+\Delta H)}} \tag{2.42}$$

The ensemble average is as usually defined by considering equivalent copies of *initial conditions*. Consequently, $\overline{A}(t)$ is a time-dependent average, because the distribution of initial condition in not the equilibrium generated by the unperturbed Hamiltonian. Expanding in $\beta\Delta H$ we have

$$\begin{aligned} \overline{A}(t) \quad &\sim \quad \frac{\text{Tr } A(t)(1-\beta\Delta H)+\ldots)e^{-\beta(H_0)}}{\text{Tr } (1-\beta\Delta H)+\ldots)e^{-\beta(H_0+\Delta H)}} \\ &= \quad \frac{\langle A(t)\rangle}{1-\beta\langle \Delta H\rangle} - \beta\frac{\langle \Delta H A(t)\rangle}{1-\beta\langle \Delta H\rangle} \end{aligned} \tag{2.43}$$

Further expanding in ΔH we find:

$$\overline{A}(t) \simeq \langle A\rangle - \beta(\langle \Delta H A(t)\rangle - \langle \Delta H\rangle) + \mathcal{O}(\Delta H^2) \tag{2.44}$$

If we now use the assumption (2.40) and take the ratio, we obtain (2.39). Hence, the kinetic of approach to thermal equilibrium of a system which has been set out of equilibrium by a small perturbation is equivalent to the kinetics of return to equilibrium of small spontaneous fluctuations. The consequence is deep: we can study the transitions which occur at equilibrium by slightly perturbing the system out of equilibrium and following the regression to thermal equilibrium.

Chapter 3

Stochastic Dynamics

This chapter is devoted to a non-relativistic discussion of the dynamics of open systems, i.e. systems coupled to an environment, or heat-bath. Examples of open classical and quantum open systems are ubiquitus in physics. In particular, the classical description these systems is the key to understand the behaviour in time of soft-condensed and biological matter. A quantum treatment of quantum systems is crucial to explore the physical implications of fundamental concepts of quantum mechanics, such as measurement and decoherence. being extensively studied, e.g. with cold-atoms.

The dynamics of open systems is shaped by irreversible phenomena, such as e.g. thermalisation. In order to analyse how irreversibility emerges starting from a microscopic, symplectic Hamiltonians we exploit the effective theory approach.

3.1 Classical Open Systems

In the next few section, we shall assume Hamilton's equation and show how, by systematically integrating out the dynamics of the fast (ultra-violet, UV) degrees of freedom from the phase-space, it is possible to generate a low-resolution effective theory for the remaining slow degrees of freedom. The loss of information resulting from integrating out part of the phase space is reflected by the fact that the resulting low-energy dynamics is non-deterministic, i.e. stochastic. Furthermore, this procedure elucidates the origin of irreversibility, i.e. explain why the time-reversal invariance of symplectic systems is effectively broken, giving raise to irreversible phenomena such as equilibration.

Let us begin our discussion by reviewing some very basic notions of analytic mechanics and statistical physics. We consider a system made by N particles, and we let $\Gamma(t) = (Q(t), P(t))$ denote the (time-dependent) phase-space coordinate, i.e. $Q = (q_1, \ldots, q_N)$ and $P = (p_1, \ldots, p_N)$.

We shall assume that the evolution of the phase-space variable is set by Hamilton's equations:

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned} \tag{3.1}$$

where $H = H(\Gamma)$ is the system's Hamiltonian.

Given an initial condition $\Gamma(0) = \Gamma_0$, the system spontaneously evolves according to this set of equations. Thus, at any time t, the state of this physical system is entirely determined by its phase-space coordinate $\Gamma(t)$.

In practice, the initial configuration of the system Γ_0 can never be determined with an infinite accuracy. Furthermore, for sufficiently complex systems the resulting dynamics is chaotic. As a result, even if we were able to solve the equation of motion with arbitrary accuracy, we would not be able to completely determine the configuration at time t.

This "no-go" statement sets the stage for the development of classical mechanics, which is based on Gibb's notion of *statistical ensemble*. This is defined as the infinite set of macroscopically equivalent realizations of the same physical system, each one corresponding to a different initial condition (microstate), $\Gamma(0)$. The ensemble contains an infinite number of copies of the given system which span the phase space specified by the macroscopic constraints. Hence, the number of copies in a particular microstate $\Gamma(t)$ at time t may be assumed to change smoothly, as a function of Γ.

Under such conditions, it is possible to define a *phase-space density distribution* $P(\Gamma, t)$, which assigns a probability density to each point in phase space:

$$\delta n(\Gamma, t) = P(\Gamma, t)\delta\Gamma \tag{3.2}$$

where $\delta n(\Gamma, t)$ is the fraction of copies which at time t have phase-space coordinates in the elementary volume $d\Gamma$ centered at the phase-space point Γ.

The general equation governing the change in time of the phase space density distribution is the well known Liouville's equation:

$$\frac{d}{dt}P = -P\frac{\partial}{\partial\Gamma}\dot{\Gamma} \equiv -P\Lambda(\Gamma) \tag{3.3}$$

Remarkably, the validity of this equation relies only on the assumption that microstates cannot be created or destroyed, while it does not require that the equations of motion of the system are generated by an Hamiltonian. However, if the system is indeed Hamiltonian, then one can immediately show that the term on the right-hand-side is $\Lambda(\Gamma) = 0$:

$$\Lambda(\Gamma) = \sum_i \frac{\partial}{\partial q_i}\cdot\frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i}\cdot\frac{\partial H}{\partial q_i} = 0. \tag{3.4}$$

Hence, for an Hamiltonian system, the Liouville's Eq. takes the simpler form:

$$\frac{dP(\Gamma, t)}{dt} = 0. \tag{3.5}$$

3.1.1 From Hamiltonian to Stochastic Dynamics

We now specialise on physical systems which display a gap in the characteristic frequency scales, i.e. a decoupling between fast and slow degrees of freedom. The purpose of this section is to show that, in this case, once the fast degrees of freedom have been integrated out from the phase-space, the equations of motion for the slow degrees of freedom acquire dissipative and stochastic terms which are associated to irreversibility.

The mathematical procedure which implements this program is known in the statistical mechanics literature as the Mori-Zwanzing projection formalism and is applicable to arbitrary hamiltonian systems. Unfortunately, its generality comes at the prize of not being very physically transparent, a feature which has significantly limited its practical applications. Thus, for sake of clarity, here we choose to specialize on the case in which the fast degrees of freedom in the system are represented by an infinite collection (or "bath") of harmonic oscillators, linearly coupled to the low-energy degrees of freedom (classical Caldeira-Laggett model). In this case, the Mori-Zwanzig formalism becomes much simpler and, most importantly, much more physically transparent, to a point that the physical origin of the stochastic behavior becomes evident. At a first sight, the choice of specialising on a bath made of harmonic oscillator models may be considered somewhat artificial or restrictive. On the other hand, as long as we are not interested in the dynamics of the environment, all we need to know are the characteristic time scales associated to the response of the heat-bath to changes in the system, i.e. the intrinsic memory effects. These physical scales can be tuned by an appropriate choice of the spectrum of frequencies ω_α and of the associated coefficients c_α. Indeed we recall that the low-energy dynamics of the system is expected to be insensitive to the *details* of the high energy physics.

For sake of simplicity, and without loss of generality, we consider the case in which there is a single slow-degree of freedom in the system. Generalisation to the multi-dimensional case is straightforward. The Hamiltonian reads (Zwanzig 1973, Ford and Kac 1987, Pollack 1986):

$$H = \frac{p^2}{2m} + U(x) + \sum_\alpha \left[\frac{p_\alpha^2}{2\mu_\alpha} + \frac{\mu_\alpha \omega_\alpha^2}{2} \left(q_\alpha - \frac{c_\alpha}{\mu_\alpha \omega_\alpha^2} x \right)^2 \right] \tag{3.6}$$

The corresponding Hamilton equations for the low-energy variable (which we shall refer to as system's variables) are:

$$\dot{x} = \frac{p}{m} \tag{3.7}$$

$$\dot{p} = -\frac{\partial}{\partial x} U(x) + \sum_\alpha c_\alpha \left(q_\alpha - \frac{c_\alpha}{\mu_\alpha \omega_\alpha^2} x \right) \tag{3.8}$$

while for the bath variables we obtain:

$$\dot{q}_\alpha = \frac{p_\alpha}{\mu_\alpha} \tag{3.9}$$

$$\dot{p}_\alpha = -\mu_\alpha \omega_\alpha^2 q_\alpha + c_\alpha x \tag{3.10}$$

The next step consists in integrating out the bath degrees of freedom. To this end we need to solve Newton's equation for bath variables, i.e. a set of ordinary inhomogeneous second order differential equations. This this goal, we need to sum the general solution of the associated homogeneous equation, to a special solution of the inhomogeneous equation. Finding the general solution of the homogeneous equation is straightforward and gives:

$$q_\alpha(t) = q_\alpha(t_0) \cos(\omega_\alpha (t - t_0)) + \frac{p_\alpha(t_0)}{\mu_\alpha} \sin(\omega_\alpha (t - t_0)) \tag{3.11}$$

In order to provide a special solution of the inhomogeneous equation it is convenient to resort on the Green's function method. Regarding $x(\tau)$ as a given function we write

$$\mu_\alpha \partial_t^2 q_\alpha(t) = -\mu_\alpha \omega_\alpha^2 q_\alpha + c_\alpha x(\tau) \tag{3.12}$$

After transforming to Fourier space:

$$-\omega^2 \mu_\alpha \tilde{q}_\alpha(\omega) + \mu_\alpha \omega_\alpha^2 \tilde{q}_\alpha(\omega) = c_\alpha \tilde{x}(\omega) \tag{3.13}$$

Thus

$$\tilde{q}_\alpha(\omega) = \frac{c_\alpha \tilde{x}(\omega)}{\mu_\alpha(\omega_\alpha^2 - \omega^2)} \tag{3.14}$$

Transforming back to time representation

$$q_\alpha(t) = \frac{c_\alpha}{\mu_\alpha} \int_{t_0}^{\infty} d\tau K(t-\tau)\ x(\tau) \tag{3.15}$$

where the lower bound comes from the initial condition. The Green's function reads

$$K(t-\tau) = \int \frac{d\omega}{2\pi} \frac{e^{-i\omega(t-\tau)}}{(-\omega^2 + \omega_\alpha^2 + i\epsilon)} = \theta(t-\tau)\frac{1}{\omega_\alpha}\sin(\omega_\alpha(t-\tau)) \tag{3.16}$$

We emphasise that the Fourier transform is supplemented by a "$i\epsilon$" prescription to enforce causality. This is done as usual by excluding the contribution from the pole that describes to propagation for $t - \tau < 0$.

The complete solution is then

$$\begin{aligned} q_\alpha(t) &= q_\alpha(t_0)\cos(\omega_\alpha(t-t_0)) + \frac{p_\alpha(t_0)}{\mu_\alpha}\sin(\omega_\alpha(t-t_0)) \\ &\quad + \frac{c_\alpha}{\mu_\alpha\omega_\alpha}\int_{t_0}^{t} d\tau\ x(\tau)\ \sin(\omega_\alpha(t-\tau)) \end{aligned} \tag{3.17}$$

Plugging this back into the equation of motion for the slow variables yields

$$\begin{aligned} \dot{p} &= -\frac{\partial}{\partial x}U(x) - \sum_\alpha \frac{c_\alpha^2}{\mu_\alpha\omega_\alpha^2} x(t) \\ &\quad + \sum_\alpha \frac{c_\alpha^2}{\mu_\alpha\omega_\alpha^2}\int_{t_0}^{t} ds\ \omega_\alpha \sin\left[\omega_\alpha(t-s)\right]\ x(s) \\ &\quad + \sum_\alpha c_\alpha \left[q_\alpha(t_0)\cos\left[(\omega_\alpha(t-t_0)\right] + \frac{p_\alpha(0)}{\mu_\alpha\omega_\alpha}\sin\left[\omega_\alpha(t-t_0)\right]\right] \end{aligned} \tag{3.18}$$

The last term will be denoted with $F(t)$ and defines a fluctuating force. Second and last term can be combined to give:

$$\dot{p} = -\frac{\partial}{\partial x}U(x) + \sum_\alpha \frac{c_\alpha^2}{\mu_\alpha\omega_\alpha^2}\left[\int_{t_0}^{t} ds\ x(s)\frac{\partial}{\partial s}\cos(\omega_\alpha(t-s)) - x(t)\right] + F(t) \tag{3.19}$$

Finally, after integrating by parts the second term we obtain the so-called generalized Langevin equation:

$$\dot{p} \;=\; -\frac{\partial}{\partial x}U(x) - \frac{M}{2}\int_{t_0}^{t} dsk(t-s)\dot{x}(s) + \xi(t) \tag{3.20}$$

where

$$k(t-s) \;=\; \sum_\alpha \frac{2c_\alpha^2}{M\mu_\alpha\omega_\alpha^2}\cos\left[\omega_\alpha(t-s)\right] \tag{3.21}$$

$$\xi(t) \;=\; \sum_\alpha c_\alpha \left[\left(q_\alpha(0) - \frac{c_\alpha}{\mu_\alpha\omega_\alpha^2}x(t_0)\right)\cos\left[(\omega_\alpha(t-t_0)\right] + \frac{p_\alpha(0)}{\mu_\alpha\omega_\alpha}\sin\left[\omega_\alpha(t-t_0)\right]\right] \tag{3.22}$$

Where

$$\xi(t) = F(t) - Mk(t-t_0)x(t_0). \tag{3.23}$$

Eq. (3.20) is one of the fundamental results in the theory of open classical systems. It is basically a Newton's equation, in which the two additional terms describe the effective forces introduced by the coupling of the system to the environment.

In particular, the term $-\int_{t_0}^{t} dsMk(t-s)\dot{x}(s)$ introduces an effective dissipative friction. The so-called *memory kernel* $k(t-s)$ encodes the information about the characteristic time scales at which the environment absorbs the energy provided by the system's propagation. In the spirit of effective theory, we shall assume that the environment dynamics is much faster than that of the system, i.e that the Fourier transform of the system variable $\tilde{x}(\omega)$ is dominated by components with frequency much smaller than the frequency range where the spectral function of the heat-bath

$$\tilde{k}(\omega) \;=\; \sum_\alpha \frac{2c_\alpha}{M\mu_\alpha\omega_\alpha^2}\left(\delta(\omega-\omega_\alpha) + \delta(\omega+\omega_\alpha)\right) \tag{3.24}$$

is peaked. In this limit, several simplification occurs. Firstly $\gamma(t-t')$ can be expanded as follows

$$k(t-s) = c_0\delta(t-s) + \ldots \tag{3.25}$$

where … denotes terms proportional to derivatives of delta-functions. In particular, we consider here the so-called Ohmic limit, in which only the lowest-order term is retained in this expansion. In this limit, the time scale it takes the oscillators to thermalize around a given system configuration is instantaneous, thus

$$k(t-s) \simeq 2\gamma\ \delta(t-s) \tag{3.26}$$

Once the Ohmic limit is applied to the generalized Langevin equation one recovers what in the literature is known simply as the Langevin Equation:

$$\dot{p} \;=\; -\frac{\partial}{\partial x}U(x) - M\gamma\dot{x}(s) + \xi(t) \tag{3.27}$$

The fluctuating force $\xi(t)$ describes the effects of the “kicks” provided by the environment on the system. We emphasize that this term does not depend on the instantaneous configuration $x(t)$ but only on the initial system position $x(t_0) = x_0$ and on the initial heat bath configurations and momenta $q_\alpha(0) = q_\alpha^0$ and $p_\alpha(0) = p_\alpha^0$, which are unknown. Furthermore, in the Ohmic limit, the second term in the right-hand-side of Eq. (3.23) vanishes for any finite time and can be dropped. In the lack of complete information on the initial condition, $\xi(t)$ must be regarded as a stochastic variable.

In order to determine the first moments of its distribution, we take the thermodynamic limit (large number of oscillators) and assume that the initial variables $q_\alpha(t_0)$ and $p_\alpha(t_0)$ can be considered to be distributed according to the canonical Boltzmann's distribution:

$$\rho_{eq}(p_\alpha(t_0), q_\alpha(t_0)|x(t_0)) = \frac{1}{Z} e^{-\frac{1}{k_B T}\sum_\alpha \left[\frac{p_\alpha^2(t_0)}{\mu_\alpha} + \frac{1}{2}\mu_\alpha \omega_\alpha^2 \left(q_\alpha^2(t_0) - \frac{c_\alpha}{\mu_\alpha \omega_\alpha} x(t_0)\right)\right]}. \tag{3.28}$$

Then, the mean value and time auto-correlation function of the stochastic force can be computed by performing averages over the equilibrium Boltzmann distribution. The result is

$$\langle \xi(t) \rangle = 0 \tag{3.29}$$

$$\langle \xi(t+t_0)\xi(t_0) \rangle = 2\ k_B T M \gamma \delta(t) \tag{3.30}$$

This result is referred in the literature as the fluctuation-dissipation theorem. It establishes a highly non-trivial relationship between the energy which is provided by the environment to the system by the stochastic force and the energy which is dissipated. Below we shall show that this fine energetic balance is responsible for the fact that the system eventually reaches thermal equilibrium with its heat-bath.

We derived the Langevin equation 3.27 by exploiting the decoupling of intrinsic time scales associated to the system and heat-bath dynamics. We now take this idea even further and consider a limit in which the dynamics of the system is so slow that its characteristic frequency scales are such that $\gamma/\omega \gg 1$. To monitor what happens to the Langevin dynamics, let's transform equation (3.27) to Fourier space:

$$-M\omega^2 \tilde{x}(\omega) = \int_0^t d\tau e^{-i\omega t} \frac{\partial}{\partial x} U(x) + i\omega\gamma\tilde{x}(\omega) + \tilde{\xi}(\omega) \tag{3.31}$$

If we multiply both sides by $1/\gamma$ we realize that the inertial term in the left-hand-side contains a factor $\frac{\omega}{\gamma} \ll 1$ and can be neglected. Returning to time representation we obtain the so-called over damped Langevin equation:

$$\dot{x} = -\frac{D}{k_B T}\frac{\partial}{\partial x} U(x) + \eta(t) \tag{3.32}$$

where $D = \frac{k_B T}{M\gamma}$ is called the diffusion coefficient and $\eta(t)$ is a rescaled stochastic force, which obeys the fluctuation-dissipation relationship:

$$\langle \eta(t) \rangle = 0 \tag{3.33}$$

$$\langle \eta(t)\eta(t_0) \rangle = 2D\ \delta(t - t_0) \tag{3.34}$$

3.1.2 Aspects of Stochastic Calculus

From a mathematical standpoint the Langevin equations (3.27) and (3.32) are quite puzzling. For example, on the one hand, the left-hand side of equation (3.32) contains a time derivative. On the other hand, the right-hand side contains the stochastic force which is delta-correlated, thus discontinuous for any t.

The mathematician's way to go around this problem is to abandon Rieman's calculus and introduce new definitions of integral and derivative, which are tailored for the description of stochastic processes (stochastic calculus). From a physical point of view, and in particular from a renormalisation group perspective, these mathematical issues are associated with the fact that Langevin equations define an effective theory. As such, they are only supposed to provide a description of the slow dynamics, i.e. to be meaningful and applicable for times much longer than an ultraviolet cut-off scale Δt and to be meaningless beyond such a scale. Consequently, the delta-function appearing in the stochastic force auto-correlation function needs to be smeared to acquire a width of order Δt (see also the discussion of renormalization in Lepage's lectures [16]).

In practice, the constraint set by the effective theory is implemented by stating that Eq. (3.32) is equivalent to the following stochastic process:

$$x(t+\Delta t) = x(t) - \frac{D\Delta t}{k_B T}\frac{\partial}{\partial t}U[x(t)] + \int_{t_i}^{t_i+\Delta t} \eta(\tau)d\tau, \tag{3.35}$$

while no statement made can be made for time intervals smaller than Δt. Equation (3.35) is completely general, but does not yet define the stochastic derivative. To this end, we need to specify the convention used to define the discretization of the derivate and when the stochastic variable is sampled, i.e. to provide a rule on how to compute the integral in the right hand side. For example, in the so-called Ito Calculus, the first order stochastic differential equation (3.35) is defined by the following rule:

$$x(t+\Delta t) = x(t) - \frac{D\Delta t}{k_B T}\nabla U[x(t)] + \eta(t)\Delta t \tag{3.36}$$

where $\eta(t)$ is a random number sampled from a distribution which obeys then

$$\begin{aligned} \langle \eta(t) \rangle &= 0 & (3.37)\\ \langle \eta(t_i)\eta(t_j)\rangle &= \frac{2D}{\Delta t}\delta_{ij} & (3.38) \end{aligned}$$

Notice that $\delta_{ij}/\Delta t$ provides a representation of delta-function $\delta(t-t')$, smeared at the UV cut-off time scale Δt.

It is convenient to absorb the $1/\Delta t$ factor into the definition of the stochastic variable: $W(t) \equiv \eta(t)\Delta t$. Then the probability distribution for the stochastic displacements $W(t)$ obeys

$$\begin{aligned} \langle W(t) \rangle &= 0 & (3.39)\\ \langle W(t_i)W(t_j)\rangle &= 2D\Delta t\ \delta_{ij} & (3.40) \end{aligned}$$

The simplest choice of distribution which satisfies these relationships is a Gaussian of variance $\sqrt{2D\Delta t}$:

$$\mathcal{P}[W] = \frac{1}{\sqrt{4\pi D\Delta t}} e^{-\frac{W^2}{4D\Delta t}} \tag{3.41}$$

An alternative choice of discretization of the over damped Langevin equation follows the so-called Stratonovich convention:

$$x(t+\Delta t) = x(t) - \frac{D\Delta t}{k_B T}\frac{1}{2}\nabla\left(U[x(t)] + U[x(t+\Delta t)]\right) + W(t) \tag{3.42}$$

So far we discussed the notion of stochastic derivative. It is useful to define also the stochastic generalization of integration. Given a time series $\{x_k\}_k$ the so-called Ito-Integral is defined as follows

$$(I)\int_{x_0}^{x_f} dx f(x) \equiv \lim_{\substack{\Delta t\to 0\\ N\to\infty}} \sum_{k=0}^{N-1} \left(x(t_{k+1}) - x(t_k)\right)\, f(x_k). \tag{3.43}$$

A similar expression exists for the Stratonovich definition.

Interestingly, the fundamental theorem of Calculus can been shown to lead to different relationships, depending on the specific choice of discretisation convention. In particular, Ito integrals of derivatives can be shown to obey the modified relationship:

$$(I)\int_{x_0}^{x_f} dx f'(x) = f(x_f) - f(x_i) - D\int_0^t ds f''[x(s)] \tag{3.44}$$

where $x(t) = x_f$ and $x(t_0) = x_i$ and the last term denotes a standard (i.e. Riemann) integral and the path $x(s)$ is obeys the boundary values $x(s=0) = x_0$, $x(t) = x_f$. It is important to empasize that this results holds in the probabilistic sense, i.e. after averaging over many independent stochastic processes. In contrast, in the Stratonovich convention, the fundamental theorem of calculus coincides with that of Riemann's Calculus.

$$(S)\int_{x_0}^{x_f} dx f'(x) = f(x_f) - f(x_i) \tag{3.45}$$

From this point on we shall assume the Ito convention, which has the advantage of defining a Markov process and thus is most commonly adopted in numerical simulations.

3.2 Mori-Zwanzig Projection Formalism

In this section we analyze the derivation of stochastic dynamics from an underlying Hamiltonian dynamics from a more general and formal standpoint, by illustrating the Mori-Zwanzig projection formalism. The goal is to derive the dynamics of an arbitrary function of the phase variable $A(\Gamma)$, for system evolving under Eq.s of motion which preserve the equilibrium distribution function. The fundamental result of Mori-Zwanzig Theory (MZT) is that the phase variable obeys a generalized Langevin equation. For this reason, we shall refer to A as to the Langevin (phase) variable.

For sake of definiteness, we shall focus on the case in which the equilibrium distribution is the canonical distribution:

$$f_c(\Gamma) = \frac{e^{-\beta H(\Gamma)}}{\int d\Gamma e^{-\beta H(\Gamma)}} \tag{3.46}$$

However, one should bare in mind that MZT holds for arbitrary equilibrium distributions (e.g. micro-canonical).

The starting point consists in defining a Hilbert space spanned by the phase variables, and equipped with an bilinear operator:

$$(V_1, V_2) \equiv \int d\Gamma V_1(\Gamma)\; V_2(\Gamma) f_c(\Gamma). \tag{3.47}$$

Note that the inner product $(A^*(\Gamma), B(\Gamma, t))$ defines the equilibrium time correlation function between the phase variable B and the Langevin variable A:

$$C_{BA}(t) = \int d\Gamma B(\Gamma, t) A^*(\Gamma) f_c(\Gamma) = (B(\Gamma, t), A^*(\Gamma)) \tag{3.48}$$

Using the above inner product it is possible to define an operator $\hat{Q}$ which selects the "component" of any phase variable B which has no correlation with the Langevin variable A. To this end, we first introduce the *projection operator* $\hat{P}$, defined as follows

$$\hat{P} B(\Gamma, t) \equiv \frac{(B(\Gamma, t), A^*(\Gamma))}{(A(\Gamma), A^*(\Gamma))} A(\Gamma) \tag{3.49}$$

The operator $\hat{Q}$ is then defined as the complement of P. i.e.

$$\hat{Q} = 1 - \hat{P} \tag{3.50}$$

It is immediate to show that $\hat{Q}B$ is the component of B which is uncorrelated (i.e. random) with respect to A:

$$C_{\hat{Q}B,A}(t) = (\hat{Q} B(\Gamma, t), A^*(\Gamma)) = 0. \tag{3.51}$$

It is possible to show that both P and Q are Hermitian and nilpotent ($\hat{P}\hat{P} = \hat{P}$, $\hat{Q}\hat{Q} = \hat{Q}$) and that $\hat{P}\hat{Q} = \hat{Q}\hat{P} = 0$. Note that the propagator e^{iQL} generates a dynamics which is uncorrelated to the Langevin variable A. For this reason it is called the random propagator.

The rate of change of the Langevin variable A with time is given by:

$$\frac{d}{dt} A(t) = iLe^{iLt} A(0) = e^{iLt}(Q + P) iLA(0), \tag{3.52}$$

Now we recall the definition of the projection operator P and find:

$$e^{iLt}\; P\; iLA(0) = e^{iLt}\; \frac{(iLA, A^*)}{(A, A^*)} A(0) = \frac{(iLA(0), A^*)}{(A, A^*)} e^{iLt}\; A(0) \equiv i\Omega A(t), \tag{3.53}$$

Note that the quantity

$$i\Omega = \frac{(iLA, A^*)}{(A, A^*)} \tag{3.54}$$

has the dimension of a frequency and represents an equilibrium property of the system. Substituting this result into (3.52) we find:

$$\frac{d}{dt} A(t) = i\Omega A(t) + e^{iLt} QiLA(0), \tag{3.55}$$

To derive the generalized Langevin Eq., we now consider the Dyson representation (2.25) where the operator A is iQL and the operator B is iPL:

$$\begin{aligned} e^{iLt} = e^{(QiL+PiL)t} &= e^{QiLt} + \int_0^t d\tau e^{QIL\tau} PiLe^{(QiL+PiL)(t-\tau)} \qquad (3.56)\\ &= e^{QiLt} + \int_0^t d\tau \; e^{QiL\tau} P \; iL \; e^{iL(t-\tau)} \qquad (3.57)\\ &\overset{(t_1 \equiv t-\tau)}{=} e^{QiLt} + \int_0^t dt_1 \; e^{QiL(t-t_1)} P \; iL \; e^{iLt_1}. \qquad (3.58) \end{aligned}$$

Hence, we find

$$\frac{d}{dt}A(t) = i\Omega A(t) + \int_0^t dt_1 \; e^{QiL(t-t_1)} P \; iL \; e^{iLt_1} \; Q \; iL \; A(0) + e^{QiLt} QiLA(0). \tag{3.59}$$

The quantity

$$F(t) = e^{QiLt} QiLA(0) \tag{3.60}$$

can be identified as a random force, i.e. a force uncorrelated to the dynamics of the Langevin phase variable A. In fact, the time correlation function of $F(t)$ with A is

$$(F(t), A^*) = (e^{QiLt} QiLA, A^*) = (e^{QiLt} QQiLA, A^*) = (QF(t), A^*) = 0. \tag{3.61}$$

We stress the fact that the evolution from $F(0)$ to $F(t)$ is generated by the random propagator. Using the hermiticity of L and Q one finds:

$$P\,iL\,e^{iQLt}\,Q\,iL\,A(0) = -\frac{(F(t), (QiLA)^*)}{(A, A^*)} A(0) = -\frac{(F(t), (F(0))^*)}{(A, A^*)} A(0). \tag{3.62}$$

We now introduce a memory kernel for the random force, defined as:

$$K(t) \equiv \frac{(F(t), (F(0))^*)}{(A, A^*)} \tag{3.63}$$

So we finally find:

$$\begin{aligned} \frac{d}{dt}A(t) &= i\Omega\, A(t) - \int_0^t d\tau e^{iL(t-\tau)} K(\tau) A(0) + F(t) \qquad (3.64)\\ &= i\Omega\, A(t) - \int_0^t d\tau K(\tau) A(t-\tau) + F(t). \qquad (3.65) \end{aligned}$$

Hence, we have shown that the Langevin phase variable A obeys a generalized Langevin Eq. with a memory kernel.

3.2.1 Stochastic Path Integral

The Markov process defined by the Ito discretisation provides a simple procedure to integrate the overdamped Langevin equation: At each time-step one samples a stochastic displacement W_i from Eq.(3.41) and then sums the deterministic drift force $-\frac{D\Delta t}{k_B T}\frac{\partial}{\partial t}U[x(t)]$. This procedure generates stochastically trajectories in configuration space, which physically represent the Brownian trajectories of the system under the influence of the drifting force and of multiple and uncorrelated collisions with the environment. Our goal is to compute the probability density that the system initially prepared in configuration x_i at time t_0 reaches configuration x at time t_N

To this goal, we note that any given trajectory in configuration space generated by the Markov process (3.36), $x(t_0) \to x(t_1) \to x(t_2) \to \ldots \to x(t_{N-1})$, is uniquely by the time series of random displacement, $W(t_0) \to W(t_1) \to W(t_2) \to \ldots W(t_{N-1})$ which are sampled from the Gaussian distribution in the N integration steps. The probability of realising such a sequence of random displacements is given by the product of Gaussians:

$$\mathcal{P}[W(t_0) \to \ldots \to W_{N-1}] = \left(\frac{1}{\sqrt{4\pi D\Delta t}}\right)^{N-1} \prod_{k=0}^{N-1} e^{-\frac{W_k^2}{4D\Delta t}} \tag{3.66}$$

To compute the probability of the corresponding trajectory in configuration space we can use Eq. (3.36) to relate each W_k to a displacement in configuration space $x(t_{k+1}) - x(t_k)$:

$$\mathcal{P}[x(t_0) \to \ldots \to x_{N-1}] = \left(\frac{1}{\sqrt{4\pi D\Delta t}}\right)^{N-1} e^{-\frac{1}{4D\Delta t}\sum_{k=0}^{N-1}\left(x(t_{k+1})-x(t_k)+\frac{D\Delta t}{k_BT}\nabla U[x(t_k)]\right)^2} J\left[\frac{\partial W}{\partial x}\right], \tag{3.67}$$

In this equation $J\left[\frac{\partial W}{\partial x}\right]$ schematically represents the Jacobian of the transformation from the W to the x variables. It is straightforward to show (see e.g. [9]) that it is trivially a constant,

$$J\left[\frac{\partial W}{\partial x}\right] = \left[\frac{K_BT}{D\Delta t}\right]^{N-1}. \tag{3.68}$$

Finally, the conditional probability density $P(x_f, t|x_0)$ of observing the particle in some final location x_f after time $t = \Delta t N$ is obtained integrating Eq. (3.67) over all the undetermined intermediate positions:

$$P(x_f, t|x_0) = \mathcal{N} \int \left[\prod_{k=1}^{N-1} dx_k\right] e^{-\frac{1}{4D\Delta t}\sum_{k=0}^{N-1}\left(x(t_{k+1})-x(t_k)+\frac{D\Delta t}{k_BT}\nabla U[x(t_k)]\right)^2} \tag{3.69}$$

where $\mathcal{N}$ is an irrelevant normalisation constant. The multi-dimensional integration $\int\left[\prod_{k=1}^{N-1} dx_k\right]$ is usually represented with the symbol $\int \mathcal{D}x(\tau)$. It represents an integral over all paths the system can take to go from x_0 to x_f and for this reason it is is called the stochastic path integral. After restoring the continuous *notation* to represent the term at the exponent, we conditional probability is written as follows:

$$P(x_f, t|x_0) = \mathcal{N} \int \mathcal{D}x e^{-\frac{1}{4D}\int_{t_0}^{t} d\tau\left(\dot{x}(\tau)+\frac{D}{k_BT}\nabla U[x(\tau)]\right)^2} \tag{3.70}$$

where the functional in the exponent:

$$S_{OM}[x] = \frac{1}{4D}\int_{t_0}^{t} d\tau \left(\dot{x}(\tau) + \frac{D}{k_BT}\frac{\partial}{\partial t}U[x(\tau)]\right)^2 \tag{3.71}$$

is called the Onsager-Machlup functional. We emphasize however that the functional measure symbols and the continuos integrals and derivatives entering $S_{OM}[x]$ are just a convenient symbolic way to represent the multi-dimensional integral (3.69).

There is an equivalent representation of the Onsager-Machlup (3.71) which will be used in the following chapters to discuss the analogy and differences between stochastic diffusion and quantum delocalisation. To obtain such a representation we expand the square:

$$\begin{aligned} S_{OM}[x] &= \frac{1}{4D\Delta t}\sum_{k=0}^{N-1}\left(x(t_{k+1})-x(t_k)+\frac{D\Delta t}{k_BT}\nabla U[x(t_k)]\right)^2 \\ &= \Delta t\sum_{k=0}^{N-1}\left[\frac{1}{4D}\left(\frac{x(t_{k+1})-x(t_k)}{\Delta t}\right)^2+\frac{D}{4(k_BT)^2}|\nabla U[x(t_k)]|^2\right]+ \\ &\quad -\frac{1}{2k_BT}\sum_{k=0}^{N-1}(x(t_{k+1})-x(t_k))\nabla U[x(t_k)] \end{aligned} \tag{3.72}$$

Taking the limit $\Delta t \to 0, N \to \infty$ and using the fundamental theorem of Ito Calculus (3.43) the last term can be re-written as follows

$$\begin{aligned} &\lim_{\substack{\Delta t\to 0\\ N\to\infty}}\frac{1}{2k_BT}\sum_{k=0}^{N-1}(x(t_{k+1})-x(t_k))\nabla U[x(t_k)] \\ &= \frac{1}{2k_BT}\left(U(x_f)-U(x_i)\right)-\frac{1}{2M\gamma}\int_{t_0}^{t}ds\nabla^2U[x(s)] \end{aligned} \tag{3.73}$$

Combining these expressions and restoring the continuous notation, the stochastic path integral expressing the conditional probability $P(x_f,t|x_0)$ reads

$$P(x_f,t|x_0)=\mathcal{N}\ e^{-\frac{1}{2k_BT}\left(U(x_f)-U(x_i)\right)}\int\mathcal{D}x\ e^{-\frac{1}{4D}\int_{t_0}^{t}d\tau\left(\frac{1}{4D}\dot{x}^2+V_{eff}(x)\right)} \tag{3.74}$$

where

$$V_{eff}(x)=\frac{D}{4(k_BT)^2}\left(|\nabla U|^2-2k_BT\nabla^2U(x)\right) \tag{3.75}$$

The path integral representation of the conditional probability (3.70) can be obtained also in the underdamped limit (a detailed derivation can be found e.g. in the appendix of [22]). The result is

$$P(x_f,t|x_0)=\mathcal{N}\int_{x_0}^{x_f}\mathcal{D}xe^{-\frac{\beta}{4M\gamma}\int_{t_0}^{t}d\tau(M\ddot{x}(\tau)+M\gamma\dot{x}(\tau)+\nabla U[x(\tau)])^2} \tag{3.76}$$

3.2.2 Smoluchowki-Fokker-Planck Equation

In the previous subsections we have computed the conditional probability $P(x_f,t|x_0)$. Here we address the problem of determining the partial differential equation which defines the time evolution of this probability density.

First derivation

$$\begin{aligned} P(x,t+\Delta t) &= \int dy P(x,\Delta t+t|y,t)P(y,t) \\ &= \int dy \left(\frac{1}{4\pi D\Delta t}\right)^{3N/2} e^{-\frac{(x-y+\beta D\Delta t\nabla U(y))^2}{4D\Delta t}} P(y,t) \end{aligned} \tag{3.77}$$

Let us now expand the right-hand-side for small Δt.

$$\begin{aligned} \int dy \left(\frac{1}{4\pi D\Delta t}\right)^{3N/2} e^{-\frac{(x-y+\beta D\Delta t\nabla U(y))^2}{4D\Delta t}} P(y,t) &= \int dy \left(\frac{1}{4\pi D\Delta t}\right)^{3N/2} e^{-\frac{(x-y)^2}{4D\Delta t}} \\ & e^{-\Delta t V_{eff}(y)} e^{-\frac{\beta}{2}(U(x)-U(y))} P(y,t) \end{aligned} \tag{3.78}$$

where we have used the Fundamental Theorem of Ito Calculus and

$$V_{eff}(y) = \frac{D\beta^2}{4} \left(|\nabla U(y)|^2 - 2/\beta \nabla^2 U(y)\right) \tag{3.79}$$

Now notice that

$$\left(\frac{1}{4\pi D\Delta t}\right)^{3N/2} e^{-\frac{(x-y)^2}{4D\Delta t}} = \int \frac{d^N p}{(2\pi)^N} e^{ip(x-y)} e^{-D\Delta t q^2} \tag{3.80}$$

Thus, expanding to leading orders in a Taylor series for small Δt:

$$\begin{aligned} \left(\frac{1}{4\pi D\Delta t}\right)^{3N/2} e^{-\frac{(x-y)^2}{4D\Delta t}} &= \int \frac{d^N q}{(2\pi)^N} e^{iq(x-y)} (1 - D\Delta t q^2) \\ &\simeq \delta(x-y) + D\Delta t \nabla^2 \delta(x-y) \end{aligned} \tag{3.81}$$

Re-inserting into Eq. (3.78) we obtain an expression for $P(x,t+\Delta t)$ which is accurate to order Δt:

$$\begin{aligned} P(x,t+\Delta t) &= \int dy \left[\delta(x-y) + D\Delta t \nabla^2 \delta(x-y)\right]\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ e^{-\Delta t V_{eff}(y)} P(y,t) \\ &= P(x,t) - \Delta t V_{eff}(x) P(x,t) + D\Delta t \int dy\ \nabla^2\delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \end{aligned} \tag{3.82}$$

Let us now focus on the last term

$$\begin{aligned} & D\Delta t \int dy\ \nabla^2 \delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) = D\Delta t \nabla^2 \int dy\ \delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \\ & -\frac{\beta^2 D\Delta t}{4} \int dy \delta(x-y) (|\nabla U(y)|^2 - \frac{2}{\beta}\nabla^2 U(y))\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \\ & +\beta D\Delta t\ \nabla U(x) \int dy \nabla\delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))} P(y,t) \\ = \ & D\Delta t \left(\nabla^2 P(x,t) - V_{eff}(x)P(x,t) + \beta\nabla U(x)\nabla P(x,t) + \frac{\beta^2}{2}|\nabla U(x)|^2 P(x,t)\right) \end{aligned} \tag{3.83}$$

Notice that in the last line we have used:

$$\begin{aligned}
& D\Delta t\beta\nabla U(x)\int dy\nabla\delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \\
= \;& D\Delta t\beta\nabla U(x)\nabla\int dy\delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \\
& +D\Delta t\frac{\beta^2}{2}|\nabla U(x)|^2\int dy\delta(x-y)\ e^{-\frac{\beta}{2}(U(x)-U(y))}\ P(y,t) \\
= \;& D\Delta t\beta\nabla U(x)\nabla P(x,t)+D\Delta t\frac{\beta^2}{2}|\nabla U(x)|^2P(x,t)
\end{aligned} \tag{3.84}$$

Combining all terms we find the following expression to order Δt.

$$P(x,t+\Delta t)=P(x,t)+\Delta t\frac{\partial}{\partial t}P(x,t)=D\Delta t\left(\nabla^2+\beta\nabla U(x)\nabla+\beta\nabla^2U(x)\right)P(x,t) \tag{3.85}$$

which leads to the Smoluchowski-Fokker-Planck equation:

$$\frac{\partial}{\partial t}P(x,t)=D\left(\nabla^2+\beta\nabla U(x)\nabla+\beta\nabla^2U(x)\right)P(x,t). \tag{3.86}$$

Second derivation

An alternative derivation of the Smoluchowski-Fokker-Planck equation can be made by recalling that, in chapter 1 we introduced Feynman path integral formulation of quantum mechanics. In particular, we have shown that the path integral representation of the quantum propagator $K(x_f,t|x_0)=\langle x_f|e^{-\frac{i}{\hbar}Ht}|x_0\rangle$ takes the following form:

$$K(x_f,t|x_0)=\mathcal{N}\int\mathcal{D}x\ e^{\frac{i}{\hbar}\int_{t_0}^{t}d\tau\left(\frac{1}{2}m\dot{x}^2-U(x)\right)}. \tag{3.87}$$

Upon analytically continuing this expression to imaginary time $t=i\tau$, then we find an expression for the so-called Euclidean propagator

$$K_E(x_f,t|x_0)=\mathcal{N}\int\mathcal{D}x\ e^{-\int_{t_0}^{t}d\tau\left(\frac{1}{2}m\dot{x}^2+U(x)\right)}. \tag{3.88}$$

On the other hand, we recall that the path integral describing the diffusive dynamics of a classical particle obeying the overdamped Langevin equation reads:

$$P(x_f,t|x_0)=\mathcal{N}\ e^{-\frac{1}{2k_BT}\left(U(x_f)-U(x_i)\right)}\int\mathcal{D}x\ e^{-\frac{1}{4D}\int_{t_0}^{t}d\tau\left(\frac{1}{4D}\dot{x}^2+V_{eff}(x)\right)} \tag{3.89}$$

The similarity between (3.89) and (3.88) establishes a striking analogy between the consequences of delocalisation in quantum mechanics and of thernal fluctuations in sttochastic dynamics. It tells that, at the formal level, Brownian motion with diffusion coefficient D interacting with a potential $U(x)$ is dual to quantum propagation in imaginary time of a particle of effective mass $m_{eff}=2D$ interacting with an effective modified potential $V_{eff}(x)$.

The same analogy can be used to determine the equation obeyed by the Langevin probability density. Indeed, taking the time derivative of both sides of Eq. (3.74) and assuming $t - t_0 > 0$ and using the fact that K_E is the Green's function of the imaginary time Schrödinger equation:

$$\begin{aligned}
\frac{\partial}{\partial t}P(x,t|x_0) &= -e^{-\frac{1}{2k_BT}(U(x)-U(x_i))}\frac{\partial}{\partial t}K_E(x,t|x_i) \\
&= -e^{-\frac{1}{2k_BT}(U(x)-U(x_i))}(-D\nabla^2 + V_{eff}(x))\ K_E(x,t|x_i) \\
&= -e^{-\frac{1}{2k_BT}(U(x)-U(x_i))}(-D\nabla^2 + V_{eff}(x))e^{\frac{1}{2k_BT}(U(x)-U(x_i))}\ P(x,t|x_i) \\
&= -e^{-\frac{1}{2k_BT}U(x)}(-D\nabla^2 + V_{eff}(x))e^{\frac{1}{2k_BT}U(x)}\ P(x,t|x_i)
\end{aligned} \tag{3.90}$$

After taking explicitly the derivatives in the right-hand side and re-arranging the terms we arrive to the standard form of the Smoluchowski Fokker Planck (SFP) equation

$$\frac{\partial}{\partial t}P(x,t|x_0) = -\hat{H}_{FP}P(x,t|x_0) \tag{3.91}$$

where the $\hat{H}_{FP}$ is the Fokker-Plank operator,

$$\hat{H}_{FP} = -D\nabla(\nabla + \beta\nabla U(x)) \tag{3.92}$$

SFP is a Continuity Equation

Some comments on the SFP equation are in order. First, we note that this equation can be written in the form of a continuity equation, which expresses particle conservation in this theory:

$$\frac{\partial}{\partial t}P(x,t|x_0) = D\nabla \cdot J(x,t|x_0) \tag{3.93}$$

where $J(x,t|x_0) = (\nabla + \beta\nabla U(x))\ P(x,t|x_0)$ is the probability current distribution.

Next, we emphasise that the Fokker-Planck operator is not Hermitian. As a result, the dynamics generated by it is not unitary and therefore time-invariance is broken: Fluctuation and dissipation generate an irreversible dynamics. A system in contact with a heat-bath will eventually converge to thermal equilibrium. Indeed, it is straightforward to show that the Gibbs distribution

$$P_{eq}(x) = \frac{1}{Z}e^{-\beta U(x)} \tag{3.94}$$

is the only stationary solution of the SFP equation.

The SFP equation can be solved analytically only for a limited number of simple problems. In particular, in the absence of the external potential $U(x)$ (free random walk) the solution with boundary condition x_0 is given by:

$$P(x,t|x_0) = \frac{1}{\sqrt{4\pi Dt}}\ e^{-\frac{1}{4Dt}(x-x_0)^2} \tag{3.95}$$

This result can be used to study free Brownian diffusion (random walk). It is straightforward to show that the root-mean-square displacement as a function of time reads

$$\langle(x(t) - x_0)^2\rangle = 2Dt \tag{3.96}$$

This is famous Einstein's law for Brownian motion , which in 1905 provided the ultimate proof of the existence of molecular structure of matter.

We conclude with a technical remark. Mathematically, the conditional probability is a Green's function, so it obeys the distribution-valued equation

$$\left(\frac{\partial}{\partial t} + \hat{H}_{FP}\right) P(x,t|x_i) = \delta(t)\delta(x-x_i) \tag{3.97}$$

Hence, strictly speaking, equation (3.137) is valid only for $t > 0$.

3.3 Thermodynamics from Stochastic Dynamics

3.3.1 Microscopic Formulation of the first Law of Thermodynamics

Let us consider the most general case of a (passive) particle immersed in a thermal bath and possible interacting with external conservative and non conservative forces. We can model its dynamics through a standard Smoluchowski equation:

$$.\dot{\mathbf{x}} = -\frac{1}{M\gamma}\nabla U(\mathbf{x},t) + \sqrt{2D}\eta(t) \tag{3.98}$$

$$\langle \eta(t)\eta(0)\rangle = \delta(t). \tag{3.99}$$

We emphasize that Brownian particle can be prevented from equlibrating with the thermal bath by the time variations in the deterministic forces. Following Sekimoto [26, 27] the heat δQ that the particle exchanges with the thermal bath while moving over an infinitesimal distance $d\mathbf{x}$ during a time step dt from t to $t+dt$ is quantified as the energy the thermal bath transfer to the particle along this displacement due to friction $-M\gamma\dot{\mathbf{x}}$ and fluctuations $\sqrt{2k_BTM\gamma}\eta(t)$, i.e.

$$\delta Q(t) = \left(-M\gamma\dot{\mathbf{x}}(t) + \sqrt{2k_BTM\gamma}\eta(t)\right) d\mathbf{x}(t) \tag{3.100}$$

where the inner product is intended in the Stratonovich sense. We note that, with this definition, the heat is counted as positive if received by the particle and as negative when dumped into the environment via friction. Using the Langevin equation, we get to

$$\delta Q(t) = \nabla U(\mathbf{x}(t),t) \cdot d\mathbf{x}(t) \tag{3.101}$$

The change of the particle's "internal" energy $dU(t)$ over the same displacement $d\mathbf{x}(t)$ is given by the total differential

$$dU(t) = \nabla U(\mathbf{x}(t),t) \cdot d\mathbf{x}(t) + \frac{\partial U(x,t)}{\partial t} dt \tag{3.102}$$

Thus, if the potential does not vary over time, the only contribution comes from the change in the potential energy associated with the displacement $d\mathbf{x}$. On the other hand, even if the particle does not move within dt, its internal energy can still change due to variations of the potential landscape by an externally applied time-dependent protocol. Being due to forces imposed and controlled externally, the associated contribution $\frac{\partial U(x,t)}{\partial t}dt$ is interpreted as the rate of work performed by external forces on the particle and

$$\delta W(t) = \frac{\partial U(x,t)}{\partial t}dt \tag{3.103}$$

is the total work.

Combining the last three equation leads to the first law of thermodynamics:

$$dU = \delta Q + \delta W. \tag{3.104}$$

3.3.2 Microscopic Irreversibility, Entropy Production and Detailed Balance

The next step towards a thermodynamics characterization of the Brownian motion is to introduce an entropy change or entropy production associated with individual stochastic trajectories. To this end, we connect the concept of thermodynamic irreversibility (associated with entropy production) with a statistical mechanical definition, based on the probability of observing microscopic paths in opposite order.

We consider a set up in which system is described for simplicity by a single variable $\mathbf{x}$ in contact with the bath, and obeying an overdamped Langevin equation. In the most general case, we assume the particle is evolving according to an explicitly time-dependent potential, $U(\mathbf{x}, t)$.

To establish a relationship between thermodynamic and microscopic reversibility, we consider two stochastic trajectories, which are realized in the reverse time order. In the first process, the particle starts from configuration $\mathbf{x}(0) = \mathbf{x}_i$ and ultimately reaches configuration $\mathbf{x}(t) = \mathbf{x}_f$ along some path $\mathbf{x}(\tau)$, under the influence of the potential $U(\mathbf{x}, t)$. In the reversed process, the particle travels along the reversed path $\bar{\mathbf{x}}(\tau) = \mathbf{x}(t-\tau)$, i.e. starts from $\bar{\mathbf{x}}(0) = \mathbf{x}_f$ and ultimately reaches configuration $\mathbf{x}_i$, under the influence of the *time-reversed* potential $U(\mathbf{x}, t-\tau)$ (with $\tau \in [0,t]$).

Using the formulation of stochastic path integrals, it is possible to compute the probability density of observing the "forward" path $\mathbf{x}(\tau)$, relative to that of observing the reversed one $\bar{\mathbf{x}}(\tau)$. The former density is given by the probability of observing the initial configuration $P(\mathbf{x}_0, 0)$, multiplied by the negative exponent of the Onsager-Machlup action associated to the forward path, $S_{OM}[\mathbf{x}(\tau)]$:

$$P[\mathbf{x}(\tau)] = \mathcal{N}\ e^{-S_{OM}[\mathbf{x}]}, \tag{3.105}$$

where $\mathcal{N}$ is an irrelevent normalization factor. Conversely, the probability density for the reversed process is given by the probability density $\bar{P}(\mathbf{x}_f, 0)$ of observing $\mathbf{x}_f$ at time 0, multiplied the negative exponent of the Onsager Machlup functional evaluated along the reversed path with the time-inverted external potential, $\bar{S}_{OM}[\bar{x}]$:

$$\bar{P}[\bar{\mathbf{x}}(\tau)] = \mathcal{N}\ e^{-\bar{S}_{OM}[\bar{\mathbf{x}}]} \tag{3.106}$$

The ratio between these two probability densities is

$$R[\mathbf{x}, \bar{\mathbf{x}}] = \frac{\bar{P}(\mathbf{x}_f, 0)\ \bar{P}[\bar{\mathbf{x}}(t)|\mathbf{x}_f]}{P(\mathbf{x}_i, 0)\ P[\mathbf{x}(t)|\mathbf{x}_i]} \tag{3.107}$$

Let us adopt for convenience the Stratonovich Calculus (where the Fundamental Theorem of Calculus is the same of the standard Riemann's). Then, the probability of a given path $\mathbf{x}(t)$ satisfying $\mathbf{x}(\tau) = \mathbf{x}_f$ and $\mathbf{x}(0) = \mathbf{x}_i$. is

$$P[\mathbf{x}(\tau)] \propto \exp\left[-\int_0^\tau dt \left(\frac{(\dot{\mathbf{x}}(\tau) + \frac{D}{k_B T}\nabla U(\mathbf{x}(\tau),\tau))^2}{4D} - \frac{D}{2k_B T}\nabla^2 U(\mathbf{x}(\tau),\tau)\right)\right] \tag{3.108}$$

A similar structure is obtained for the probability density of the reversed path $\bar{\mathbf{x}}(\tau)$:

$$\bar{P}[\bar{\mathbf{x}}(\tau)] \quad \propto \quad \exp\left[-\int_0^\tau d\tau \left(\frac{(\dot{\bar{x}}(\tau) + \frac{D}{k_B T}\nabla \bar{U}(\bar{\mathbf{x}}(\tau),\tau))^2}{4D} - \frac{1}{2k_B T}\nabla^2 \bar{U}(\bar{\mathbf{x}}(\tau),\tau)\right)\right] \tag{3.109}$$

It is immediate to check that most of these terms appear identically in $P[\mathbf{x}(\tau)]$ and $\bar{P}[\bar{\mathbf{x}}(\tau)]$, thus cancel out in the ratio R. For example,

$$\begin{aligned} \int_0^\tau d\tau \frac{1}{4D}(\dot{\bar{x}}(\tau))^2 &= \int_0^\tau d\tau \frac{1}{4D}(\dot{\mathbf{x}}(t-\tau))^2 \\ &= -\frac{1}{4D}\int_t^0 d\tau (\dot{\mathbf{x}}(\tau))^2 = \frac{1}{4D}\int_0^t d\tau (\dot{\mathbf{x}}(\tau))^2 \end{aligned} \tag{3.110}$$

The only surviving factor arises from the cross product between $\dot{\mathbf{x}}$ and ∇U, which involves a single time derivative, thus adds up in the ratio, rather than canceling out. The result for R is then:

$$R[\mathbf{x}, \bar{\mathbf{x}}] = \frac{\bar{P}(\mathbf{x}_f, 0)\ \ \bar{P}[\bar{\mathbf{x}}(t)|\mathbf{x}_f]}{P(\mathbf{x}_i, 0)\ P[\mathbf{x}(t)|\mathbf{x}_i]} = \frac{\bar{P}(\mathbf{x}_f, 0)}{P(\mathbf{x}_i, 0)}\ e^{\frac{1}{k_B T}\int_0^\tau dt\ \dot{\mathbf{x}}(\tau)\cdot\nabla U(\mathbf{x},\tau)} \tag{3.111}$$

To make contact with the thermodynamic concept of reversibility, we now express this relationship in terms of entropy productions. We begin by noting that, using equation (3.101), the functional at the exponent in the right-hand-side,

$$\Delta S_{bath} = \frac{1}{T}\int_0^\tau d\tau\ \dot{\mathbf{x}}(\tau)\cdot\nabla U(\mathbf{x},\tau), \tag{3.112}$$

can be written as follows:

$$\frac{1}{T}\int_0^\tau dt\ \dot{\mathbf{x}}(t)\cdot\nabla U(\mathbf{x},t) = \frac{1}{T}\int_{\mathbf{x}_i}^{\mathbf{x}_f} d\mathbf{x}(t)\cdot\nabla U(\mathbf{x},t) = -\int_i^f \frac{\delta Q(t)}{T} = \Delta S \tag{3.113}$$

Thus, ΔS can be interpreted as the entropy production in the bath, resulting from the exchange of heat with the system.

In addition to this entropy production, there is also a change in entropy associated with the change in the microscopic state of the system. We can associate an entropy density to a configuration, i.e.

$$S_{sys}(\mathbf{x}, t) = -k_B \log P(\mathbf{x}, t), \tag{3.114}$$

Indeed, with such a definition, the ensemble entropy is given by the usual expression:

$$S_{sys} = -k_B \int d\mathbf{x} P(\mathbf{x}, t)\log P(\mathbf{x}, t). \tag{3.115}$$

In the specific case of a Langevin process, $P(\mathbf{x},t)$ is the solution of the Smoluckowski equation associated to the Langevin equation.

The change in system entropy along the trajectory is therefore given by

$$\Delta S_{sys} = -k_B \log P(\mathbf{x},\tau) + k_B \log P(\mathbf{x},0), \tag{3.116}$$

The total entropy change in the system and bath is then:

$$\Delta S_{tot} = \frac{1}{T}\int_0^t d\tau\ \dot{\mathbf{x}}(\tau)\cdot\nabla U(\mathbf{x},\tau) - k_B \log P(\mathbf{x},t) + k_B \log P(\mathbf{x},0), \tag{3.117}$$

We then use the fact that

$$\begin{aligned} R[\mathbf{x},\bar{\mathbf{x}}] &= \frac{\bar{P}(\mathbf{x}_f,0)}{P(\mathbf{x}_i,0)} e^{\Delta S_{bath}} = \frac{P(\mathbf{x}_f,t)}{P(\mathbf{x}_i,0)} e^{\Delta S_{bath}} \\ &= e^{-k_B \log P(\mathbf{x}_f,t) + k_B \log P(\mathbf{x}_i,0)} e^{\Delta S_{bath}} \\ &= e^{\Delta S_{bath} + \Delta S_{sys}} \end{aligned} \tag{3.118}$$

Therefore the ratio R is related to the total change in entropy in the forward process:

$$R = e^{\frac{\Delta S_{tot}}{k_B}} \tag{3.119}$$

Note that if the system is initially not in an equilibrium condition, i.e. $P(\mathbf{x},t) \neq \frac{1}{Z}e^{-\beta U(\mathbf{x}_i)}$ or is subject to a time-varying external force, the total change in entropy is non-null and the probability for the forward and backward processes are not identical, hence the irreversibility.

Conversely, if the system is in equilibrium and is not subject to time-dependent external drives, then $\int_0^\tau dt\ \dot{\mathbf{x}}(t)\cdot\nabla U(\mathbf{x},t) = U(\mathbf{x}_f) - U(\mathbf{x}_0)$, thus $\Delta S_{tot} = 0$. In this case, the total change in entropy vanishes and the probability of the forward and backward process coincide, i.e.

$$R[\mathbf{x},\bar{\mathbf{x}}] = 1 \tag{3.120}$$

In this case, the system is said to obey a microscopic reversibility condition.

Finally, we note after integrating over all paths with identical boundary conditions, from Eq. (3.111) we obtain the so-called *detailed balance condition*:

$$\rho_{eq}(\mathbf{x}_0)P(\mathbf{x}_f,\tau|\mathbf{x}_0) = \rho_{eq}(\mathbf{x}_f)P(\mathbf{x}_0,\tau|\mathbf{x}_f). \tag{3.121}$$

where $\rho_{eq}(\mathbf{x}) = 1/Ze^{-\beta U(\mathbf{x})}$ is the equilibrium distribution.

3.3.3 Fluctuation Theorems

The most useful results of thermodynamics are usually derived in the adiabatic limit, i.e. when the parameters which define the external potential are assumed to change infinitely slowly along some path g from an initial point A to a final point B in the parameter space. In this case, the total external work W performed on the system is equal to the Helmholtz free energy difference ΔF between the initial and final configurations:

$$W = \Delta F = F(B) - F(A) \tag{3.122}$$

Away from the adiabatic limit, i.e. when the parameters are switched along g at a finite rate, the work W will depend on the microscopic initial conditions of the system and reservoir, and will, on average, exceed the difference in Helmotz free energy:

$$\langle W\rangle = -\delta Q + \delta U > -T\Delta S + \Delta U = \Delta F, \tag{3.123}$$

where we have used the Clausius inequality, stating that for a general irreversible process $T\Delta S > \delta Q$. The difference $\langle W\rangle - \Delta F$ is the dissipated work, which is associated with the increase of entropy during an irreversible process.

The average in Eq. (3.123) is performed over an ensemble of different measurements of W, each of which is performed by first letting the system thermalize with the bath (at temperature T) and then perfoming a transition form A to B by varying the parameter.

In a relatively recent celebrated paper, Jarzynsky [18] derived the following *equality*:

$$\langle e^{-\beta W}\rangle = e^{-\beta\Delta F} \tag{3.124}$$

Shortly after, Crooks [19] came up with a closely related relationship

$$\frac{P_F(\Delta S)}{P_R(-\Delta S)} = e^{\left(\frac{\Delta S}{k_B}\right)}. \tag{3.125}$$

In this equation $P_F\,(\Delta S)$ is the probability of producing a certain amount of entropy in a non-equilibrium process in which the system is driven by some external force according to some schedule, starting from configuration sampled from some initial (e.g. equilibrium distribution). $P_R\,(-\Delta S)$ is the probability of generating an entropy $-\Delta S$ in the reversed process, in which the schedule of the external force has been inverted too. We stress that from the discussion in the previous section it follows that the entropy is an odd function of time.

The Crook's inequality is more general that Jarzynski's one. Indeed, the latter can be derived from the former. To see this we first note that Eq. (3.125) implies

$$\langle e^{-\frac{1}{k_B}\Delta S}\rangle_F = \int_{-\infty}^{\infty} d\Delta S\ P_F(\Delta S)\ e^{-\frac{1}{k_B}\Delta S} = \int_{-\infty}^{\infty} d\Delta S\ P_R(-\Delta S) = 1 \tag{3.126}$$

In this equation, the average is performed over many independent realization of the same process. Jarzynsky fluctuation theorem (3.124) is then readily obtained by using the relationship (3.123) and recalling that the Helmotz free energy F is a state-function, thus can be taken outside the average:

$$1 = \langle e^{-\frac{1}{k_B}(S_B-S_A)}\rangle_F = \langle e^{-\beta(W_{AB}-(F_B-F_A))}\rangle_F \tag{3.127}$$

$$\Rightarrow e^{-\beta(F_B-F_A)} = \langle e^{-\beta W_{AB}}\rangle_F \tag{3.128}$$

Both theorems (3.124) and (3.125) are quite remarkable. Indeed, they are among the few exact relationship in statistical mechanics which connect an equilibrium concept such as the Helmotz free-energy to *non-equilibrium* averages over different realizations of a dynamical process. We emphasize that, unlike most non-equilibrium relationships of statistical mechanics (such as Onsager's regression hypotheses), these results are not confined to small non-equilibrium perturbations, but instead apply to systems which are driven arbitrarily far from equilibrium. They state that it is possible to extract equilibrium information ΔF from the ensemble of non-equilibrium (finite-time) measurements.

The proof of Crooks theorem follows directly from the discussion about microscopic reversibility we presented in the previous section. In the following, we simply repeat the same argument with a notation closer to that of the original paper of Crooks. Namely, we assume for simplicity an overdamped Langevin dynamics and consider the instantaneous microstate of the system obeying Langevin dynamics to be specified by x and λ, where x represents a point in configuration space and $\lambda(t)$ is a controlled time-dependent parameter associated with an external force applied on the system. A particular realization of the transformation on the system induced by the external force is described by a path in configuration space, thus is specified by the pair of functions $(x(t), \lambda(t))$.

For convenience, we consider a symmetric time interval $t \in (-\tau, \tau)$. In addition, in analogy with the discussion in the previous section, we also introduce a time-reversed path $(\overline{x}(-t), \overline{\lambda}(-t))$. The overbar has been introduced to stress that quantities which are odd under a time-reversal, such as the momenta, have also changed sign. Recalling Eq.s (3.111) and (3.113) then

$$\frac{\mathcal{P}[x(+t)|\lambda(+t)]}{\mathcal{P}[(x(-t)|\lambda(-t)]} = e^{-\beta Q[x(+t),\lambda(+t)]} \tag{3.129}$$

Q is the heat, i.e. the amount of energy transferred from the bath into the system. As it is evident from Eq. (3.111) the heat is a functional of the path, and odd under a time reversal $Q[x(+t), \lambda(+t)] = -\bar{Q}[\bar{x}(-t), \bar{\lambda}(-t)]$. Recalling Eq.s (3.113) and (3.116) we find the expression for the total entropy production along a given path in the forward process:

$$\Delta S_F = k_B \log P(x, -\tau) - k_B \log P(x, \tau) - \beta Q[x(+t), \lambda(+t)] \tag{3.130}$$

Let us now compare the entropy production probability distribution of a process defined by a given path $x(t)$ with the entropy production distribution of the corresponding time-reversed process $\bar{x}(-t)$. To allow this comparison of forward and backward processes, we will rely on the fact that the entropy production is odd under a time reversal, $\Delta S_F = -\Delta S_R$. Considering Eq. (3.130) and recalling that the heat is an odd function, then the $\Delta S_F = -\Delta S_R$ condition is equivalent to requiring that the final distribution of the forward process, $P(x, \tau)$ is the same (after a time reversal) as the initial phase-space distribution of the reverse process, $P(\bar{x}, -\tau)$. This condition is satisfied when the system begins and ends in equilibrium or when the system begins and ends in the same time symmetric non-equilibrium steady state. Combining Eq.s (3.129) and (3.130) we find:

$$\frac{P(x, -\tau)\ \mathcal{P}[x(+t)|\lambda(+t)]}{P(\bar{x}, \tau)\mathcal{P}[(x(-t)|\lambda(-t)]} = e^{\frac{1}{k_B}\Delta S_F} \tag{3.131}$$

Let us now compute the probability of observing a particular value of entropy production in a forward process:

$$P_F(\Delta S) = \langle \delta(\Delta S - \Delta S_F) \rangle_F = \int dx_i \int dx_f \int_{x_i}^{x_f} \mathcal{D}x\ \delta(\Delta S - \Delta S_F)$$
$$\mathcal{P}[x(+t)|\lambda(+t)] P(x_i, -\tau) \tag{3.132}$$

Using Eq. (3.131) it is possible to re-write this in terms of an integral over the reversed process:

$$\begin{aligned} P_F(\Delta S) = e^{+\frac{1}{k_B}\Delta S} \int dx_i \int dx_f \int_{x_i}^{x_f} \mathcal{D}x\ \delta(\Delta S + \Delta S_R) \\ \mathcal{P}[\bar{x}(-t)|\bar{\lambda}(-t)] P(\bar{x}_i, +\tau) \\ = e^{+\frac{1}{k_B}\Delta S} \langle \delta(\Delta S + \Delta S_F) \rangle_R \\ = e^{+\frac{1}{k_B}\Delta S} P_R(-\Delta S) \end{aligned} \qquad (3.133)$$

thus prooving Crooks relationship (3.125).

To better appreciate the significance of Crooks theorem, let us now specialize on the case in which a system that is in equilibrium from time $-\infty$ to $-\tau$ is driven out of equilibrium by a change in the controlled parameter, λ till time τ. Then, it is allowed to relax, so that it once again reaches equilibrium at $t \to \infty$. For the forward process the system starts in the equilibrium ensemble specified by $\lambda(-\infty)$, and ends in the ensemble specified by $\lambda(\infty)$, which are Gibbs distributions:

$$P_{eq}(x|\lambda) = \frac{e^{-\beta H(x,\lambda)}}{\int dx e^{-\beta H(x,\lambda)}} = e^{\beta F(\lambda) - \beta H(x,\lambda)} \qquad (3.134)$$

After substituting these expression into Eq . (3.130) and using the first principle of thermodynamics we find

$$\Delta S_F = -\beta \Delta F + \beta W \qquad (3.135)$$

Since the change in free energy is path independent, the probability of entropy production is essentially the probability of exerting a given work W. thus, Crooks theorem implies

$$P_F(\beta W)/P_R(-\beta W) = e^{-\beta \Delta F + \beta W} \qquad (3.136)$$

From an average over the work performed in forward and backward processes it is possibile in principle to compute the free energy difference ΔF.

3.4 Stochastic Dynamics at Low Resolution

The dynamics of complex systems, such as biomolecules, glasses or liquids, is strongly influenced by the presence of a large number of metastable states, separated by all others by energy barriers. Barrier crossing transitions are rare events. Indeed, at least at the classical level, energy barriers can only be overcome when a peculiar sequence of thermal fluctuations occurs. These fluctuations must be such to drive system all the way to the top of the free energy barrier (the so-called transition state). Once the system has passed the transition state it is likely to diffuse to new meta-stable state, than to return to the original one.

Clearly, when the height of the energy is large compared to the typical energy of thermal fluctuations we expect that such thermally activated transitions should be rare. In particular, according to Kramers-Arrhenius theory [1] such a rate of

---

[1] see e.g. [14]

transitions is exponentially suppressed with the height of the free energy barrier . $k = \tau_0\ \exp[-\beta\Delta G]$ (here τ_0 is a typical time-scale associated with thermal relaxation within the initial state).

Therefore, the presence of energy barriers determines the decoupling between the time scales associated to the exploration of local metastable minima $\sim \tau_0$ and the slow time-scales associated to transitions between different basins and, ultimately, to global thermalisation, $\sim 1/k$. In this chapter, we exploit this gap in time scales to develop a rigorous effective low-energy description of the dynamics of complex classical systems which goes under the name of Markov State Model approach. From a RG inspired standpoint, we set a resolution frequency scale λ such that the transitions between local metastable states occur at a rate $k \ll \lambda$, while the fast relaxation dynamics within each state is UV physics, occurring at frequencies $k \gg \lambda$ and thus is parametrised in terms of few coefficients in the effective theory.

3.4.1 Hermitian formulation of Fokker-Planck Dynamics and Spectral Decomposition

The starting point to develop such an approach is to return to Fokker-Planck (FP) equation (3.91)

$$\frac{\partial}{\partial t}\ P(x,t) = -H_{\mathrm{FP}} P(x,t) \tag{3.137}$$

The non-hermitian character of H_{FP} implies that its left- and right- eigenvalues do not coincide.

As sketched already in Eq. (3.90) it is possible to define a (non-unitary) transformation acting on the FP operator and the probability distribution, which puts the FP equation an hermitian form:

$$\begin{cases} H_h = e^{\frac{\beta}{2}U(x)}\ H_{\mathrm{FP}}\ e^{-\frac{\beta}{2}U(x)} = -D\nabla^2 + \frac{D\beta^2}{4}\ \left[(\nabla U(x))^2 - \frac{2}{\beta}\,\nabla^2 U(x)\right] \\ \psi(x,t) = e^{-\frac{\beta}{2}U(x)} P(x,t) \end{cases} \tag{3.138}$$

Under this transformation, the FP equation turns into a Schrödinger equation in imaginary time:

$$\frac{\partial}{\partial t}\psi(x,t) = -H_h\ \psi(x,t). \tag{3.139}$$

In the following, the function $\psi(x,t)$ will be referred to as the *hermitian component* of the probability density $P(x,t)$. As a consequence, the Green's function of the FP operator, i.e. the conditional probability $P(x,t|x_0)$ which obeys

$$(\partial_t + H_{FP})P(x,t|x_0) = \delta(t)\delta(x-x_0) \tag{3.140}$$

is related to the Feynman imaginary time propagator of the fictitious quantum system defined by the hermitian operator H_h, i.e.

$$P(x,t|x_0) = e^{-\frac{\beta}{2}U(x)+\frac{\beta}{2}U(x_0)}\ \langle x|e^{-\hat{H}_h t}|x_0\rangle \equiv e^{-\frac{\beta}{2}U(x)+\frac{\beta}{2}U(x_0)}\ K(x,t|x_0) \tag{3.141}$$

where

$$(\partial_t + H_h)K(x,t|x_0) = \delta(t)\delta(x-x_0). \tag{3.142}$$

It is straightforward to show that H_h and $H_{\rm FP}$ have the same spectrum, which is non-negative definite and contains a null eigenvalue, $0 = \lambda_0 < \lambda_1 < \ldots$. We have already mentioned that the right zero-mode $R_0(x)$ is the Gibbs distribution:

$$H_{\rm FP} R_0(x) = 0, \qquad R_0(x) = \frac{1}{Z}\, e^{-\beta U(x)}, \tag{3.143}$$

where $Z = \int dx\; e^{-\beta U(x)}$ is the system's canonical partition function.

The hermitian components of the left- and right- eigenstates of H_{FP} are in one-to-one correspondence to the eigenstates $\phi_i(x)$ of the hermitian operator H_h. In other words, defining

$$\hat{H}_h \phi_n(x) = \lambda_n \phi_n(x) \tag{3.144}$$

with the normalization

$$\int dx \phi_n^*(x) \phi_m(x) = \delta_{nm} \tag{3.145}$$

Then, the functions

$$\begin{aligned} R_n(x) &= \frac{1}{\sqrt{Z}} e^{-\frac{\beta}{2}U(x)}\; \phi_n(x) && (3.146)\\ L_n(x) &= \sqrt{Z}\; e^{\frac{\beta}{2}U(x)}\; \phi_n(x)\,. && (3.147)\end{aligned}$$

are eigenstates of $\hat{H}_{FP}$ and $\hat{H}^\dagger_{FP}$, respectively, obeying the normalization relationship:

$$\int dx R_n(x) L_m(x) = \delta_{nm}. \tag{3.148}$$

We note that the normalization factors in Eqs. (3.146) and (3.147) have been introduced to ensure that the right zero-mode R_0 has the correct normalization. Namely, since $\phi_0 = \frac{1}{\sqrt{Z}} \exp(-\beta/2 U(x))$, then the definition (3.146) leads to

$$R_0(x) = \frac{1}{Z} e^{-\beta U(x)}. \tag{3.149}$$

Correspondingly, the left zero mode is just the identity

$$L_0(x) = 1. \tag{3.150}$$

Finally, we note that the conditional probability density $P(x, t|x_0)$ (i.e. the Green's function of the FP operator) can be expanded as a series of right and left eigenfunctions of $H_{\rm FP}$ or, equivalently, eigenfunctions of H_h:

$$\begin{aligned} P(x,t|x_0) &= e^{-\frac{\beta}{2}U(x)+\frac{\beta}{2}U(x_0)}\; \langle x| e^{-\hat{H}_h t} |x_0\rangle && (3.151)\\ &= e^{-\frac{\beta}{2}U(x)+\frac{\beta}{2}U(x_0)}\; \sum_n \phi_n(x)\; \phi_n(x_0) e^{-\lambda_n t} && (3.152)\\ &= \sum_n R_n(x) L_n(x_0) e^{-\lambda_n t} && (3.153)\end{aligned}$$

where we have made use of the fact that the eigenstates $\phi_n(x)$ are real-valued functions. This expression shows that the lowest non-vanishing eigenvalues of the FP operator is associated to the inverse time required to attain global thermalization.

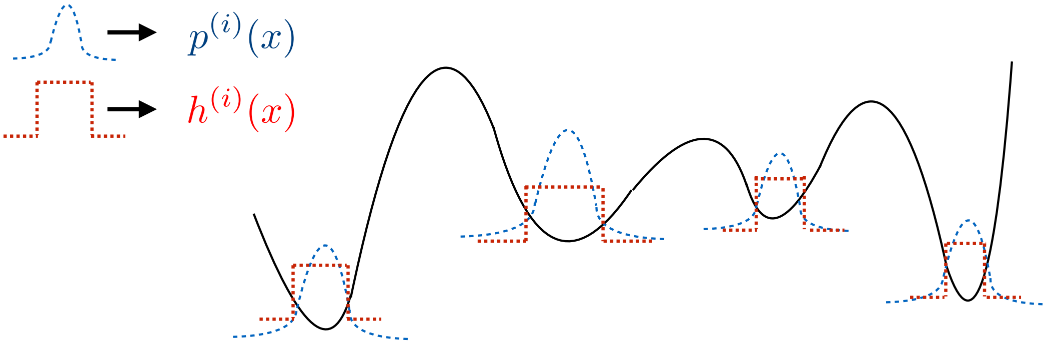

Figure 3.1: Graphical representation of the definition of microstate distributions $p^{(i)}(x)$ and characteristic functions $h^{(i)}(x)$. The continuous line denotes the energy landscape.

3.4.2 FP Dynamics in Rugged Energy Landscapes

Let us now specialize on systems characterized by a rugged energy landscape, i.e. by a potential energy $U(x)$ with many local minima, separated by barriers. In this case the spectrum of the FP operator will be gapped, with a set of high-frequency modes associated to local relaxation within the states and a few low-frequency modes associated to hopping between states and global relaxation (Figure 3.2). Our goal is to construct a low-energy ET which only deals with dynamics at time intervals much longer than the inverse of the lowest eigenvalue above the gap, Λ. Thus, Λ represents a typical value for the hard-frequency cut-off scale of the effective theory, i.e. $t \gg dt \equiv 1/\Lambda$.

To explicitly construct such an effective theory, starting from the microscopic Fokker-Planck dynamics, we quote an important (yet very recent) result of mathematical physics[2]: In the spectrum of the FP operator is gapped with N eigenvalues below the gap, there exist exactly N linear combinations of the corresponding right-eigenstates,

$$p^{(i)}(x) = \sum_{j=1}^{N} C_{ij} \; R_j(x) \tag{3.154}$$

which simultaneously satisfy the following three properties:

1. Non-negativity, i.e. $p^{(i)}(x) \geq 0$
2. Disjointness, i.e. $p^{(i)}(x) p^{(j)}(x) = 0$ for $i \neq j$ and $\forall x$
3. Local Gibbseanity, i.e.

$$p^{(i)}(x) = \frac{1}{z_i} e^{-\beta U(x)} h^{(i)}(x) \qquad \text{with} \qquad z_i = \int dx \; h^{(i)}(x) \; e^{-\beta U(x)}, \tag{3.155}$$

   where $h^{(i)}(x)$ is negligible everywhere and equal to 1 in the vicinity of one and only one of the local meta-stable states. $p^{(i)}(x)$ is then interpreted as

---

[2] The interested reader can find a discussion in [10]

the probability distribution associated to the $i-$th so-called Markov state and $h^{(i)}(x)$ denotes the corresponding characteristic function (see Fig. 6.1).

It is also relevant to consider the following linear combinations of the left eigenstates $L_i(x)$, with $1 \leq i \leq N$:

$$a^{(i)}(x) = \frac{1}{Z}\sum_{j=1}^{N} z_i C_{ij}\ L_j(x). \tag{3.156}$$

where the convenience of the $1/Z$ normalization and the z_i factor will become clear below. The interest in these linear combinations resides in the fact that the $a^{(i)}(x)$ functions are proportional to the characteristic functions of the microstates, thus are approximatively 1 where $p^{(i)}(x)$ is locally Gibbsean, and are negligible elsewhere. To see this, it is sufficient to isolate the hermitian component of the $a^{(i)}(x)$ distributions:

$$\begin{aligned}
a^{(i)}(x) &= \frac{1}{Z}\sum_{j=1}^{N} C_{ij} z_i\ \sqrt{Z}\ \phi_j(x)\ e^{\frac{\beta}{2}U(x)} \\
&= \frac{1}{Z}\sum_{j=1}^{N} z_i C_{ij}\ \sqrt{Z}\ \frac{\sqrt{Z}\ e^{-\frac{\beta}{2}U(x)}}{\sqrt{Z}\ e^{-\frac{\beta}{2}U(x)}}\phi_j(x)\ e^{\frac{\beta}{2}U(x)} \\
&= e^{\beta U(x)}\sum_{j=1}^{N} z_i C_{ij}\ R_j(x) \\
&= z_i p^{(i)}(x)\ e^{\beta U(x)} \simeq h^{(i)}(x)
\end{aligned} \tag{3.157}$$

In the hermitian formalism, the left- and right- state distributions are replaced by a single distribution $\pi^{(i)}(x)$:

$$\pi^{(i)}(x) \equiv \frac{h_i(x)}{\sqrt{z_i}}\ e^{-\frac{\beta}{2}U(x)} = \begin{cases} \frac{1}{\sqrt{z_i}} e^{-\frac{\beta}{2}U(x)}\ a^{(i)}(x) \\ \sqrt{z_i} e^{+\frac{\beta}{2}U(x)}\ p^{(i)}(x) \end{cases} . \tag{3.158}$$

Once the dynamics has been coarse-grained at the level of micro-states, the configuration space representation becomes redundant. Our goal is to derive an equation for the time evolution of the probability to observe the system in the $i-$th microstate, i.e.

$$n_i(t) = \int dx h_i(x) P(x,t) = \int dx\ a_i(x)\ P(x,t) \tag{3.159}$$

where $P(x,t)$ is the solution of the FP equation, given some initial distribution. We note that in the long time limit we obtain the expected equilibrium result:

$$n_i(t) \to \frac{z_i}{Z}. \tag{3.160}$$

To derive the equations for $n_i(t)$ we consider these densities at time $t + \Delta t$:

$$n_i(t+\Delta t) = \int dx h_i(x) P(x, t+\Delta t) \tag{3.161}$$

We now exploit the Markovian character of Langevin dynamics, which implies

$$n_i(t+\Delta t) = \int dx h_i(x) \int dy P(x,\Delta t|y) P(y,t) \tag{3.162}$$

We now use the fact that, $P(y,t)$ can be approximatively expressed as

$$P(y,t) = \sum_{n=1}^{N} R_n(y) c_n \; e^{-\lambda_n t}, \qquad c_n = \int dx L_n(x) P(x,t) \tag{3.163}$$

since the contribution from all terms above the gap is exponentially suppressed. In turn, the $R_n(y)$ can be expressed by the locally Gibbsean distributions:

$$P(y,t) = \sum_{n=1}^{N} \sum_{j} C_{nj} p_j(y) c_n \; e^{-\lambda_n t} = \sum_{n=1}^{N} \sum_{j} C_{nj} h_j(y) \frac{1}{z_j} e^{-\beta U(y)} c_n \; e^{-\lambda_n t} \tag{3.164}$$

Then

$$n_i(t+\Delta t) = \int dx h_i(x) \int dy P(x,\Delta t|y) \sum_{n=1}^{N} \sum_{j} C_{nj} h_j(y) \frac{1}{z_j} e^{-\beta U(y)} c_n \; e^{-\lambda_n t} \tag{3.165}$$

This equation is conveniently rewritten as

$$n_i(t+\Delta t) = \sum_{j} T_{ij}^{(\Delta t)} \sum_{n=1}^{N} C_{nj} c_n \; e^{-\lambda_n t} \tag{3.166}$$

where

$$T_{ij}^{(\Delta t)} = \int dx h_i(x) \int dy P(x,\Delta t|y) \frac{1}{z_j} e^{-\beta U(y)} \; h_j(y) \tag{3.167}$$

is called the transition matrix and measures the probability to perform a transition from state j to state i in lag-time Δt. Finally, we show that the summation

$$\sum_{n=1}^{N} C_{nj} c_n \; e^{-\lambda_n t} \tag{3.168}$$

is in fact just an expression for $n_j(t)$. Indeed

$$\begin{aligned} n_j(t) &= \int dx h_j(x) P(x,t) = \int dx h_j(x) \sum_n c_n e^{-\lambda_n t} R_n(x) \\ &= \int dx h_j(x) \sum_n c_n e^{-\lambda_n t} C_{nj} h_j(x) \frac{e^{-\beta U(x)}}{z_j} \\ &= \sum_n c_n e^{-\lambda_n t} C_{nj} \end{aligned} \tag{3.169}$$

In conclusion, we arrived to the so-called Master Equation, which describes the FP evolution in the space of microstates:

$$n_i(t+\Delta t) = \sum_{j} T_{ij}^{(\Delta t)} n_j(t) \tag{3.170}$$

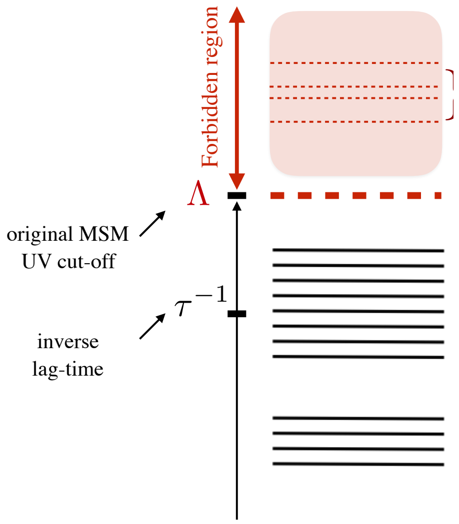

Figure 3.2: The different frequencies scales must be ordered according to a specific hierarchy for the MSM to represent an effective theory of MD. The forbidden region is inaccessible to the effective theory.

We stress that conservation of probability implies that the transition matrix is in fact left-stochastic, i.e. it obeys the sum-rule:

$$\sum_{j=1}^{N} T_{ji}^{(\Delta t)} = 1 \tag{3.171}$$

Since τ is an infinitesimal time scale it is sometimes convenient to express the Master Equation (3.170) in differential form. This is done by introducing the so-called *rate matrix* K_{ij}:

$$T_{ij}^{(\Delta t)} = \delta_{ij} + \Delta t K_{ij} \tag{3.172}$$

leading to the so-called *continuous time Master Equation*:

$$\dot{n}_i(t) \quad = \quad \sum_{j=1}^{N} K_{ij} n_j(t) \tag{3.173}$$

The left-stochasticity of T_{ij} is reflected into K_{ij} as follows

$$K_{ij} = \begin{cases} k_{ij} \geq 0 & \text{if } i \neq j \\ -\sum_i k_{ij} & \text{if } i = j \end{cases} \tag{3.174}$$

We emphasize again that, in the RG perspective, time derivative in the left-hand-side of Eq. (3.173) does not represent a real derivative, but rather a finite increment. Indeed, dt cannot be set smaller than the inverse frequency cut-off scale Λ. From this discussion it follows that the lag-time τ of discrete-time MSM must obey $\tau > dt > 1/\Lambda$. The present effective Markovian coarse grained description of the stochastic dynamics is usually referred to as a Markov State Model (MSM).

Chapter 4

Statistical Mechanics of Rare Events

Complex systems often exhibit extremely slow structural re-arrangements. Example of these process include structural relaxation in glasses, but also conformational transitions in biomolecules, which can occur at time-scales which are many orders of magnitude longer than those associated to local conformational changes. Being able to predict the time-evolution of these systems by explicitly accounting for the motion of all its degrees of freedom typically provides an extremely challenging task. For example, to study the dynamics of a protein, one typically needs to solve $\sim 10^6 - 10^8$ coupled equations of motions.

But even if the we have enough computer power to be able to solve all the relevant coupled differential equations for very long time intervals, we are still left with the problem of physically interpreting the results of our numerical simulations. As we shall see, this step involves tricky conceptual questions. For example, suppose we seek to use computer simulations to predict the native state of a protein. How can we decide when we have found the native state? One possibility would be run a very long molecular dynamics simulation and monitor the potential energy for each configuration. Then native state may be identified as the configuration of minimum energy.

This approach, however, would not be entirely accurate and conceptually correct. In fact, the native state of a protein does not consist of a unique configuration, because of thermal fluctuations visits a finite portion of phase space in the structural neighborhood of the minimum energy configuration. In other words, the native state of a protein is a thermodynamically stable state and has a finite entropy S and represents the minimum of free energy $G = U - TS$. By contrast, the entropy of unfolded state is huge, as there are many coil configurations which can be reached by denaturing the molecule. How can one decide when two configurations belong to the same meta-stable state or to different states?

Related questions arise in the study of kinetics: integrating the equations of motion starting from different initial conditions leads to a set of trajectories which will visit different points of phase space. In general, not all these trajectories will provide information to the same physical process. For example, the small fluctuations of a protein around the minimum energy configuration, i.e. the trajectories confined within the native state, provide different information from thermally activated trajectories which connect the native state and denatured state. Let us now consider two trajectories which connect native and unfolded state. Even in this case, it is not clear that they represent equivalent realization of the same structural reaction, as they may represent kinetically distinct reactive channels, i.e. different independent ways in which the same reaction can take place.

This example illustrates a number of questions which naturally emerge when studying dynamics and kinetics of complex systems: How to quantitatively distinguish trajectories which represent different microscopical realizations of the same physical process from those which represent "truly" different processes? How to define and classify the phases of a complex systtes, as a function of e.g. temperature and urea concentration? How to predict the rate at which reactions occur and the relative kinetic weights of different reaction channels? This chapter is devoted to the rigorous theoretical framework to describe rare events in classical statistical mechanical systems. For sake of definiteness we shall focus on system obeying (overdamped) Langevin dyanamics. However, most of the results apply in general Markovian systems obeying microscopic reversibility.

4.1 First-Passage Time Distribution

A key concept which is required to investigate the kinetics of thermally activated transitions between two meta-stable states is the distribution of first-arrival times into the product, stating from configurations in the reactant. In a more mathematical language, we are interested in the distribution of first passage times across a hyper-surface ∂W which is defined to wrap the product.

To this end, we define a new Green's function, which has the following interpretation in terms of conditional probability densities:

$$P^*_W(x_f, t_f|x_i) \quad = \quad \text{Prob.(going from } x_i \text{ to } x_f \text{ in time } t, \text{without ever touching } \partial W) \tag{4.1}$$

where ∂W denotes the boundary of the region W.

In the framework of the overdamped Langevin dynamics, this new Green's functiocan be given the following path integral representation:

$$P^*_W(x,t|x_i) \quad = \quad e^{-\frac{\beta}{2}(U(x)-U(x_i))} \int_{x(t_i)=x_i}^{x(t)=x} \mathcal{D}q e^{-\int_0^t d\tau \ \beta \ \frac{\dot{x}^2}{4D}+V_{eff}[x]+\Omega_W[x]} \tag{4.2}$$

where $\Omega_W(x)$ is a characteristic function which vanishes outside the region W and is infinite inside.

Imposing that $\Omega_W(x)\, P^*_W(x,t|x_i) = 0$, implies that $P(x,t|x_0)$ and $P^*_W(x,t|x_0)$ satisfy the same FP equation, i.e.

$$\frac{\partial}{\partial t} P^*_W(x,t|x_i) - D\nabla \left[\nabla P^*_W(x,t|x_i) + \beta \nabla U(x) \ P^*_W(x,t|x_i)\right] = \delta(t)\delta(x-x_i). \tag{4.3}$$

On the other hand, P^*_W and P are in general different distributions, since P^*_W must satisfy the extra requirement to vanish at the boundary of the region W.

The integral over the entire space of P^*_W yields the survival probability, i.e. the probability of not having entered yet the region delimited by ∂W, at time t.

$$S(\partial W, t|x_i) = \int dx\ P^*_W(x,t|x_i). \tag{4.4}$$

Clearly, $1-S(\partial W, t|x_i)$ represents the probability of having crossed ∂W at some time before t, starting from an initial point x_i at a finite distance from ∂W. This observation can be used to compute the probability distribution $F(\partial W, t|x_i)$ of crossing ∂W for the first time a time t. Indeed,

$$1 - S(\partial W, t|x_i) = \int_0^t dt' F(\partial W, t|x_i), \tag{4.5}$$

thus

$$\begin{aligned} F(\partial W, t; x_i) &\equiv \frac{\partial}{\partial t}(1 - S(\partial W, t|x_i)) = -\frac{\partial}{\partial t} S(\partial W, t|x_i) \\ &= -\int dx\ \frac{\partial}{\partial t} P^*_W(x,t|x_i) \\ &= -D \int dx\ \nabla(\nabla P^*_W(x,t|x_i) + \beta \nabla U(x)\ P^*_W(x,t|x_i)) \\ &= D \int_{x\in\partial W} \mathbf{d}\sigma \cdot (\nabla P^*_W(x,t|x_i) + \beta\nabla U(x) P^*_W(x,t|x_i)) \\ &= \int_{x\in\partial W} \mathbf{d}\sigma \cdot \mathbf{J}^*_W(x,t|x_i) \end{aligned} \tag{4.6}$$

where

$$J^*_W(x,t|x_i) = D(\nabla P^*_W(x,t|x_i) + \beta\nabla U(x) P^*_W(x,t|x_i)) \tag{4.7}$$

is the probability current associated to the modified propagator $P^*_W(x,t|x_i)$ and $\mathbf{d}\sigma$ represents an infinitesimal surface element of the boundary of W, oriented outwards. This equation expresses the first passage time distribution as a integral over the flux through the surface ∂W of the conditional probability, with absorbing boundary condition. Notice that this Eq. yields the well-known expression for the mean-first-passage time across ∂W [15]:

$$\langle t \rangle_{\partial W} = \int_0^\infty dt S(\partial W, t|x_i) \tag{4.8}$$

4.2 Transition Path Theory

Transition Path Theory (TPT) provides a rigorous theoretical framework to characterize classical thermally activated transitions between two given metastable state, here denoted with reactant state R and product state P, embedded in the system's configuration space Ω (see Fig. 1). While it is in principle possible to formulate this theory for any dynamics which obeys the fundamental requests of e markovianity (possibly in phase-space) and microscopical reversibility (which implies ergodicity with respect to some equilibrium distribution), for sake of simplicity, we shall here specialize on systems obeying the underdamped Langevin equation.

The first step in TPT consists in providing a rigorous mathematical definition of transition paths (e.g. represented by the solid lines in Fig. 1) To this goal it is convenient to introduce the following time-dependent functions:

$$\begin{cases} t^{+}(t) = & \text{smallest time } t' \geq t \text{ such that } x(t') \in R \cup P \\ t^{-}(t) = & \text{largest time } t' \leq t \text{ such that } x(t') \in R \cup P \end{cases} \tag{4.9}$$

The function $t^{+}(t)$ returns the first instant after time t at which the trajectory $x(t)$ is in either the reactant or the product state. Similarly, $t^{-}(t)$ returns the last time before t at which the trajectory was in either of these states (see Fig. 1).

The ergodic trajectory is said to be *reactive* at some specific instant $\bar{t}$ if the following three conditions are simultaneously satisfied:

1. The instantaneous configuration at time $\bar{t}$ lies outside both the reactant and product states (i.e. in the so-called transition region): $\Omega_T \equiv \Omega/(R \cup P)$.

2. The trajectory $x(t)$ last entered the transition region by leaving the reactant state: $x[t^{-}(\bar{t})] \in R$

3. The trajectory $x(t)$ will leave the reactive region by entering the product state: $x[t^{+}(\bar{t})] \in P$.

A *transition path* is a continuous piece of the ergodic trajectory $x(t)$ which is entirely reactive. The *transition path ensemble* is formed by the set of all transition paths which are contained in the ergodic trajectory. Finally, the rate of a reaction can be defined as the average number of reactive pathways N_R observed per unit time interval T:

$$k_{RP} = \lim_{T \to \infty} \frac{N_R(T)}{T}. \tag{4.10}$$

4.2.1 Committor functions

Any generic point x of configuration space can be visited by both reactive and non-reactive parts of the ergodic trajectory. The probability $P_{\text{react.}}(x)$ that a given segment of the ergodic trajectory which visits point x is reactive can be conveniently expressed by introducing the so-called (forward) *committor function* $q^{+}(x)$, which measures the probability that a trajectory initiated at some point $x \in \Omega_T$ will enter the product state P before returning to the reactant R. Thus, by definition, $q^{+}(x)$ obeys the boundary condition:

$$\begin{cases} q^{+}(x) = 0, & x \in \partial R \\ q^{+}(x) = 1, & x \in \partial P, \end{cases} \tag{4.11}$$

where ∂R and ∂P denote the boundaries of the reactant and product regions, respectively. Similarly, one can define the backward committor $q^{-}(x)$ as the probability that a trajectory passing through point x will next return to the reactant R before landing into the product P. By microscopic time-reversibility, $q^{-}(x)$ can be expressed in terms of the forward committor: $q^{-}(x) = 1 - q^{+}(x)$.

We now show that $q^+(x)$ (and thus $q^-(x)$) obeys the so-called stationary backward Kolmogorov equation (see e.g. [**?**]):

$$D\nabla^2 q^{+/-}(x) - \beta\nabla U(x)\cdot\nabla q^{+/-}(x) = 0. \tag{4.12}$$

To prove this statement, we consider the Green's function $P^*_{\partial W}(x,t|x_i)$ introduced in (4.1), where the absorbing boundary conditions are defined on border of the region defined by the reactant

$$\partial W \equiv \partial R \cup \partial P. \tag{4.13}$$

By definition, the committor probability is related to the flux of the associated probability current $J^*_{\partial W}$ through the border of the product. More precisely,

$$\begin{aligned} q(x) &= \int_0^\infty dt \int_{\partial P} \mathbf{d}\sigma' \ \cdot \mathbf{J}^*_{\partial W}(x',t|x) \\ &= D\int_0^\infty dt \int_{\partial P} \mathbf{d}\sigma' \ \cdot (\nabla' + \beta\nabla U(x'))P^*_{\partial W}(x',t|x) \end{aligned} \tag{4.14}$$

Indeed, this equation represents the probability of arriving at the border of the product state at any time, *without having touched the border of the reactant state.*

Let us now recall that all FP propagators also obey the backward FP equation:

$$\begin{aligned} -\partial_{t_0}P(x_f,t|x,t_0) &= D(\nabla_x^2 - \beta\nabla U(x)\nabla_x)P(x_f,t|x,t_0) \\ &\equiv \hat{H}_b \ P(x_f,t|x,t_0) \end{aligned} \tag{4.15}$$

Assuming homogeneity (i.e. $P(x,t|x_0,t_0) = P(x,t-t_0|x_0,0)$), the equation (4.15) is rewritten as

$$\begin{aligned} \partial_t P(x_f,t|x,t_0) &= D(\nabla_x^2 - \beta\nabla U(x)\nabla_x)P(x_f,t|x,t_0) \\ &\equiv \hat{H}_b \ P(x_f,t|x,t_0) \end{aligned} \tag{4.16}$$

Finally applying the backward Fokker-Planck operator to both sides of Eq. (4.14) and using (4.16) we obtain:

$$\begin{aligned} \hat{H}_b \ q(x) &= \hat{H}_b \int_0^\infty dt \int_{\partial P} \mathbf{d}\sigma' \ \cdot \mathbf{J}^*_{\partial W}(x',t|x) \\ &= D\int_0^\infty dt \int_{\partial P} \mathbf{d}\sigma' \ \cdot (\nabla' + \beta\nabla U(x')) \ \hat{H}_b \ P^*_{\partial W}(x',t|x) \\ &= D\int_0^\infty dt \ \frac{\partial}{\partial t}\int_{\partial P} \mathbf{d}\sigma' \ \cdot (\nabla' + \beta\nabla U(x'))P^*_{\partial W}(x',t|x) \\ &= F(t=\infty) - F(t=0) = 0 \end{aligned} \tag{4.17}$$

The last line denotes the flux of the probability current through the hyper surface at very large times minus that at the initial time. The former flux vanishes because, in an ergodic system, at infinite time the system must have visited the product. The latter vanishes assuming the initial condition is at a finite distance from ∂W. Finally, the $P_{\text{react.}}(x)$ probability is written as $P_{\text{react.}}(x) = q^+(x)\ q^-(x)$, thus

$$P_{\text{react.}}(x) = q^+(x)\left(1 - q^+(x)\right). \tag{4.18}$$

Iso-committor hypersurfaces – defined as the set of points for which $q^+(x) = \bar{q}$, with $\bar{q} \in (0,1)$ – provide a particularly useful foliation of the configuration space. To see this, let us first consider a generic hypersurface ∂S (not in general an iso-committor one) and compute the probability $p_{\partial S}(x)$ that a *reactive trajectory* crosses it at some point $x \in \partial S$. This can be written as the probability density that any piece of the ergodic trajectory (reactive or not) visits the point x (i.e. the Gibbs distribution), times the probability $P_{\text{react.}}(x)$ for such piece of trajectory to be reactive:

$$p_{\partial S}(x) \quad = \quad \frac{1}{Z_{\partial S}}\ e^{-\beta U(x)}\ P_{\text{react.}}(x), \tag{4.19}$$

where the normalization factor is defined by the surface integral $Z_{\partial S} = \int_{\partial S} d\sigma q^+(x)\ (1- q^+(x))\ e^{-\beta U(x)}$.

Let us now specialize on case in which ∂S is some iso-committor hyper-surface defined by $q^+(x) = \bar{q}$. Then Eq. (4.18) implies

$$p_{\partial S_{\bar{q}}}(x) = \frac{1}{\mathcal{Z}_{\partial S_{\bar{q}}}}\ e^{-\beta U(x)} \qquad \mathcal{Z}_{\partial S_{\bar{q}}} = \int_{\partial S_{\bar{q}}} d\sigma\ e^{-\beta U(x)}. \tag{4.20}$$

Therefore, the probability that a reactive trajectory crosses an iso-committor surface at some point x coincides with the equilibrium probability restricted to this surface. We emphasize that this result establishes a highly non-trivial relationship between probability densities defined in equilibrium and dynamical conditions, and identifies the iso-commitor function as the ideal reaction coordinate. For this reason $q^+(x)$ provides an idal definition of reaction coordinate, i.e. of a one-dimensional projection of the reaction trajectory, because on the hyperplane locally orthogonal to the tangent of $q(x)$ the distribution of points is by definition at equilibrium. More generally, a "good" reaction coordinate can be defined as any collective variable $\xi(x)$ which parameterizes the committor, i.e for which

$$q^+(x) = q^+[\xi(x)]. \tag{4.21}$$

4.2.2 Transition density distribution and transition current

The ergodicity assumption of TPT can be exploited to define a time-independent distribution which measures the probability for transition paths to visit a specific configuration x. In mathematical terms, given an arbitrary function $F(x)$ and denoting with $h_R(x)$, $h_P(x)$ and $h_{\Omega_T}(x)$ as the characteristic functions of the reactant state, product state and transition region, then $m_T(x)$ is defined by the following equality

$$\lim_{T\to\infty} \frac{\int_{-T/2}^{T/2} d\tau\ F[x(\tau)]\ h_R[x(t^-(\tau))]\ h_P[x(t^+(\tau))]\ h_{\Omega_T}[x(\tau)]}{\int_{-T/2}^{T/2} d\tau\ h_R[x(t^-(\tau))]\ h_P[x(t^+(\tau))]\ h_{\Omega_T}[x(\tau)]} \equiv \int\ dx\ F(x)\ m_T(x) \tag{4.22}$$

To explicitly evaluate $m_T(x)$ we make use again of the fact that to the probability that any piece of ergodic trajectory (reactive or not) visits configuration x at time t corresponds to the Gibbs distribution times the probability for trajectories passing at x to be reactive, i.e.

$$m_T(x) = \frac{1}{Z_T} e^{-\beta U(x)}\ q^+(x)\ (1 - q^+(x)), \tag{4.23}$$

where Z_T is the appropriate normalization factor: $Z_T = \int_{\Omega_T} dx e^{-\beta U(x)}\ q^+(x)\ (1 - q^+(x))$.

The transition probability density $m_T(x)$ carries information about the configurations in which the system is most likely to be found while it is performing a transition. The complementary information about the reaction kinetics is encoded in the so-called *transition current* $\mathbf{J}_T(x)$, a $3N$-dimensional vector field which provides the flux of all transition paths across arbitrary hyper-surfaces ∂S which enclose a region S of the configuration (see Fig. 1). In analogy with the transition density $m_T(x)$, also the definition of $\mathbf{J}_T(x)$ is based on exploiting the ergodic hypothesis:

$$\lim_{\Delta\tau\to 0}\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} d\tau(\ h_S[x(\tau)]\ h_{\Omega/S}[x(\tau+\Delta\tau)] - h_{\Omega/S}[x(\tau)]\ h_S[x(\tau+\Delta\tau)])\ \ h_R[x(t^-(\tau))]\ h_P[x(t^+(\tau$$
$$\equiv \int_{\partial S} d\sigma(x)\ \hat{\mathbf{n}}(x)\cdot\mathbf{J}_T(x). \qquad (4.2$$

In this equation, $h_S(x)$ is the characteristic function associated to the region S of configuration space enclosed by the hyper-surface ∂S, Ω/S is its complementary region of configuration and $\hat{n}(x)$ is a unit-norm vector orthogonal to the surface ∂S at point x (see Fig. 1). Like the transition probability density $m_T(x)$, also the transition current can be written in a form which involves the Gibbs distribution and the committor funciton:

$$\mathbf{J}_T(x) = D\nabla q^+(x)\ \frac{e^{-\beta U(x)}}{Z} \tag{4.25}$$

The proof of this equation is too involved to be reported here, and can be found e.g. in Ref. [**?**].

4.3 Potential of Mean-Force

A possible approach which tries to address the questions listed in the previous paragraph follows from exploiting (or boldly assuming) a separation of time scales. For example, the dynamics of the water molecules takes place at a significant shorter time scale that of the torsional angles in the chain, which in turn is much slower than the vibrational dynamics of the bond lengths, and so on. In the presence of separation of time scales, it is useful to introduce the notion of reaction coordinates. While the rigorous definition of reaction coordinate is still a matter of scientific debate, here we shall regard them as collective variables which (i) can be considered order parameters and (ii) evolve according to a dynamics significantly slower than that of all other degrees of freedom in the system. If this is the case, then for each value of the reaction coordinate, the remaining degrees of freedom can be considered as equilibrated.

The knowledge of the system's reaction coordinates paves the way to an effective "low-energy" coarse grained description, defined in terms of a set of "slow" degrees of freedom. This can be done by averaging out the effect of the "fast" degrees of freedom, which are expected to be distributed with Boltzmann. For example, the gyraton radius of a polymer is smaller in the collapsed phase than it is in the coil phase, thus provides a decent order parameter. Assuming it is also a slow variable, we can regard it as a reaction coordinate for studying the collapse of a polymer in a bad solvent.

As the temperature (or the concentration of some denaturing agent such as urea) is increased, one expects R_g to undergo a drastic increase and reach its equilibrium value in the denatured state. Similarly, trajectories connecting different collapsed configurations must have comparable gyration radius, while denaturation trajectories will contain a drastic change in R_g.

Let us now be more quantitative. Let the entire molecule-solvent system be described by a Hamiltonian $\mathcal{H}(p,q)$. At equilibrium, the distribution of generalized coordinates and momenta characterizing the position of the molecule in phase-space will obey Boltzmann's law

$$\mathcal{P}(p,q) = const \times \exp\left(-\frac{1}{k_B T}\mathcal{H}(q,p)\right), \tag{4.26}$$

where the constant is a normalization factor. Let us now $l_1, ...l_{N_l}$ be a set of reaction coordinates defined in terms of the Hamiltonian variable by the relationships

$$l_i = f_{l_i}(p_1, q_1, ..., p_N, q_N) \tag{4.27}$$

We introduce the free energy as a function of $l_1, ..., l_N$, $F(l_1, ..l_{N_l})$:

$$\begin{aligned}\exp\left(-F(l_1,..l_{N_l})/k_BT\right) &= \int d^{3N}q\; d^{3N}p \exp\left(-\frac{1}{k_BT}\mathcal{H}(q,p)\right) \\ &\times \prod_{i=1}^{N_l} \delta[l_i - f_{l_i}(p_1, q_1, ..., p_N, q_N)]\end{aligned} \tag{4.28}$$

If $l_1, ..., l_N$ are order parameters, they can be used to identify the thermodynamical state. For example, in Fig. () we

For example, macromolecules are usually solvated in water, which can be considered as a thermalized heat bath. The collisions with the water molecules occur at a rate of other order one every few fractions of ns, i.e. quite faster than the typical time scale associated with the conformational changes of the system — which is of the order of few ns—. Such collisions can subtract or inject energy in the molecule. If the kinetic energy per degree of freedom of the molecule is much larger than the thermal energy k_BT of the heat bath, the atoms in the molecule will preferentially loose energy, getting cooler. On the other hand, if the initial kinetic energy of the atoms in molecule is much smaller than the thermal energy k_BT, the atoms will on average increase their energy by colliding with the solvent. In both cases, after a transient regime, a complete thermal equilibrium will be reached. and $\mathcal{H}(q,p)$ is the Hamiltonian.

The analysis of the thermodynamics based on the choice of reaction coordinates does not entirely settle the questions we started from. In fact, the choice of the reaction coordinate is somewhat ambiguous and the outcome of the analysis may depend on such a choice. On the other end, the selection of the set of relevant reaction coordinate can be guided by physical intuition. Let us now discuss an example of the application of these concepts to the study of the thermodynamics of the beta hairpin presented in the left panel of Fig. (**??**). In this context, one may introduce the reaction coordinates the gyration radius R_g and the value of the angle at the turn ϕ. The values of R_g and ϕ is expected be smaller in the folded state than in the coil state. The right panel displays the corresponding free energy $F(R_g, \phi)$ evaluated in some model. Clearly, this model predicts a unique thermodynamically stable state, at room temperature. Repeating this calculation at larger temperatures one may be able to reconstruct the structure of phase diagram.

There is of course no restriction to the number of reaction coordinates which can be used. In particular, one can apply the same construction to integrate out only the solvent degrees of freedom and obtain a microscopic description of the system in terms of the molecule internal coordinates only. In this case, let us denote with x the set of generalized coordinates specifying the configuration of the protein and y be the set of generalized coordinates of the water molecules. The equilibrium thermodynamics of the molecule-solvent system is completely contained in the partition function

$$Z(T) = \int d^{3N_x}x \; d^{3N_y}y \; e^{-\frac{1}{k_BT}U(x,y)}, \tag{4.29}$$

where N_x (N_y) is the number of molecule (water) generalized coordinate. The probability of finding the protein in a given configuration $\hat{x}$ is then given by

$$e^{-G(\hat{x})/k_BT} = \int d^{3N_x}x \; d^{3N_y}y \; e^{-\frac{1}{k_BT}U(x,y)}\delta(x-\hat{x}) = \int d^{3N_y}y \; e^{-\frac{1}{k_BT}U(\hat{x},y)}. \tag{4.30}$$

In this context, the free energy $F(\hat{x})$ is sometimes called the *potential of mean force.* The name comes from the fact that the differentiation of F with respect to x leads to

$$-\nabla_i G(x) = -\frac{\int d^{3N_x}x \; d^{3N_y}y \; \nabla_i U(x,y) e^{-\frac{1}{k_BT}U(x,y)}}{\int d^{3N_x}x \; d^{3N_y}y \; e^{-\frac{1}{k_BT}U(x,y)}} = \langle F_i(x) \rangle \tag{4.31}$$

The function $-\nabla_i G(x)$ represents therefore the mean force acting at the configuration x, when the effect of the solvent is averaged out.

Chapter 5

Statistical Computing for Complex Systems

In Chapter 3, we have introduced a theoretical framework to describe stochastic processes in high dimensional configuration space. We have also seen how Markov State Models provide a rigorous low-resolution representation of the complex dynamics in high-dimensional configuration spaces. In this Chapter, we tackle the problem of how extracting information from huge data set obtained integrating the dynamics of high-dimensional complex system.

One of the problem we tackle consists in explicitly constructing the MSMs representation starting from an underlying microscopic dynamics. The main issue there is finding the Markov states using the information available from simulation data. In low dimensional spaces (say, in 1D or 2D) this is a straightforward task. One simply needs to plot the frequency histogram of the configurations visited by long ergodic trajectories. High density peaks correspond to Markov States. Then the kinetic matrix can be constructed by counting the number of transition per unit time from different states.

In high-dimensional spaces, however, the same logical construction is very far from trivial, because of two main technical problems: on the one hand, in the presence of thermal activation transitions are exponentially suppressed, and simulations may not be able to record relevant events (this is generally referred to as the sampling problem). In addition, in the very high-dimensionality limit, any finite data set is effectively infinitely sparse and the distance between points in configuration space tends to infinity. As a result, ergodicity is always badly broken, and the problem arises of how to extract relevant kinetic and thermodynamical information in the absence of it.

Fortunately, finite data set often do not evenly cover the entire configuration space. For example, in macromolecules, because of steric constraints and bond constraints, molecular data at finite statistics may effectively distribute along lower dimensional curved manyfolds (up to small thermal noise). Following this view, one may think of ways to project onto a smaller set of (potentially curvilinear) coordinates which are relevant for the slow dynamics and then use these low-dimensional representations to identify Markov States and the associated dynamical processes.

This quest raises the challenge of finding rigorous dimensionality reduction techniques for data sets representing stochastic dynamics in high dimensional datasets.

This problem often crosses with that of navigating free energy landscapes of complex systems and using this information to infer the structure of the manyfold embedded in configuration space where the dynamics effectively takes place.

Once a satisfactory low-dimension representation has been achieved, the next challenge consists in finding reliable ways of clustering the data, in order to recover regions of high density, which are can identified with Markov States. Finally, the last conceptual steps consists in analyzing the transitions in the trajectories in order to compute the rates between the states. Only after all these steps have been achieved, then the microscopic dynamics can be represented in terms of a network were the directed edges between the nodes are determined by the transition rates between the Markov States.

In the following section we present a brief introduction to how is it possible to tackle all of these problems.

5.1 Diffusion Distance

Let us address the problem of providing a definition for an ideal low-dimensionality representation for the stochastic dynamics of some complex system. To this end, we shall assume that the system's dynamics obeys a Fokker-Planck equation with a decoupling of relaxation time scales, i.e. a gap between soft and hard modes. In this case, it is key to introduce a norm which identifies how quickly diffusion can connect two configurations x, y. This is done through the notion of diffusion distance (sometimes called kinetic distance):

$$d_\tau(x,y) = \sqrt{\int dz (P(z,\tau|x) - P(z,\tau|y))^2 \frac{1}{\pi(z)}} \tag{5.1}$$

where $\pi(x) = \frac{1}{Z} e^{-\beta U(x)}$ is the Boltzmann's distribution. This equation, in fact defines a whole family of norms, parametrized by the (short) time scale τ. $d_\tau(x,y)$ is zero any time a stochastic dynamics initiated from the two configurations x and y converges to the same distribution within a time scale τ. In other words, if x and y are in the same basing d_τ vanishes as long as τ is chosen longer than the typical relaxation time within the basin. Conversely $d(x,y)$ remains finite any time x and y belong to kinetically separated states, with interconversion time $t \gg \tau$.

After performing the spectral decomposition of Eq. (5.1) and exploiting the orthogonality relationship $\int dy L_n(y) R_m(y) = \delta_{nm}$ one arrives to:

$$d_\tau(x,y) = \sqrt{\sum_{n=1}^{M-1} e^{-2\lambda_n \tau} \, (L_n(x) - L_n(y))^2} \tag{5.2}$$

where $M-1$ is the label associated to the largest eigenfrequency *below the gap*. In the limit of large gap, $e^{-2\lambda_n \tau} \sim 1$, thus the diffusion distance reduces to a standard Euclidean distance, in the space of collective variables represented by the lowest M left-eigenvalues of the Fokker-Plank operator. In this sense, they represent the ideal global coordinates. The problem now arises on how is it possible to obtain estimates of these collective variables using a finite and sparse simulation data set.

To this end, we shall begin by introducing some useful mathematical definitions, which will be then used to obtain estimates for $L_n(x)$.

5.1.1 Mathematical Prologue

We consider a stochastic dynamics defined by a general Fokker-Plank equation

$$\partial_t P(x,t) = -\hat{H}_{FP} P(x,t) \tag{5.3}$$

and let $P(x,t|y,t)$ be the corresponding Green's function.

We now define two convenient Hilbert spaces. The first, denoted with $\mathcal{L}^2_\pi$, corresponds to the linear space of square-integrable functions $\mathcal{L}^2$, equipped with the following inner product:

$$\langle a|b\rangle_\pi = \int dx a(x) b(x) \pi(x), \tag{5.4}$$

where $\pi(x) = \frac{1}{Z} e^{-\beta U(x)}$ is the Boltzmann's measure. The second, denoted with $\mathcal{L}^2_{\pi^{-1}}$ the linear space of square-integrable functions $\mathcal{L}^2$, equipped with following inner product:

$$\langle a|b\rangle_{\pi^{-1}} = \int dx a(x) b(x) \frac{1}{\pi(x)}, \tag{5.5}$$

Let us now define some useful operators on these two Hilbert spaces. In particular, we focus on the following operators:

$$\text{Propagating Operator } \hat{P}_\tau \quad : \quad \hat{P}_\tau \rho(x,t) = \int dy P(x,t+\tau|y,t)\rho(y,t) \tag{5.6}$$

$$\text{Transfer Operator } \hat{T}_\tau \quad : \quad \hat{T}_\tau \rho(x,t) = \frac{1}{\pi(x)} \int dy\ P(x,t+\tau|y,t)\ \rho(y)\pi(y) \tag{5.7}$$

It is important to note that, since the Fokker-Planck dynamics obeys detailed balance, the propagating operator is Hermitian in $\mathcal{L}_2^{\pi^{-1}}$:

$$\begin{aligned}
\langle v|(\hat{P}_\tau u)\rangle_{\pi^{-1}} &= \int dy\ v(y) \frac{1}{\pi(y)} \int dx P(y,t+\tau|x,t) u(x) \\
&= \int dy\ v(y) \int dx \frac{P(x,t+\tau|y,t)}{\pi(x)}\ u(x) \\
&= \langle (\hat{P}v)|u\rangle_{\pi^{-1}}
\end{aligned} \tag{5.8}$$

The propagating and transfer operators can be formally written in terms of the Forkker-Plank operator and the backward Kolmogorov operator. Indeed, from Eq. (5.3) it immediately follows

$$\hat{P}^\tau \rho(x,t) = e^{-\hat{H}_{FP}\tau} \rho(x,t). \tag{5.9}$$

To obtain a similar expression for $\hat{T}^\tau$ we note that if $\rho(x,t)$ is a solution of the Fokker-Planck equation, then the combination

$$u(x,t) = \rho(x,t)/\pi(x) \tag{5.10}$$

is a solution of the so-called backward Kolmogorov equation,

$$\partial_t u(x,t) = -\hat{H}^\dagger_{FP} u(x,t), \tag{5.11}$$

where the adjoint of the Fokker-Planck operator,

$$\hat{H}_{FP}^{\dagger} = D(\nabla^2 - \beta \nabla U \nabla) \tag{5.12}$$

is called the backward Kolmogorov operator. The proof of (5.11) can be immediately obtained by direct inspection. Then,

$$\hat{T}^{\tau} = e^{-\hat{H}_{FP}^{\dagger}\tau}. \tag{5.13}$$

The relationship between the propagator and transfer operators follow reflects the fact that they are related by a unitary multiplicative operator $\hat{M}$:

$$M : \mathcal{L}^2_{\pi^{-1}} \to \mathcal{L}^2_{\pi} : \hat{M}\rho(x) = \pi(x)\rho(x) \tag{5.14}$$

Indeed,

$$\begin{aligned} \hat{T}_{\tau}\rho(x,t) &= \frac{1}{\pi(x)} \int dy P(x, t+\tau|y,t)\ \rho(y)\pi(y) \\ &= \frac{1}{\pi(x)} \int dy P(x, t+\tau|y,t)\hat{M}\ \rho(y) \\ &= \hat{M}^{-1} \int dy P(x, t+\tau|y,t)\hat{M}\ \rho(y), \end{aligned} \tag{5.15}$$

from which, we get the operator equation $\hat{T}_{\tau} = \hat{M}^{-1}\hat{P}_{\tau}\hat{M}$.

It is easy to prove that this relation implies that $\hat{T}_{\tau}$ and $\hat{P}_{\tau}$ share the same spectrum, and that the eigenfunctions $|l_i\rangle$ transfer operator are related to those $|r_i\rangle$ of $\hat{P}_{\tau}$ by the relation $|l_i\rangle = \hat{M}^{-1}|r_i\rangle$. Indeed,

$$\hat{T}_{\tau}|l_i\rangle = \lambda_i|l_i\rangle \Rightarrow M^{-1}\hat{P}_{\tau}\hat{M}|l_i\rangle = \lambda_i|l_i\rangle \Rightarrow \hat{P}_{\tau}(\hat{M}|l_i\rangle) = \lambda_i\ (\hat{M}|l_i\rangle). \tag{5.16}$$

Then, defining $|r_i\rangle = \hat{M}|l_i\rangle$, we obtain

$$\hat{P}_{\tau}|r_i\rangle = \lambda_i|r_i\rangle. \tag{5.17}$$

Finally, from the expressions (5.9) and (5.13), recalling the fact that the spectrum of $\hat{H}_{FP}$ is non negative, it follows that $\lambda_0 = 0 > \lambda_1 > \ldots$.

Variational Estimate of the Diffusion Distance

Let us now return to our original problem of estimating the diffusion distance between configurations. We recall that, due to Eq. (5.2), this problem is basically equivalent to finding approximation for the left eigenstates of the Fokker-Plank operator or equivalently, in view of Eq. (5.13), the eigenstates of the transition operator $\hat{T}^{\tau}$. Equipped with the mathematical notions we have just introduced, we can now tackle this problem.

Let us consider the following matrix element:

$$\begin{aligned} G_f(\tau) &\equiv \langle f|\hat{T}_{\tau}|f\rangle_{\pi} \qquad (5.18) \\ &= \int dx \int dy f(x) f(y)\ P(x,\tau|y)\ \pi(y) \qquad (5.19) \end{aligned}$$

So, $G_f(\tau)$ is the time auto-correlation of the function $f(x)$, along stochastic trajectories evolving according to the Fokker-Planck dynamics. We note that, in practice, $G_f(\tau)$ can be estimated directly from simulations using the central limit theorem:

$$G_f(\tau) \simeq \frac{1}{N}\sum_{k=1}^{N} f[x(t_k)]\; f[x(t_k+\tau)]. \tag{5.20}$$

Let us now consider a specific choice for the function $f(x)$, namely on in which it coincides with an eigenstate of the Transfer operator, $f(x) = l_i(x)$. In this case, for any given choice of the lag-time τ, the time autocorrelation function $G_{l_i}(\tau)$ is simply the corresponding eigenvalue λ_i:

$$G_{l_i}(\tau) = \langle l_i^\tau|\hat{T}_\tau|l_i^\tau\rangle = \lambda_i^\tau \tag{5.21}$$

where we have emphasized that the eigenstates and the eigenvalues parametrically depend on the lag-time τ.

Let us assume now we have fund some *approximation* (proxy) $\tilde{l}_1^\tau(x)$ of the 1−st non-trivial eigenstate $l_i^\tau(x)$ of $\hat{T}$ (the ground-state is trivially $\lambda_0^\tau(x) = 1$). This proxy is requested to obey the correct orthogonality condition to $l_0^\tau(x)$, with respect to the $\langle\rangle_\pi$ inner product:

$$\langle l_0^\tau|\tilde{l}_i^\tau\rangle_\pi = 0 \tag{5.22}$$

In practice, this implies that $l_1^\tau(x)$ is a mean-free function, according to the Boltzmann's measure:

$$\langle \tilde{l}_1^\tau|\hat{l}_0^\tau\rangle = \int dx \tilde{l}_1^\tau \pi(x) = 0. \tag{5.23}$$

Then, it is immediate to prove the following variational bond:

$$G_{\tilde{l}_1^\tau}(\tau) \leq \lambda_1^\tau. \tag{5.24}$$

The proof is immediate: since the transfer operator is Hermitian with respect to the $\langle\cdot\rangle_\pi$ inner product, we can Fourier decompose $\tilde{l}_1^\tau(x)$:

$$\tilde{l}_1^\tau(x) = \sum_i c_i l_i^\tau(x) \tag{5.25}$$

Then

$$G_{\tilde{l}_1^\tau}(\tau) = \sum_{ij} c_i c_j \langle l_i^\tau|\hat{T}_\tau|l_j^\tau\rangle_\pi = \sum_{ij} c_i c_j \lambda_i^\tau \delta_{ij} = \sum_i c_i^2 \lambda_i^\tau \leq \lambda_1^\tau \sum_i c_i^2 = \lambda_1^\tau. \tag{5.26}$$

Let us now discuss the implication of this variational bound. To this end, let us suppose to find a basis set formed by N orthogonal functions. We recall that orthogonality with the given scalar product corresponds to the condition of being uncorrelated with respect to the realization of the stochastic process:

$$\begin{aligned} &\langle \chi_n|\chi_m\rangle_\pi = 0 \\ &\int dx \chi_n(x)\chi_m(x)\pi(x) = \frac{1}{N}\sum_j \chi_n[x(\Delta t j)]\chi_m[x(\Delta t j)] \end{aligned} \tag{5.27}$$

Let us now compute the correlation matrix:

$$\tilde{T}^{\tau}_{nm} = \langle \chi_n | \hat{T}_{\tau} | \chi_m \rangle \tag{5.28}$$

Microscopic reversibility implies that this matrix must be symmetric. When estimated with stochastic descriptors at finite statistics, the matrix is no longer symmetric but may be symmetrized "by hand":

$$\tilde{T}^{\tau}_{nm} \to \frac{1}{2} \left(\tilde{T}^{\tau}_{nm} + \tilde{T}^{\tau T}_{nm} \right) \tag{5.29}$$

Let us now imagine to diagonalize this matrix:

$$\sum_{n} \tilde{T}^{\tau}_{nm} \, b_m = \tilde{\lambda}^{\tau}_{n} b_n, \tag{5.30}$$

Since the $\tilde{T}^{\tau}$ matrix is an approximate representation fo the transfer operator in the truncated space of orthogonal functions $\{\chi_n\}_n$, its eigenvectors $\{b^m_k\}$ yield a proxy of the transition operator eigenfunctions:

$$\tilde{l}^{\tau}_{n}(x) \simeq \sum_{k} b^{m}_{k} \chi^{\tau}_{m}(x). \tag{5.31}$$

In practical situations, however, the functions χ^{τ}_{n} are simply collective variables guessed upon chemico/physical intuition. As such, they are usually not in general mutually orthogonal. In this case, one needs to solve the generalized eigenvalue problem:

$$\sum_{m} \tilde{T}_{nm} b_m = \lambda^{\tau}_{n} \left(\sum_{k} S_{nk} b_k \right) \tag{5.32}$$

where $S_{nm} = \langle \chi_n | \chi_m \rangle$.

Once the eigenvalues and eigenstates have been obtained, then the inequality (5.26) implies the first eiqenvalue $\tilde{\lambda}^{\tau}_{1}$ can be used as variational parameter to estimate the goodness of the approximation. This is exploited in the so-called Time Independent Component Analysis (TICA) algorithm, which seeks the best linear combination of input features approximating the eigenstates of the transfer operator, thus the collective coordinates required to define the diffusion distance (5.2).

An illustrative example (taken from Wikipedia) of the MSM obtained from the analysis of massively distributed molecular simulations of protein folding (based on Folding@home[1] platform) is provided in Fig. (5.1.1).

5.1.2 Markov State Model Construction

Let us finally address the question of how the spectral methods discussed so far can be used to construct a Markov State Model (MSM) representation of the stochastic dynamics of a complex system. As a paradigmatic reference problem, we imagine to developing such a model by analyzing the result of molecular dynamics simulations of some complex macromolecular system. The following summary is closely taken

---

[1] https://foldingathome.org/

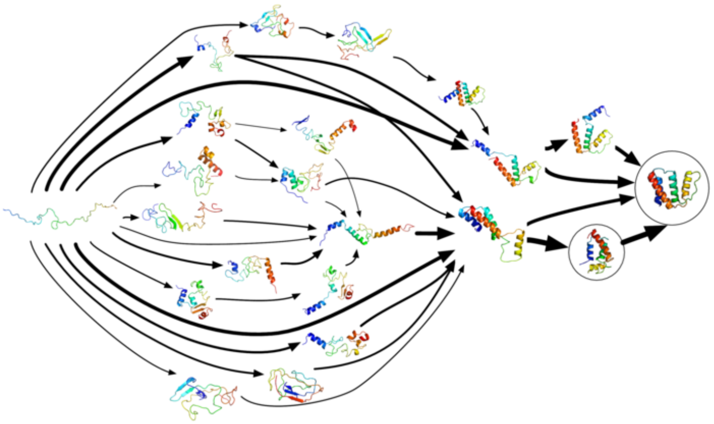

Figure 5.1: Folding@home uses MSMs, like the one diagrammed here, to model the possible shapes and folding pathways a protein can take as it condenses from its initial randomly coiled state (left) into its native 3D structure (right). (Figure and caption taken from Wikipedia)

from a recent review paper[20] introducing a special issue of the Journal of Chemical Physics, specidically devoted to this problem.

Most MSM construction approaches proceed by a sequence, or pipeline, of data processing steps that typically includes the following:

1. *Featurization:* The MD coordinates are transformed into a set of featuring functions $R_1(x), R_2(x), \ldots, R_N(x)$. Features are chosen to represent the essential dynamics of the system, i.e. the assumption is that the lowest-lying eigenstates of the backward Kolmogorov operator depend on the configurations through these function, i.e. $L_n(x) = L_n(R_1(x), \ldots, R_N(x))$. As an example, for simulations of protein dynamics example of features may include proteins inter- residue distances, the contact maps, or torsion angles

2. *Dimensionality reduction:* Spectral methods developed in the previous sections is used to reduced the dimensionality to typically 2-100 slow collective variables (CVs), e.g. corresponding to the lowest left-eigenstates of the Kolmogorov operator.

3. *Discretization:* The low-dimensional space constructed at the previous point may be discretized by clustering the projected data, typically resulting in 100-1000 discrete "microstates."

4. *MSM estimation:* A transition matrix or rate matrix describing the transition probabilities or rate between the discrete states at some lag time τ.

5. *Coarse-graining:* In order to get an easier interpretable kinetic model, the

MSM from step 5 is often coarse-grained to a few states (for an example of systematic, renormalization group based coarse graining, see [21].

5.2 Diffusion Maps

Let us now discuss a closely related spectral method which is often used for interpretation and feature extraction of large data sets and also used in the framework of molecular simulations.

The basic idea behind this approach is to construct *a posteriori* an artificial random walk on our points so to recover information about the dynamical process that generated them. To this purpose we introduce the following objects[2]

- An isotropic diffusion kernel

  $$K_\varepsilon(x_i, x_j) = \frac{1}{\sqrt{4\pi}} \exp\left(-\frac{\|x_i - x_j\|^2}{4\varepsilon}\right) \tag{5.33}$$

  where $||\ldots||$ is the Euclidean distance in configuration space and ε is a fixed time scale. This kernel introduces some notion of affinity between points in configuration space and we have chosen an exponentially decaying function relying on the idea that, in many applications, only high correlation values correspond to meaningful information on the data set.

- A normalization

  $$p_\varepsilon(x_i) = \sum_{n=1}^{N} K_\varepsilon(x_i, x_n)$$

  . We will show that this is an approximation of the true probability distribution $p(x)$, for small values of ε.

- A new symmetric kernel

  $$\tilde{K}_\varepsilon(x_i, x_j) = \frac{K_\varepsilon(x_i, x_j)}{\sqrt{p_\varepsilon(x_i)p_\varepsilon(x_j)}} \tag{5.34}$$

  and another normalization factor

  $$d_\varepsilon(x_i) = \sum_{n=1}^{N} \tilde{K}_\varepsilon(x_i, x_n) \tag{5.35}$$

---

[2] Most of the literature on Diffusion Map and related schemes is being produced by mathematicians. In their community the conventional notation for Green's function and conditional probabilities is one in which the initial condition is written in leftmost position, followed by the final position, e.g. $P_t(x, y) = \text{Prob.}(x \to y)$. On the other hand, physicists tend to adopt the opposite notation, in analogy with the bra-ket notation of quantum mechanics. Indeed, the propagator $K_t(x_f; x_i) = \langle x_f | e^{-i\frac{i}{\hbar}\hat{H}t} | x_i \rangle$ represents the amplitude to perform a transition from x_i to x_f in time t. Throughout these notes, we have been following the physicists' notation. However, in this section, we shall switch to the matematicians' convention, to facilitate the reader in making contact with the existing literature.

- A Markov matrix
$$M_\varepsilon(x_i|x_j) = \frac{\tilde{K}_\varepsilon(x_i, x_j)}{d_\varepsilon(x_i)}, \tag{5.36}$$
which satisfies the sum rule $\sum_{j=1}^{N} M_\varepsilon(x_i|x_j) = 1$. Matrices that satisfy this condition are referred to as right stochastic matrices. Notice that we used the notation for conditional probability, but since M is meant to be a *backward* kernel, the probability is conditioned on events which happen later in time. [3] Therefore, the sum rule simply translates into conservation of probability, namely the total probability to end up at x_j starting from any other point is 1.

- The evolution operator T, whose kernel is our Markov matrix. Its action on test functions is, by definition,
$$T[\varphi]\,(x_j) = \sum_{j=1}^{N} M(x_i|x_j)\varphi(x_j) \tag{5.37}$$

Now, we examine the limit $N \to \infty$, that leads to a diffusion process continuous in space but discrete in time. Exploiting the assumption that our data set was obtained by N realization of a stochastic process defined a given probability distribution, we can rely on the Central Limit theorem:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(x_n) = \mathbb{E}[g(x)] = \int g(x)\; p(x)dx \tag{5.38}$$

This allows us to convert the sums in our discrete definitions into integrals with measure $p(x)$, leading to

- A continuous formulation of the normalization condition,
$$p_\varepsilon(x) = N \int K_\varepsilon(x, z)\; p(z)dz. \tag{5.39}$$

- A new representation for the symmetric kernel:
$$\tilde{K}_\varepsilon(x, y) = \frac{K_\varepsilon(x, y)}{\sqrt{p_\varepsilon(x)p_\varepsilon(y)}}, \tag{5.40}$$
and a continuous normalization factor,
$$d_\varepsilon(x) = N \int \tilde{K}(x, z)\; p(z)dz,. \tag{5.41}$$

- In the new formulation the Markov matrix elements are
$$M_\varepsilon(x|y) = \frac{\tilde{K}_\varepsilon(x, y)}{d_\varepsilon(x)}, \tag{5.42}$$
and the associated right stochasticity condition reads:
$$\int M(x|y)\; p(y)dy = 1. \tag{5.43}$$

---

[3] M represents the probability for the system that is eventually found in configuration x_j arrives there starting from point x_i

- In the continuous formulation, the evolution operator is now defined to act upon test functions as follows:

$$T[\varphi]\,(x) = N \int M(x|y)\varphi(y)\; p(y)dy. \tag{5.44}$$

To appreciate the physical meaning of what we are doing, let us specialize on the case in which our points are sampled from an equilibrium probability distribution. We will see that $p_\varepsilon(x)$ will be soon proven to become the true probability density plus terms of order ε.

In the following, we shall uncover the relation between our evolution operator T and he backward Kolmogorov operator:

$$\lim_{\varepsilon\to 0} \frac{1-T}{\varepsilon}\varphi(x) = \beta\nabla U \cdot \nabla\varphi(x) - \nabla^2\varphi(x) = H_b\,\varphi(x). \tag{5.45}$$

This result implies that, by diagonalizing the matrix $M(x_i|x_j)$ we can find an approximation for eigenvalue and eigenfunctions of $\hat{H}_b = \hat{H}^\dagger_{FP}$, which bear the information about the relevant slow dynamics in the system and the diffusion distance.

To provide the proof of Eq. (5.45), let us introduce the continuous time limit, $\varepsilon \to 0$. Since we are also assuming the limit of large number of points ($N \to \infty$) , we need to specify the relative scaling between ϵ and N. We choose $\varepsilon \sim \frac{1}{N^2}$, which corresponds to assuming that the time intervals required to diffuse between nearest-neighbor points in our dataset scales like the square of their distance, i.e. the typical law of Brownian dynamics.

The next step towards proving Eq. (5.45) is to show that $p_\varepsilon(x)$ approximates the equilibrium distribution $p(x)$, to leading order in ε, i.e.

$$p_\varepsilon(x) = p(x) + \varepsilon\nabla^2 p(x) + O(\varepsilon^2). \tag{5.46}$$

To this end, we perform a Fourier transform of both sides of Eq. (5.39) Recalling that the kernel is a Gaussian function (see Eq. (5.33)), we find:

$$p_\varepsilon(q) = e^{-q^2\varepsilon}\; p(q) = \left(1 - q^2\varepsilon\right) p(q) + O(\varepsilon^2). \tag{5.47}$$

Eq.(5.46) follows to anti-transforming back, from momentum to coordinate space.

Let us now plug this result into the expression for the normalization constant $d_\varepsilon(x)$:

$$\begin{aligned} d_\varepsilon(x) &= N \int \tilde{K}(x,y)\; p(y)dy = N \int \frac{K(x,y)}{\sqrt{p_\varepsilon(x)p_\varepsilon(y)}}\; p(y)dy \\ &\simeq N \int \frac{K(x,y)}{\sqrt{p(x)\left(1+\varepsilon\frac{\nabla^2 p(x)}{p(x)}\right)p(y)\left(1+\varepsilon\frac{\nabla^2 p(y)}{p(y)}\right)}}\; p(y)dy \\ &\simeq p^{-\frac{1}{2}}(x)\left(1 - \frac{\varepsilon}{2}\frac{\nabla^2 p(x)}{p(x)}\right) N\; \times \\ &\quad \left(\int K(x,y)p^{\frac{1}{2}}(y)\; dy - \frac{\varepsilon}{2}\int K(x,y)p^{\frac{1}{2}}(y)\frac{\nabla^2 p(y)}{p(y)}\; dy\right). \end{aligned} \tag{5.48}$$

Now, at first order in ε we have

$$\begin{aligned} N\int K(x,y)p^{\frac{1}{2}}(y)\ dy &\simeq p^{\frac{1}{2}}(x)\left(1+\varepsilon\frac{\nabla^2 p^{\frac{1}{2}}(x)}{p^{\frac{1}{2}}(x)}\right)\\ \frac{\varepsilon}{2}N\int K(x,y)p^{\frac{1}{2}}(y)\frac{\nabla^2 p(y)}{p(y)}\ dy &\simeq \frac{\varepsilon}{2}p^{\frac{1}{2}}(x)\frac{\nabla^2 p(x)}{p(x)}. \end{aligned}$$

In conclusion, we find

$$d_\varepsilon(x) = 1+\varepsilon\left(\frac{\nabla^2 p^{\frac{1}{2}}(x)}{p^{\frac{1}{2}}(x)} - \frac{\nabla^2 p(x)}{p(x)}\right) + O(\varepsilon^2). \tag{5.49}$$

Finally, we are in a condition to examine the action of the evolution operator T

$$T[\varphi]\,(x) = N\int M(x|y)\varphi(y)\ p(y)dy = \frac{1}{d_\varepsilon(x)}N\int \tilde{K}(x,y)\varphi(y)\ p(y)dy. \tag{5.50}$$

After poerforming the same steps that lead to (5.49), we find

$$T[\varphi]\,(x) = \frac{\varphi(x)}{d_\varepsilon(x)}\left[1+\varepsilon\left(\frac{\nabla^2\left(\varphi(x)p^{\frac{1}{2}}(x)\right)}{\varphi(x)p^{\frac{1}{2}}(x)} - \frac{\nabla^2 p(x)}{p(x)}\right)\right] + O(\varepsilon) \tag{5.51}$$

Now plugging in the expression (5.49) for $d_\varepsilon(x)$, we obtain:

$$\begin{aligned} T[\varphi]\,(x) &= \varphi(x) + \varepsilon\left(\frac{\nabla^2\left(\varphi(x)p^{\frac{1}{2}}(x)\right)}{p^{\frac{1}{2}}(x)} - \varphi(x)\frac{\nabla^2 p^{\frac{1}{2}}(x)}{p^{\frac{1}{2}}(x)}\right) + O(\varepsilon^2)\\ &= \varphi(x) + \varepsilon\left(\frac{2\nabla\varphi(x)\nabla p^{\frac{1}{2}}(x)}{p^{\frac{1}{2}}(x)} + \nabla^2\varphi(x)\right) + O(\varepsilon^2). \end{aligned} \tag{5.52}$$

We are ready to compute the generator of this transformation, whose action on test functions is

$$\lim_{\varepsilon\to 0}\frac{1-T}{\varepsilon}\varphi(x) = -\frac{2\nabla\varphi(x)\nabla p^{\frac{1}{2}}(x)}{p^{\frac{1}{2}}(x)} - \nabla^2\varphi(x). \tag{5.53}$$

Let us now specify on the case in which the equilibrium probability distribution is the Boltzmann's measure, $p(x) = e^{-\beta U(x)}$. Plugging this into (5.53) we finally get

$$\lim_{\varepsilon\to 0}\frac{1-T}{\varepsilon}\varphi(x) = \beta\nabla U\cdot\nabla\varphi(x) - \nabla^2\varphi(x) = H_b\,\varphi(x), \tag{5.54}$$

thus proving Eq. (5.45).

Therefore, the diffusion map formalism enables an estimate for the eigenvectors of the H_b operator, given a sample of points distributed according to some equilibrium density $p(x)$. Given a data point $\bar{x}$, one obtains a corresponding set of values $L_1(\bar{x}), L_2(\bar{x},\ldots$. Therefore, it establishes a non-linear mapping of the data set onto a linear subspace in which the Euclidean distance corresponds to the diffusion distance (see Eq. ()). In this representation, points are close if the system can rapidly diffuse between them. This feature yields the information about the presence of metastable states (basis) and qualitative information about the reaction process.

An illustrative example (taken from Wikipedia) is provided in Fig. (**??**).

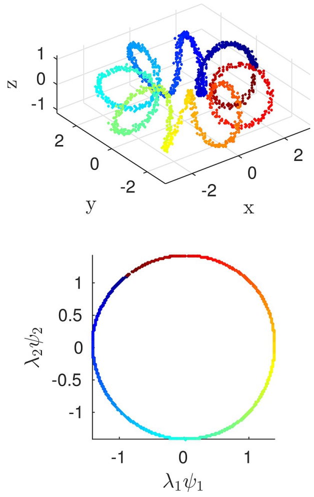

Figure 5.2: Given non-uniformly sampled data points on a toroidal helix (top), the first two Diffusion Map coordinates are plotted (bottom). The Diffusion Map unravels the toroidal helix recovering the underlying intrinsic circular geometry of the data. (Figure and caption taken from Wikipedia)

5.3 Geometric Methods

The clustering and data representation methods sketched in the previous section rely on spectral properties of the stochastic dynamics, thus are applicable only to systems obeying the microscopic reversibility condition and, at least in principle, displaying a gap in the spectrum of relaxation time scales. In many situations, however, these methods are either not applicable, or are not reliable. For example, in order to lower the computational costs of exploring energy landscapes, many algorithms include artificial biasing forces which violate the microscopic reversibility condition. Clustering of those data cannot be performed using r any spectral method, since precisely the purpose of these algorithm is that of reducing the gaps to speed up relaxation dynamics. In addition, at least in principle, the variational estimates used to perform dimensionality reduction in TICA may introduce uncontrolled systematic errors. Finally, even relying on microscopically reversible sampling, the spectrum of the Fokker Planck operator for the underlying physical system may not contain significant gaps, thus posing theoretical concerns about the very relevance of spectral projection-based approaches.

In this section, we illustrate a different theoretical framework to perform dimensionality reduction and analysis of large dimensional data sets. This particular scheme was developed by Laio and co-workers [24, 25]. These methods do not rely on any specific spectral property, but rather are derived by combining differential geometry and statistics. As a results, these clustering schemes apply to a broader set of problems and approaches.

5.3.1 Intrinsic Dimensionality

We have already discussed how the issue of dimensionality reduction is related to the fact that most of the time high dimensional datasets contain most of the information within a lower dimensional manifold.

A key issue is therefore that of estimating the dimensionality of the embedding manifold, often called Intrinsic Dimensionality (ID). For example, in the case of samples generated by molecular dynamics simulations, the real dimensionality of the data set is obviously $3N$ (where N is the number of atoms). However, due to interactions (e.g. topological constrains induced by covalent bonds or steric clashes between neighboring atoms) most the directions are prohibited, thus the directions along the system can move significantly are much fewer than $3N$. Then, an estimator of the ID should give the number of soft directions characterizing the effective dynamics of the system. Here, we sketch the derivation of an algorithm derived by Laio and co-workers which enables to efficiently solve this task

Let us consider a point in the dataset (arbitrarily identified by the index $i = 0$) and the list of it first k nearest neighbors. Let $r_1, r_2, \dots, r_k$ be a sorted list of their distances from the reference point. For example, r_1 is the distance between the point $i = 0$ and its nearest neighbor, r_2 its distance with the second nearest neighbor, and so on.

Let us consider the hyper-spherical volume enclosed by two successive neighbors l and $l + 1$ setting as centre of the hyper-spheres the reference point $i = 0$:

$$\Delta\nu_l = \omega_d (r_{l+1}^d - r_l^d), \tag{5.55}$$

where ω_d is the volume of a unitary radius hypersphere in d dimension, i.e.

$$\omega_d = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2} + 1\right)} \tag{5.56}$$

It is possible to show that, in the case of uniform distribution of density ρ, the values of the shell volume $\Delta\nu$ are distributed according to

$$P(\Delta v \in [\nu, \nu + d\nu]) = \rho e^{-\rho\nu} d\nu \tag{5.57}$$

Let us then consider two shells, $\Delta\nu_i$ and $\Delta\nu_i$ and take the ratio $R = \frac{\Delta\nu_i}{\Delta\nu_j}$. The values of R are distributed according to:

$$\begin{aligned} P(R \in [\bar{R}, \bar{R} + d\bar{R}] &= \int_0^\infty d\nu_j \int_0^\infty d\nu_i \; \rho^2 e^{-\rho(\nu_1 + \nu_2)} \; \chi_{[\bar{R}, \bar{R} + d\bar{R}]} \left(\frac{\nu_i}{\nu_j}\right) \\ &= d\bar{R} \frac{1}{(1 + \bar{R})^2}, \end{aligned} \tag{5.58}$$

where we have denoted with $\chi_A(x)$ the characteristic function of the set A, i.e.

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases} \tag{5.59}$$

Hence, in any dimension and for any choice of i and j, the probability density function of values of the ratio R, $g(R)$, is given by:

$$g(R) = \frac{1}{1 + R^2} \tag{5.60}$$

The idea is to estimate from data a quantity which depends on $g(R)$ and on the dimensionality, thus by computing this quantity and knowing the exact expression for $g(R)$ one obtains an estimate of the effective manyfold's dimensionality over which the points are (locally) uniformly distributed. In order to assess such a local information, we focus on the neighborhood of the point $i = 0$, i.e. we choose $i = 1$ and $j = 2$. This means we are interested in comparing the ratio of volume of hyperspheres between the 1st and 2nd and the 2nd an 3rd nearest neighbors, $R = \frac{\Delta\nu_2}{\Delta\nu_1}$. We then compute the cumulative distribution function of

$$\mu = \frac{r_1}{r_2} \tag{5.61}$$

It is then possible to show that μ and R are related by simple geometrical relationship:

$$R = \mu^d - 1 \tag{5.62}$$

Since the distribution of R is known, it is possible to compute the cumulative distribution of μ, $F(\mu)$:

$$F(\mu) = (1 - \mu^{-d})\chi_{[1,\infty]}(\mu) \tag{5.63}$$

The key point here is that the cumulative distribution $F(\mu)$ depends on d but not on the density ρ. Moreover, the following simple relation applies

$$-\frac{\log(1 - F(\mu))}{\log(\mu)} = d. \tag{5.64}$$

By estimating the function on the left-hand side, one can infer the intrinsic dimensionality.

5.3.2 Density Estimator

In this subsection, we discuss a point-adaptive k-nearest neighbour (PAk) estimator, an algorithm used to compute the local value of the probability density function, from a sparse data set. It has been developed starting from the classic k-nearest neighbor estimator k-NN, in which the local density at a point is measured from the volume in configuration space occupied by the k nearest neighbors of a point. The value k to consider represents a cut-off on the length scale of "influence" of a single point. In the case of highly inhomogeneous data-sets, this global parameter may induce systematic errors. Indeed, the value of k should be such that within the k-nearest neighbors the density does not vary much. In complex cases it is difficult to find a global value for k describing the geometry of the whole dataset.

PAk represents a development of k-NN in two important ways. First, in PAk, k which is adapted at every point in the dataset, rather than globally fixed. Indeed, the algorithm performs a procedure based on finding, for each point, the size of the neighborhood in which the probability density function is constant. This way the the optimal value of k for every point on the dataset is identified. The second important difference is that this procedure is performed in the sub-manifold of intrinsic dimension.

Eq. (5.57) provides the probability density that the j-th and $j+1$-th define a shell of volume ν_j, given a (locally) uniform density ρ. This distribution can be used to defined the log-likelihood function of the parameter ρ, given the realizations of the k-nearest neighbor distances from point i:

$$\mathcal{L}(\rho|\{\nu_i^{(l)}\}_{l\leq K}) = k\log\rho - \sum_{l=1}^{k}\nu_i^{(l)} \equiv k\log\rho - \rho V_{i,k} \tag{5.65}$$

This function is defined in such a way that by imposing stationarity with respect to variation of the density field ρ, $\frac{\delta\mathcal{L}}{\delta\rho}=0$, one gets

$$\rho = \frac{k}{V_{i,k}} \tag{5.66}$$

with an error $e_p = \rho/\sqrt{k} = \sqrt{k}/V_{i,k}$. To be noticed that ν_i are computed with respect the intrinsic dimensionality of the system. Thus, the method computes the density by implicitly projecting the system in the lower dimensional manifold.

The error on the estimate decreases with k. On the other hand, as one increases the size of the neighborhood (choosing larger it is more likely having density variation. The authors of the approach have proposed a way to determine an "optimal" value for k, which corresponds to the biggest value within which density variation is not statistically signicant. Without entering too much in the details, they find $k_{opt} = \hat{k}_i$ considering two models: one for which the density is constant from points i until the neighbours $k+1$ and one for which density at point i is different with respect to the one at point $k+1$. Then, the log-likelihoods of the two models are compared by taking the difference, D_k. If D_k is larger than a threshold D_{thr} then two models are distinguishable, with an accuracy depends on the choice of D_{thr}. The optimal value for k is the biggest value such that it holds $D_k < D_{thr}$.

Chapter 6

Quantum Dynamics and Thermodynamics in Open Systems

In this chapter we extend to the quantum level some of the concepts and results discussed in the previous chapters concerning dynamics in open environment. This problem is relevant also for applications to complex biological or organic macromolecular systems, where electronic excitations (let them be electrons, holes or excitons) can propagate in a molecular environment, which plays the role of a "heat-bath".

We first show how the path integral formalism can be used to define the quantum canonical partition function of a system of non-degenerate particles (sometimes called "Bolzmanions"). Next, we turn our attention to dynamical problems: We discuss the Feynman-Vernon path integral approach to evaluating the time evolution of the density matrix in an open quantum system. We then provide a brief introduction to the general structure of the Quantum Master equation and its Markovian limit.

6.1 Equilibrium Dynamics and Thermodynamics

6.1.1 Path Integral Representation of the Quantum Partition Function

Let us begin by considering a quantum particle evolving according to the Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2M} + V(\hat{q})$. We recall that the real-time quantum evolution of this system is exhaustively described by the Feynman propagator:

$$\begin{aligned} K(q'', t''|q', t') &= \langle q''| e^{-\frac{i}{\hbar}(t''-t')\hat{H}} |q'\rangle \\ &= \int \mathcal{D}q(t)\; e^{\frac{i}{\hbar}S[q]}, \end{aligned} \tag{6.1}$$

where the paths $q(t)$ are obey the boundary conditions $q(t'') = q''$, and $q(t') = q'$ and $S[q]$ denotes the classical action evaluated along the path $q(t)$:

$$S[q] = \int_{t'}^{t''} dt \left(\frac{M}{2}\dot{q}^2 - V[q(t)] \right). \tag{6.2}$$

Let us now consider the same system in equilibrium with a thermostat at temperature $1/\beta$ (we shall adopt a system of units in which $k_B = 1$). The starting point to define its thermodynamics is to define the partition function:

$$Z(\beta) = \text{Tr}e^{-\beta H} = \sum_n e^{-\beta E_n}, \tag{6.3}$$

where the trace is evaluated on a complete set of eigenvalues of $\hat{H}$,

$$\hat{H}|n\rangle = E_n|n\rangle. \tag{6.4}$$

Using the properties of the trace, we can represent the thermal operator $e^{-\beta H}$ on the basis set formed by the eigenvectors of the position operator:

$$Z(\beta) = \int dq \langle q|e^{-\beta H}|q\rangle. \tag{6.5}$$

Hence, the partition function can be represented as a periodic path integral in imaginary time, with $t'' - t' = -i\beta$:

$$Z(\beta) = \int dq\ K(q, -i\beta|q, 0) = \oint \mathcal{D}q(\tau)\ e^{-S_E[q]}, \tag{6.6}$$

where $S_E[q]$ is the Euclidean action:

$$S_E[q] = \int_0^\beta d\tau\ \left(\frac{M}{2}\dot{q}^2(\tau) + V[q(\tau)]\right). \tag{6.7}$$

and the symbol $\oint$ refers to the periodic boundary conditions $q(\beta) = q(0)$.

Let us now discuss the evaluation of thermal averages, in the form.

$$\langle A\rangle_\beta = \frac{1}{Z(\beta)}\text{Tr}[\hat{A}e^{-\beta\hat{H}}]. \tag{6.8}$$

In particular, we focus on the imaginary-time correlation functions, such as e.g. the two-point function:

$$\langle T_E[\hat{q}(-i\tau_1)\hat{q}(-i\tau_2)]\rangle_\beta \equiv \frac{1}{Z(\beta)}\text{Tr}[e^{-\beta\hat{H}}\hat{q}(-i\tau_1)\hat{q}(-i\tau_2)], \tag{6.9}$$

where $\hat{q}(-i\tau)$ is the position operator in the Heisemberg picture:

$$\begin{aligned} \hat{q}(t) &= e^{iHt}\ \hat{q}\ e^{-iHt} && (6.10)\\ \hat{q}(-i\tau) &= e^{H\tau}\ \hat{q}\ e^{-H\tau} && (6.11) \end{aligned}$$

and $T_E[\ldots]$ denotes the time-ordered product which operates in imaginary time.

In order to compute such correlation functions it is convenient to define a generating functional $Z(\beta, j)$, which reduces to the partition function $Z(\beta)$ for vanishing source $j(\tau)$:

$$Z(\beta, j) = \oint \mathcal{D}q(\tau)\ e^{-S_E[q]+\int_0^\beta d\tau q(\tau)j(\tau)} \tag{6.12}$$

Multiple functional differentiation of this functional yields arbitrary correlation functions. For example, the two-point imaginary time propagator reads:

$$\lim_{j(\tau)\to 0} \frac{1}{Z(\beta)} \frac{\delta^2 Z(\beta;j)}{\delta j(\tau_1)\delta j(\tau_2)} = \frac{1}{Z(\beta)} \oint \mathcal{D}q(\tau)\ q(\tau_1)q(\tau_2)\ e^{-S_E[q]}, \tag{6.13}$$

i.e. precisely the path integral representation of the the two-point correlation function. We note that, using the periodicity of the paths, it follows that

$$\langle T_E[\hat{q}(-i\beta)\hat{q}(-i\tau)]\rangle_\beta = \langle T_E[\hat{q}(0)\hat{q}(-i\tau)]\rangle_\beta \tag{6.14}$$

So far, the imaginary time variable τ has been defined in the strip $\tau \in [0,\beta]$. On the other hand, we shall see soon that it can be defined in the larger region $\tau \in [-\beta,\beta]$. Then it is possible to define the function

$$\Delta(\tau) = \langle T_E[\hat{q}(-i\tau)\hat{q}(0)]\rangle_\beta, \tag{6.15}$$

for any $\tau \in [0,\beta]$. Note that, from the Eq.s (6.10) and (6.11), using the cyclic properties of the trace, it follows that $\Delta(\tau)$ is a periodic function:

$$\Delta(\tau-\beta) = \langle T[\hat{q}(-i\tau)\hat{q}(\beta)]\rangle_\beta = \langle T[\hat{q}(-i\tau)\hat{q}(0)]\rangle_\beta = \Delta(\tau). \tag{6.16}$$

6.1.2 Real-time Quantum Propagators

Our ultimate goal is that of describing in a unified picture both thermodynamical quantities and real-time correlation functions. To this end, it proofs convenient to introduce a set different two-point functions which will be shown to encode information about to the spectral properties of the many-body system and the relaxation times. Hence, let us introduce the functions:

$$\begin{aligned} D^>(t,t') &= \langle\hat{q}(t)\hat{q}(t')\rangle_\beta && (6.17)\\ D^<(t,t') &= \langle\hat{q}(t')\hat{q}(t)\rangle_\beta, && (6.18)\end{aligned}$$

where, in the most general case, t and t' denote variables in the complex plane. Note that $D^<(t,t') = D^>(t',t)$ and that the products of field operators $\hat{q}(t)$ appearing in these function are not time-ordered.

A spectral representation of $D^>(t,t')$ may be obtained by inserting two resolutions of the identity:

$$D^>(t,t') = \frac{1}{Z(\beta)} \sum_{n,m} e^{-\beta E_n}\ e^{i\beta E_n(t-t')}\ e^{-i\beta E_m(t-t')} |\langle n|\hat{q}(0)|m\rangle|^2 \tag{6.19}$$

Such a representation makes it manifest that, $D^>(t,t')$ is an analytic function within the strip

$$-\beta \le \Im m(t-t') \le 0, \tag{6.20}$$

and becomes a distribution at the border of such an analytic region. Similarly, $D^>(t',t)$ is analytically defined in the strip

$$0 \le \Im m(t-t') \le \beta. \tag{6.21}$$

Now, using the cyclicity of the trace and the fact that the thermal operator $e^{-\beta \hat{H}}$ can be interpreted as an imaginary time evolution operator, we obtain the Kubo-Martin-Schwinger (KMS) relation

$$\begin{aligned} D^{>}(t,t') &= \mathrm{Tr}[e^{-\beta H} e^{iHt} \hat{q}(0) e^{-iHt} \hat{q}(t')] \\ &= \mathrm{Tr}[e^{iH(t+i\beta)} \hat{q}(0) e^{-iH(t+i\beta)} e^{-\beta H} \hat{q}(t')] \\ &= \mathrm{Tr}[e^{-\beta H} \hat{q}(t') \hat{q}(t+i\beta)] \\ &= D^{<}(t+i\beta, t'). \end{aligned} \tag{6.22}$$

We also note that, in the interval $[0,\beta]$ one has that the Euclidean time propagator defined in Eq. (6.15) is actually

$$\Delta(\tau) = D^{>}(-i\tau, 0) \tag{6.23}$$

Finally, for real values of t and t' we define the time-ordered propagator:

$$\langle T(\hat{q}(t)\hat{q}(t'))\rangle_\beta = \theta(t-t')\ D^{>}(t,t') + \theta(t'-t)\ D^{<}(t,t') \tag{6.24}$$

where $T[\ldots]$ denotes the time ordering in real time.

Time translational invariance implies that propagators $D^{>}(t,t')$ and $D^{<}(t,t')$ depend only on the time different $t'-t$ and not on t and t' individually. As a result, in the following we shall adopt the simplified notation

$$D^{>}(t) \equiv D^{>}(t,0), \qquad D^{<}(t) \equiv D^{<}(t,0). \tag{6.25}$$

It is also convenient to introduce their Fourier transform:

$$D^{>}(k_0) = \int_{-\infty}^{\infty} dt\ e^{ik_0t} D^{>}(t) \tag{6.26}$$

$$D^{<}(k_0) = \int_{-\infty}^{\infty} dt\ e^{ik_0t} D^{<}(t) = \int_{-\infty}^{\infty} dt\ e^{ik_0t} D^{>}(t-i\beta), \tag{6.27}$$

The last equality has been derived using the KMS condition. It can be shown that hermiticity of $\hat{q}(t)$ and time translational invariance imply thet $D^{>}(k_0)$ and $D^{<}(k_0)$ are real functions of k_0. From Eq. (6.18) one immediately has:

$$D^{<}(k_0) = D^{>}(-k_0) \tag{6.28}$$

In addition, from Eq.(6.27)

$$\begin{aligned} D^{<}(k_0) &= \int_{-\infty}^{\infty} dt\ e^{ik_0t} D^{>}(t-i\beta) \\ &= \int_{-\infty}^{\infty} dt\ e^{-\beta k_0}\ e^{ik_0(t-i\beta)} D^{>}(t-i\beta) \\ &= e^{-\beta k_0}\ D^{>}(k_0), \end{aligned} \tag{6.29}$$

It can be shown that this relationship expresses a detailed balance condition.

6.1.3 Spectral Function

Real time and imaginary propagators are entirely specified in terms of a quantity called the spectral function, which is defined as follows

$$\rho(k_0) \equiv D^>(k_0) - D^<(k_0) = \int_{-\infty}^{\infty} dt\; e^{ik_0 t}\; \langle [q(t), q(0)] \rangle_\beta \tag{6.30}$$

Using the detailed balance condition (6.29) and the definition of spectral function (6.30) one immediately finds an expression which relates directly the propagators $D^>(k_0)$ and $D^<(k_0)$ to the spectral function

$$\begin{aligned} D^>(k_0) &= (1 + f(k_0))\; \rho(k_0), & (6.31)\\ D^<(k_0) &= f(k_0)\; \rho(k_0), & (6.32)\end{aligned}$$

where $f(k_0) = (e^{\beta k_0} - 1)^{-1}$. Hence, the problem of computing the spectral function is in fact equivalent to that of determining $D^>(k_0)$ or $D^<(k_0)$.

Using the spectral representation (6.19) the definition of the spectral function (6.30) and the detailed balance condition (6.29) it is straightforward to work out a spectral representation for $\rho(k_0)$:

$$\rho(k_0) = \frac{2\pi}{Z(\beta)} \sum_{m,n} e^{-E_n\beta}\; \left(\delta(k_0 - E_n + E_m) - \delta(k_0 - E_m + E_n)\right) |\langle n|q(0)|m\rangle|^2 \tag{6.33}$$

This representation is quite useful as it explicitly shows that

- $\rho(k_0)$ is a real odd function: $\rho(-k_0) = -\rho(k_0)$.
- $\rho(k_0)$ is positive for positive k_0.

An useful sum rule for $\rho(k_0)$ can be deduced from the equal time commutation relationship for the field $\hat{q}(t)$:

$$\int_{-\infty}^{\infty} \frac{dk_0}{2\pi} k_0 \rho(k_0) = 1. \tag{6.34}$$

To prove this relationship, let us consider the canonical commutation relationship:

$$[\hat{q}(t), \frac{d}{dt}\hat{q}(t')]_{t=t'} = i \tag{6.35}$$

But then

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dk_0}{2\pi}\; k_0\; e^{-ik_0 t} \left(D^>(k_0) - D^<(k_0)\right) &= i\frac{d}{dt}(D^>(t) - D^<(t)) \\ &= i \lim_{t\to t'} \frac{d}{dt}\langle [q(t'), q(t)] \rangle = 1 \end{aligned} \tag{6.36}$$

Finally, it is instructive to consider the case of the harmonic oscillator, $V = \frac{1}{2}\omega^2 q^2$. In this case, the position operator can be written as:

$$\hat{q} = \frac{1}{\sqrt{2\omega}} \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}\right) \tag{6.37}$$

Then the combination leading to the spectral function is:

$$D^>(t) - D^<(t) = \frac{1}{2\omega}\langle [a, a^\dagger] e^{-i\omega t} + [a^\dagger, a] e^{i\omega t}\rangle_\beta \tag{6.38}$$

Using $[a, a^\dagger] = 1$

$$D^>(t) - D^<(t) = \frac{1}{2\omega}\langle e^{-i\omega t} + e^{i\omega t}\rangle_\beta \tag{6.39}$$

then

$$\rho(k_0) = \frac{2\pi}{2\omega}(\delta(\omega - k_0) + \delta(\omega + k_0)) = 2\pi\epsilon(k_0)\ \delta(k_0^2 - \omega^2) \tag{6.40}$$

This expression is called the free-spectral density. The name is due to the fact that a free relativistic quantum field theory corresponds to a collection of harmonic oscillators.

6.1.4 Matsubara Propagator

We now consider the Fourier transform of the imaginary-time propagator $\Delta(\tau)$. Because of its periodicity, we can restrict the time integration to the first cell $[0, \beta]$ and the Fourier integral reduces to a sum:

$$\begin{aligned} \Delta(i\omega_n) &\equiv \int_0^\beta e^{i\omega_n \tau}\Delta(\tau), & (6.41)\\ \Delta(\tau) &= \frac{1}{\beta}\sum_n e^{-i\omega_n\tau}\Delta(i\omega_n) & (6.42)\end{aligned}$$

where the Fourier frequencies are given by $\omega_n = \frac{2\pi}{\beta}n$ and will be referred to as Matsubara frequencies.

We now seek for a spectral representation of $\Delta(i\omega_n)$. To this end, we recall that, if τ lies in the interval $[0, \beta]$, then

$$\Delta(\tau) = D^>(-i\tau) = \int_{-\infty}^{\infty}\frac{dk_0}{2\pi}e^{-ik_0(-i\tau)}D^>(k_0) = \int_{-\infty}^{\infty}\frac{dk_0}{2\pi}e^{-k_0\tau}D^>(k_0) \tag{6.43}$$

Then using the relationship (6.31) between $D^>(k_0)$ and $\rho(k_0)$ we find:

$$\begin{aligned} \Delta(i\omega_n) &\equiv \int_0^\beta e^{i\omega_n\tau}\int_{-\infty}^{\infty}\frac{dk_0}{2\pi}e^{-ik_0(-i\tau)}(1 + f(k_0))D^>(k_0) & (6.44)\\ &= -\int_{-\infty}^{\infty}\frac{dk_0}{2\pi}\frac{\rho(k_0)}{i\omega_n - k_0}. & (6.45)\end{aligned}$$

Let us return to real-time propagators. From the basic definitions, it is immediate to show that the time-ordered propagators obey:

$$D(k_0) = \int dt e^{ik_0 t}\left(\theta(t)D^>(t) + \theta(-t)D^<(t)\right) \tag{6.46}$$

Furthermore, it is useful to introduce the retarded and advanced propagators:

$$\begin{aligned} D_R(t) &= \langle\theta(t)[\hat{q}(t), \hat{q}(0)]\rangle_\beta & (6.47)\\ D_A(t) &= -\langle\theta(-t)[\hat{q}(t), \hat{q}(0)]\rangle_\beta. & (6.48)\end{aligned}$$

Using the integral representation of the Heaviside funciton:

$$\theta(t) = i\int_{-\infty}^{\infty} \frac{dk_0'}{2\pi} \frac{e^{-ik_0 t}}{k_0' + i\eta} \tag{6.49}$$

Then one immediately finds:

$$\begin{aligned} D_R(k_0) &= i\int_{-\infty}^{\infty} \frac{dk_0'}{2\pi} \frac{\rho(k_0')}{k_0 - k_0' + i\eta} &(6.50)\\ D_A(k_0) &= -i\int_{-\infty}^{\infty} \frac{dk_0'}{2\pi} \frac{\rho(k_0')}{k_0 - k_0' - i\eta} &(6.51)\end{aligned}$$

Thus

$$\begin{aligned} D_R(k_0) &= -i\Delta(k_0 + i\eta) &(6.52)\\ D_A(k_0) &= i\Delta(k_0 - i\eta). &(6.53)\end{aligned}$$

6.2 Non-Equilibrium Quantum Dynamics

In this section we extend the discussion of the irreversible dynamics in open systems to the quantum level, starting from an arbitrary initial condition, which may represent an non-equilibrium state. In particular we upgrade to the quantum level the same model discussed at the classical level in section 3.1.1, defined by the quantum Hamiltonian:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{U}(x) + \sum_\alpha \left[\frac{\hat{p}_\alpha^2}{2\mu_\alpha} + \frac{\mu_\alpha\omega_\alpha^2}{2}\left(\hat{q}_\alpha - \frac{c_\alpha}{\mu_\alpha\omega_\alpha^2}\hat{x}\right)^2\right] \tag{6.54}$$

Our goal is to evaluate the evolution of the **reduced density matrix** for the system's variable, which is obtained by tracing out the bath from the full density matrix:

$$\rho(x,y,t) \equiv \sum_\alpha \int dq_\alpha \langle x, q_\alpha|\hat{\rho}(t)|y, q_\alpha\rangle \tag{6.55}$$

where $|x\rangle$ and $|y\rangle$ are system's position eigenstates while $|q_\alpha\rangle$ are bath's position eigenstates.

We consider starting from some specified initial condition. For example,

$$\hat{\rho}(0) = |x_0\rangle\langle x_0| \otimes e^{-\beta \hat{H}_B} \tag{6.56}$$

where $\hat{H}_B = \sum_\alpha \left[\frac{\hat{p}_\alpha^2}{2\mu_\alpha} + \frac{\mu_\alpha\omega_\alpha^2}{2}\left(\hat{q}_\alpha - \frac{c_\alpha}{\mu_\alpha\omega_\alpha^2}x_0\right)^2\right]$ is the bath Hamiltonian. Then the reduced density matrix is written as follows:

$$\rho(x,y,t|x_0,x_0) = \sum_\alpha \int dq_\alpha \langle x, q_\alpha| e^{-\frac{i}{\hbar}\hat{H}t}\ \hat{\rho}(0)\ e^{\frac{i}{\hbar}\hat{H}t}|y, q_\alpha\rangle \tag{6.57}$$

In order to describe the dynamics of the reduced density matrix we adopt the path integral representation. To build the path integral we need to use Trotter formula to represent the forward and backward time evolution operators entering

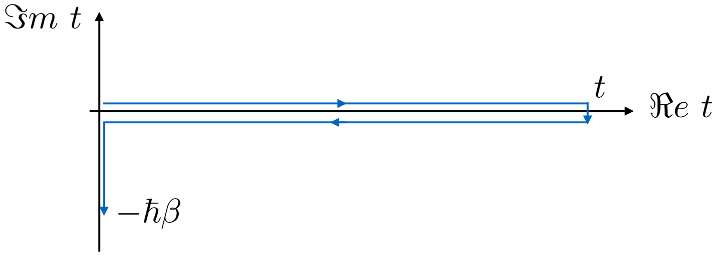

Figure 6.1: The Keldysh-Schwinger contour $\mathcal{C}$

expression (6.57). In addition we need to represent also the initial equilibrium density matrix for the bath variables, which gives raise to an imaginary-time path integral. The result is

$$\begin{aligned}\rho(x,y,t|x_0,x_0) \quad = \quad & \int dQ \int d\bar{Q}_i \int dQ_i \int_{\bar{Q}_i}^{Q} \mathcal{D}Q'_\alpha \int_{Q}^{Q_i} \mathcal{D}Q''_\alpha \int_{Q_i}^{\bar{Q}_i} \mathcal{D}Q_E \int_{x_0}^{x} \mathcal{D}x' \int_{y}^{x_0} \mathcal{D}x'' \\ & e^{\frac{i}{\hbar}(S_S[x']+S_{SB}[x'.Q'])}\ e^{-\frac{i}{\hbar}(S_S[x'']+S_{SB}[x'',Q''])} e^{-S_E[Q_E]}\end{aligned} \tag{6.58}$$

where we have collectively denoted with Q the set of coordinates of all harmonic oscillators and the action functionals in the exponent are:

$$S_E[Q] \quad = \quad \int_0^{\hbar\beta} d\tau \sum_\alpha \left[\frac{\dot{q}_\alpha^2}{2}\mu_\alpha + \frac{1}{2}\mu_\alpha\omega_\alpha^2 \, \left(q_\alpha - \frac{c_\alpha}{\mu_\alpha\omega_\alpha^2} x_0 \right)^2 \right] \tag{6.59}$$

$$S_B[Q,X] \quad = \quad \int_0^{t} d\tau \sum_\alpha \left[\frac{\dot{q}_\alpha^2}{2}\mu_\alpha - \frac{1}{2}\mu_\alpha\omega_\alpha^2 \, \left(q_\alpha - \frac{c_\alpha}{\mu_\alpha\omega_\alpha^2} x \right)^2 \right] \tag{6.60}$$

$$S_S[x] \quad = \quad \int_0^{t} d\tau \left[\frac{\dot{x}^2}{2M} - V(x) \right] \tag{6.61}$$

It is convenient to introduce a notation which enables one to collectively represent path integrals in the forward- backward- and imaginary time direction: To this end, we introduce the so called Keldysh-Schwinger time contour represented in Fig. and represent by $\int_{\mathcal{C}} d\tau$ the integral over the curvilinear abscissa which parametrizes it (see Fig. 3.3). With this notation, the path integral representation of the density matrix reads:

$$\rho(x,y,t|x_0,x_0) \quad = \quad \int_{\mathcal{C}} \mathcal{D}Q \int_{\mathcal{C}} \mathcal{D}X \ e^{\frac{i}{\hbar}(S_{\mathcal{C}}[x]+S_{\mathcal{C}}[Q,x])} \tag{6.62}$$

All path integrals over bath variables are of Gaussian type, so can be carried out analytically. The result is an effective path integral representation in terms of system's variables only. Returning provisorily to a notation with explicit distinction between forward- and backward- propagating paths, we have

$$\rho\,(x,y,t|x_0,x_0) \quad = \quad \int_{x_0}^{x} \mathcal{D}x' \int_{y}^{x_0} \mathcal{D}x'' \ e^{\frac{i}{\hbar}(S_S[x']-S_S[x'']+i\hbar\Phi[x',x''])}, \tag{6.63}$$

where $\Phi[x', x'']$ is the so called Feynman-Vernon **influence functional** and reads:

$$\begin{aligned}\Phi[x', x''] &= \int_0^t dt' \int_0^{t'} dt'' \left\{ \, (x'(t') - x''(t')) \cdot [B(t'-t'')x'(t'') - B^*(t'-t'')x''(t'')] \right\} \\ &+ \frac{i\bar{\mu}}{2} \int_0^t dt' [x'^2(t') - x''^2(t')], \qquad \left(\bar{\mu} = \frac{c_\alpha^2}{\mu_\alpha \omega_\alpha^2}\right) \end{aligned} \tag{6.64}$$

$\mathcal{B}(t)$ is a Green's function which encodes the fluctuation-dissipation induced by the heat bath and reads:

$$\mathcal{B}(t) = \sum_\alpha \frac{c_\alpha^2}{2\mu_\alpha \omega_\alpha} \left[\coth\left(\frac{\omega_\alpha \hbar}{2k_B T}\right) \cos(\omega_\alpha t) - i \, \sin(\omega_\alpha t) \right]. \tag{6.65}$$

In complete analogy with the classical discussion, let us now consider the Ohmic limit, in which the $\mathcal{B}(t)$ reduces to:

$$\mathcal{B}(t) = C_1 \delta(t) + i \; C_2 \; \frac{d}{dt}\delta(t) + \ldots \tag{6.66}$$

The coefficient C_1 and C_2 are insofar unspecified parameters. However they can be fixed by invoking the correspondence principle, i.e. requiring that , in the classical limit, the time-evolution of the density matrix returns the prediction of classical Langevin dynamics. In the following, we shall show that this choice is guaranteed if

$$C_1 = 2M\gamma k_B T/\hbar \qquad C_2 = M\gamma \tag{6.67}$$

At this point it is convenient to change variables in the path integrals and introduce the so-called classical and fluctuation paths, defined as following:

$$R(\tau) = \frac{x'(\tau) + x''(\tau)}{2} \tag{6.68}$$

$$Y(\tau) = x'(\tau) - x''(\tau) \tag{6.69}$$

we emphasize that the second term parametrizes the difference between forward and backward propagation. In terms of these functions, the Feynman-Vernon path integral reads

$$\rho\left(R_f + \frac{Y_f}{2}, R_f - \frac{Y_f}{2}, t | x_0, x_0\right) \;=\; \int_{x_0}^{x} \mathcal{D}R \int_0^0 \mathcal{D}Y \; e^{\frac{i}{\hbar}(\mathcal{W}[R,y] + \Phi'[R,y])} \tag{6.70}$$

where $R_f = \frac{x+y}{2}$, $Y_f = x - y$, $R(0) = x_0$, $Y(0) = 0$ and

$$\begin{aligned}\mathcal{W}[R, Y] &= \int_0^t d\tau \left\{ M \; \dot{R} \cdot \dot{y} - V\left(R + \frac{Y}{2}\right) + V\left(R - \frac{Y}{2}\right) \right\} \\ \Phi'[R, Y] &= \int_0^t dt' \left\{ i \frac{M\gamma k_B T}{\hbar} Y^2 + M\gamma \dot{R} \cdot Y \right\} \end{aligned} \tag{6.71}$$

Eq. (6.70) with Eq.s (6.71) and (6.71) are the quantum generalisation of the path integral representation of the classical Fokker planck dynamics given in Eq. (3.76).

An important observation is that the functional at the exponent contains a term proportional to $-\frac{Y^2}{\hbar}$. Thus, in the classical limit, fluctuations of the Y field around 0 are expected to be exponentially small. Based on this consideration, we can expand the exponent to quadratic order in Y^2 and analytically perform the resulting Gaussian integral over the Y field using the standard technique discussed in Chapter 1. Note also that, in the classical limit, the non-diagonal elements in the density matrix are suppressed. The final expression for the diagonal elements , which represent the probability of observing the particle at position R at time t is $P(R_f, t, x_0) = \rho_R(R_f, R_f; t|x_0, x_0)$:

$$P(R_f, t, x_0) = \mathcal{N} \int_{x_0}^{x_f} \mathcal{D}x e^{-\frac{\beta}{4M\gamma}\int_{t_0}^{t} d\tau (M\ddot{x}(\tau) + M\gamma\dot{x}(\tau) + \nabla U[x(\tau)])^2} \tag{6.72}$$

hence, we have recovered the familiar result for a classical underdamped Langevin dynamics.

6.3 The Quantum Master Equation and its Markovian Limit

In the previous section we have discussed the path integral representation of the time-dependent density matrix, which provides the quantum mechanical qeneralisation of the conditional probability generated by the Langevin equation. In the previous chapter, we have shown that for a classical particle obeying the over damped Langevin equation (Markovian limit) the conditional probability obeys the SFP equation (3.91). It is then natural to ask the question: what is the quantum counterpart of such an equation?

We recall that the total density operator of the system and heat bath obeys the Liouville-von Neumann equation:

$$i\hbar\dot{\hat{\rho}}(t) = [\hat{H}_{tot}, \hat{\rho}(t)], \tag{6.73}$$

which is equivalent to a Schrödinger equation.

The system's reduced matrix, which we recall is defined as the trace of $\hat{\rho}$ over the bath degrees of freedom, does no longer obey Eq. (6.73). It can be shown that, in general, the trace over the heat-bath generates an additional *non-unitary* term:

$$i\hbar\frac{\partial}{\partial t}\hat{\rho}(t) = [\hat{H}_{tot}, \hat{\rho}(t)] + i\hat{\mathcal{D}}\hat{\rho}(t) \tag{6.74}$$

which is called the (Quantum) Master Equation. In this equation, $\hat{\mathcal{D}}$ is a super-operator which provides non-unitary character to the time-evolution of the density matrix.

Lindblad showed that in the Markovian limit (which is equivalent to the over damped limit for the Langevin equation) the quantum Master Equation takes the general form:

$$\dot{\hat{\rho}} = [\hat{H}_{tot}, \hat{\rho}(t)] + \frac{1}{2\hbar}\sum_{\mu}\left([\hat{L}_\mu\hat{\rho}, \hat{L}_\mu^\dagger] + \left[\hat{L}_\mu, \hat{\rho}\hat{L}_\mu^\dagger\right]\ \right) \tag{6.75}$$

where $\{\hat{L}_\mu\}_\mu$ denotes a generic set of operators in the Hilbert space. Hence, in this limit, the super-operator $\hat{\mathcal{D}}$ reads

$$\hat{\mathcal{D}}\rho = \frac{1}{2\hbar}\sum_{\mu}\left([\hat{L}_\mu\hat{\rho}, \hat{L}^\dagger] + [\hat{L}_\mu, \hat{\rho}\hat{L}^\dagger]\right) \tag{6.76}$$

This equation is referred to as the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation or sometimes simply the Linblad equation. It is often used to investigate the effect of decoherence and dissipation at the quantum level.

Part II

Statistical Fields

Chapter 7

Non-Relativistic Quantum and Statistical Fields

Let us consider the problem of computing *quantum* canonical averages of the system described by the Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2M} + U(\hat{q})$. If this system is coupled to a thermostat at temperature $1/\beta$ (we shall adopt a system of units in which $k_B = 1$) it reaches a state of thermal equilibrium, whose properties can be entirely deduced from its (quantum) partition function.

$$Z(\beta) = \mathrm{Tr} e^{-\beta H} = \sum_n e^{-\beta E_n}, \tag{7.1}$$

where the trace is evaluated on a complete set of eigenvalues of $\hat{H}$,

$$\hat{H}|n\rangle = E_n|n\rangle. \tag{7.2}$$

Using the properties of the trace, we can represent the thermal density matrix operator $e^{-\beta H}$ on the basis set formed by the eigenvectors of the position operator:

$$Z(\beta) = \int dy \langle y|e^{-\beta H}|y\rangle. \tag{7.3}$$

Hence, the partition function can be represented as a periodic path integral in imaginary time, with $t'' - t' = -i\beta$:

$$Z(\beta) = \int dy\ K(y, -i\beta|y) = \oint \mathcal{D}y(\tau)\ e^{-S_E[y]}, \tag{7.4}$$

where $S_E[q]$ is the systems' Euclidean action and the symbol $\oint$ refers to the periodic boundary conditions $q(\beta) = q(0)$. Thus, the quantum canonical ensemble is mapped into the classical partition function of "ring polymers", with length set by the inverse temperature β. This mathematical statement illustrate that the cost of upgrading the statistical mechanics of a bosonic system from the classical to the quantum level is to introduce into the partition function an infinite set of coupled copies of the classical system, one for each incremental imaginary time step.

7.1 Non-Relativistic Bosonic Quantum Fields

This section contains a brief introduction to second-quantized formulation of non-relativistic quantum mechanics of many-body systems. An obvious application of this formalism is the theoretical discussion of quantum many particle systems. In this course, however, we are mostly concerned with two alternative applications of field theory: (i) modelling of the equilibrium statistical classical mechanics of polymers and (ii) modelling the irreversible non-equilibrium dynamics of *isolated* quantum excitations, propagating in an open environment.

The first application (to be discussed in the next chapter) is based on exploiting the formal duality between quantum dynamics in imaginary time and classical equilibrium statistical mechanics, which was established in the previous section. The second application (to be discussed e.g. in [6, 7, 8]) concerns the evolution of single-body excitations, which are obviously insensitive to Pauli principle. Thus, both such applications can be developed by considering only bosonic degrees of freedom. The reader interested to the generalization to fermionic fields is referred, e.g. to Negele-Orland's classic textbook [5].

N-particle bosonic states: In the non-relativistic limit, pair-creation annihilation effects are suppressed and the number of particles is conserved. Consequently, the Hilbert space $\mathcal{H}$ for a generic N-particle quantum system can constructed combining N single-particle Hilbert spaces: Namely, if $|\lambda\rangle$ represents a complete orthogonal basis set for the single particle space, then the $N-$ particle states

$$|\lambda_1, \ldots, \lambda_N\} \equiv |\lambda_1\rangle \otimes \ldots \otimes |\lambda_N\rangle \tag{7.5}$$

provide a complete orthonormal basis for the N-particle Hilbert space H_N. Notice that, following [5], we have chosen to the note the "ket" symbol by a curly bracket. This is to keep track that the states defined in (7.6) do not obey Bose symmetry, hence cannot represent physical bosonic states. A symmetrized version of the the N-boson state is denoted as

$$|\lambda_1, \ldots, \lambda_N\rangle \equiv \sum_{\text{perm.}} \mathcal{P}\left[|\lambda_1\rangle \otimes \ldots \otimes |\lambda_N\rangle\right] \tag{7.6}$$

where the sum is over all permutations of particle indexes, thus establishes Bose symmetry.

The bosonic Fock space is defined as the direct sum of infinitely many of such Hilbert spaces, each one corresponding to a different number of particles:

$$\mathcal{F} \equiv H_0 \oplus H_1 \oplus H_2 \ldots = \sum\nolimits_{\oplus\ k=0}^{\infty} H_k \tag{7.7}$$

The boson **creation operators** are defined by:

$$\hat{a}^\dagger_\lambda |\lambda_1, \ldots, \lambda_N\} = |\lambda, \lambda_1, \ldots, \lambda_N\}, \tag{7.8}$$

Equivalently, using symmetrized basis states, one has:

$$\hat{a}^\dagger_\lambda |\lambda_1, \ldots, \lambda_N\rangle = \sqrt{n_\lambda + 1}|\lambda, \lambda_1, \ldots, \lambda_N\rangle \tag{7.9}$$

where n_λ is the occupation number of the quantum state with quantum number λ. The **vacuum state** is defined by

$$\hat{a}^\dagger_\lambda|0\rangle = |\lambda\rangle. \tag{7.10}$$

It is immediate to verify that Bose symmetry implies the commutation relationship:

$$[\hat{a}^\dagger_\lambda, \hat{a}^\dagger_{\lambda'}] = 0. \tag{7.11}$$

The unsymmetrized and symmetrized multi-boson state may be generated by acting with the creation operator on the vacuum:

$$\begin{aligned} |\lambda_1, \ldots, \lambda_N\} &= \hat{a}^\dagger_{\lambda_1} \ldots \hat{a}^\dagger_{\lambda_N}|0\rangle && (7.12)\\ |\lambda_1, \ldots, \lambda_N\rangle &= \frac{1}{\sqrt{\prod_{k=1}^N n_{\lambda_k}!}}\hat{a}^\dagger_{\lambda_1} \ldots \hat{a}^\dagger_{\lambda_N}|0\rangle && (7.13)\\ & && (7.14) \end{aligned}$$

The creation operators are not self-adjoint. Hence, **annihilation operators** can be defined as $\hat{a}_\lambda = (a^\dagger_\lambda)^\dagger$. By taking the adjoint of Eq. (7.8), it is immediate to prove that:

$$\langle\alpha_1, \ldots, \alpha_m|\hat{a}_\lambda|\beta_1, \ldots, \beta_n\rangle = \langle\lambda,\ \alpha_1, \ldots, \alpha_m|\beta_1, \ldots, \beta_n\rangle. \tag{7.15}$$

The right-hand-side of this equation is 0 unless $m + 1 = n$. Hence, the annihilation operator reduces the number of particles in a space. In particular, when acting on the vacuum state we have:

$$\hat{a}_\lambda|0\rangle = 0 \qquad \langle 0|a^\dagger_\lambda = 0 \qquad (\forall\ \lambda). \tag{7.16}$$

To close the algebra of these operators we need to evaluate $[a^\dagger_\lambda, a^\dagger_\xi]$ and $[a_\lambda, a^\dagger_\xi]$. By direct applying $\hat{a}$ and $\hat{a}^\dagger$ on a generic multi-boson state, it is a trivial exercise to obtain

$$\begin{aligned} [\hat{a}_\lambda, \hat{a}_\xi] &= 0 && (7.17)\\ [\hat{a}_\lambda, \hat{a}^\dagger_\xi] &= \delta_{\lambda\xi} && (7.18) \end{aligned}$$

Finally, we define the **normal ordering** of an arbitrary sequence of creation and annihilation operators as the rearrangement which moves all the annihilation operators to the rightmost side, for example

$$: a_\lambda a^\dagger_\xi a_\mu a_\nu a^\dagger_\sigma := a^\dagger_\xi a^\dagger_\sigma a_\lambda a_\mu a_\nu \tag{7.19}$$

If the number of particles in a system is held fixed to N, a **generic quantum state** can be written as:

$$|\Psi\rangle = \int d^3x_1 \ldots \int d^3x_N\ \psi(x_1, \ldots, x_N)\ |x_1, \ldots, x_N\rangle \tag{7.20}$$

where $\psi(x_1, \ldots, x_N)$ is the wave-function projected onto the coordinate basis, since

$$\langle x_1, \ldots, x_N|\Psi\rangle = \psi(x_1, \ldots, x_N). \tag{7.21}$$

Note also that

$$\langle p_1,\ldots,p_N|\Psi\rangle = \int d^3x_1 \ldots \int d^3x_N \; e^{i\sum_i p_i\cdot x_i} \; \psi(x_1,\ldots,x_N). \tag{7.22}$$

On the other hand, If the number of particles in a system is not fixed (e.g. in relativistic quantum systems) a generic state $|\psi\rangle$ has to be projected onto Fock spaces built out of different particle numbers:

$$\begin{aligned} |\Psi\rangle = \int d^3x_1 c_1(x_1)|x_1\rangle + \int dx_1 dx_2 c_2(x_1,x_2)|x_1x_2\rangle \\ + \int dx_1 dx_2 dx_3 c_2(x_1,x_2,x_3)|x_1x_2x_3\rangle + \ldots \end{aligned} \tag{7.23}$$

Given the creation and annihilation operators associated to a given complete orthonormal basis $|\lambda\rangle$, it is immediate to obtain the creation operators associated to a different basis, $|\alpha\rangle$:

$$\hat{a}^\dagger_\alpha = \sum_\lambda \langle\lambda|\alpha\rangle \; \hat{a}^\dagger_\lambda \tag{7.24}$$

$$\hat{a}_\alpha = \sum_\lambda \langle\alpha|\lambda\rangle \; \hat{a}_\lambda. \tag{7.25}$$

In particular, a very useful basis set is that which combines position eigenstates x and possibly other internal discrete quantum numbers (e.g. spin and iso-spin); $|X\rangle \equiv |x\sigma\tau\rangle$. In this case, the creation operators are traditionally denoted with $\hat{\psi}(X)$ and $\hat{\psi}^\dagger(X)$ and are called the **field operators**. The general expansion of field operators on another basis set $|\alpha\rangle$ is given by:

$$\hat{\psi}^\dagger(X) = \sum_\lambda \; \langle\lambda|X\rangle \; \hat{a}^\dagger_\lambda \tag{7.26}$$

$$\hat{\psi}(X) = \sum_\lambda \; \langle X|\lambda\rangle \; \hat{a}_\lambda \tag{7.27}$$

For example, for a continuous system of spin-0 boson with no other internal degree of freedom, choosing $|\lambda\rangle$ to be the momentum eigenstates basis (i.e. eigenstates of the free Hamiltonian), then we have:

$$\hat{\psi}^\dagger(x) = \int \frac{d^3p}{(2\pi)^3} e^{ip\cdot x} \; \hat{a}^\dagger_p \tag{7.28}$$

$$\hat{\psi}(x) = \int \frac{d^3p}{(2\pi)^3} e^{-ip\cdot x} \; \hat{a}_p. \tag{7.29}$$

i.e. the non-relativistic limit of the field operators routinely introduced in the description of high-energy scattering, in which anti-particle degrees of freedom are neglected.

A convenient way to represent a generic operator in terms of field operators consists in first giving its expression in a basis in which it is diagonal, and then transform to the position eigenstate basis. For example the kinetic energy operator is diagonal in the momentum basis:

$$\hat{T} = \int \frac{d^3p}{(2\pi)^3} \hat{\psi}^\dagger(p) \; \left(\frac{p^2}{2m}\right) \; \hat{\psi}(p) \tag{7.30}$$

hence, in coordinate space we have

$$\hat{T} = \int d^3x \; \hat{\psi}^\dagger(x) \left(\frac{-\nabla^2}{2m} \right) \; \hat{\psi}(x). \tag{7.31}$$

We conclude this section with an illustrative application. We show that when the following normal-ordered field-theoretic Hamiltonian

$$\hat{H} = \int d^3x \; \hat{\psi}^\dagger(x) \frac{-\nabla^2}{2m} \hat{\psi}(x) + \int d^3x \int d^3y \; U(x,y) \;\; : \hat{\psi}^\dagger(x)\hat{\psi}(x) \; \hat{\psi}^\dagger(y)\hat{\psi}(y) : . \tag{7.32}$$

acts on the N particle Hilbert space H_N, it reduces to the first quantised Hamiltonian

$$\hat{H} = \sum_i \frac{-\nabla_i^2}{2m} + \sum_{i<j} U(x_i, x_j). \tag{7.33}$$

To show this result, first the kinetic energy operator projected onto the position basis:

$$\begin{aligned}
\langle x_1, \dots, x_N | \hat{T} | \Psi \rangle &= \langle x_1, \dots, x_N | \int \frac{d^3p}{(2\pi)^3} \; \frac{p^2}{2m} \hat{\psi}_p^\dagger \hat{\psi}_p \\
&\quad \cdot \left[\prod_l \int \frac{d^3k_l}{(2\pi)^3} \right] \psi(k_1, \dots, k_N) | k_1, \dots, k_N \rangle \\
&= \int \frac{d^3p}{(2\pi)^3} \; \left(\frac{p^2}{2m} \; \sum_{i=1}^N \delta(p - k_i) \right) \langle x_1, \dots, x_N | \\
&\quad \cdot \left[\prod_l \int \frac{d^3k_l}{(2\pi)^3} \right] \psi(k_1, \dots, k_N) | k_1, \dots, k_N \rangle \\
&= \left[\prod_l \int \frac{d^3k_l}{(2\pi)^3} \right] \left(\sum_{i=1}^N \frac{k_i^2}{2m} \right) \psi(k_1, \dots, k_N) \; e^{i \sum_m p_m \cdot x_m} \\
&= \sum_{i=1}^N \frac{-\nabla_i^2}{2m} \; \psi(x_1, \dots, x_N).
\end{aligned} \tag{7.34}$$

The potential energy operator is diagonal in the position basis, hence:

$$\begin{aligned}
\langle x_1, \dots, x_N | \hat{U} | \Psi \rangle &= \frac{1}{2} \int d^3z \int d^3y \langle x_1, \dots, x_N | : \hat{\psi}_y^\dagger \hat{\psi}_y \; \hat{\psi}_z^\dagger \hat{\psi}_z : | \Psi \rangle \; U(z - y) \\
&= \sum_{i<j=1}^N U(x_i, x_j) \psi(x_1, \dots, x_N).
\end{aligned} \tag{7.35}$$

Note that, without normal ordering, this sum would contain also N self interacting terms with $i = j$. For many potentials, these contribution would lead to an infinite energy.

7.2 Path Integral on Coherent Fields

Coherent states provide a convenient complete basis for the Fock space. Indeed, they display the useful feature that each coherent state spans the entire Fock space, i.e. contains components of with an arbitrary number of particles. Coherent states are defined as the eigenstates of the annihilation operator,

$$\hat{a}_\alpha |\phi\rangle = \phi_a \; |\phi\rangle \tag{7.36}$$

where the eigenvalue ϕ_α is in l a complex number.

It is easy to verify that

$$\begin{aligned} |\phi\rangle &= e^{\sum_\alpha \phi_\alpha \hat{a}^\dagger_\alpha} |0\rangle, & (7.37)\\ \langle\phi| &= \langle 0| e^{\sum_\alpha \phi^*_\alpha \hat{a}_\alpha}, & (7.38)\end{aligned}$$

provide a constructive definition of coherent states (and of their duals).

Some important properties follow immediately from the definition of these states: First, we have that

$$\langle\phi|\hat{a}^\dagger_\alpha = \langle\phi|\phi^*_\alpha \tag{7.39}$$

Next, we observe

$$\hat{a}^\dagger_\alpha |\phi\rangle = \hat{a}^\dagger_\alpha \; e^{\sum_\alpha \phi_\alpha a^\dagger_\alpha} \; |0\rangle = \frac{\partial}{\partial \phi_\alpha} \; |\phi\rangle \tag{7.40}$$

and its adjoint relation

$$\langle\phi|\hat{a}_\alpha = \frac{\partial}{\partial \phi^*_\alpha} \; \langle\phi| \tag{7.41}$$

The orthogonality condition on the states $|\alpha\rangle$ created by $a^\dagger_\alpha$ implies

$$\langle\phi|\phi'\rangle = e^{\sum_\alpha \phi^*_\alpha \; \phi'_\alpha}, \tag{7.42}$$

Thus, coherent states are not orthonormal!

A crucial property of the coherent states is their over completeness in the Fock space. This implies that any vector in the Fock space can be expanded in coherent states. On the other hand, the over completeness implies that the resolution of the identity reads

$$1 = \int \prod_\alpha \frac{d\phi_\alpha d\phi^*_\alpha}{2\pi i} \; e^{-\sum_\alpha \phi^*_\alpha \phi_\alpha} |\phi\rangle\langle\phi| \tag{7.43}$$

The proof of this relation is straightforward and can be found e.g. in Ref. [5], where the integration measure in (7.43) is defined by the Jacobian hence reads

$$\frac{d\phi_\alpha d\phi^*_\alpha}{2\pi i} = \frac{d(\Re e\phi_\alpha) \; d(\Im m\phi_\alpha)}{\pi} \tag{7.44}$$

and the integration runs over the real and complex axes.

A quantum state $|\psi\rangle$ can be represented in the coherent basis as follows

$$|\psi\rangle = \int \prod_{\alpha} \frac{d\phi_\alpha d\phi^*_\alpha}{2\pi i} \; e^{-\sum_\alpha \phi^*_\alpha \phi_\alpha} |\phi\rangle\langle\phi|\psi\rangle \tag{7.45}$$

where ϕ denotes the set of all component $\phi_1, \ldots, \phi_N$.

A powerful property of coherent states which will be used shortly is that matrix elements of normal-ordered operators between coherent states take a particularly simple form: denoting with $A(\hat{a}^\dagger_\alpha, \hat{a}_\beta)$ a normal-ordered operator, then:

$$\langle\phi|A(\hat{a}^\dagger_\alpha, \hat{a}_\beta)|\phi'\rangle = A(\phi^*_\alpha, \phi'_\beta) \; e^{\sum_l \phi^*_l \phi'_l}. \tag{7.46}$$

For example, if $A(\hat{a}^\dagger_\alpha, \hat{a}_\beta)$ then

$$\langle\phi|\hat{a}^\dagger_\alpha \hat{a}_\beta|\phi'\rangle = \phi^*_\alpha \phi'_\beta \; \langle\phi|\phi'\rangle = \phi^*_\alpha \phi'_\beta \; e^{\sum_l \phi^*_l \phi'_l}. \tag{7.47}$$

We now use the coherent state basis to construct another equivalent representation of the path integral, in which the integration is carried over the configurations of complex classical fields which represent the eigenvalues of the coherent states.

The eigenvalues of the coherent states are then complex functions of the coordinates (i.e. complex classical fields):

$$\hat{\psi}(x)|\phi\rangle = \phi(x)|\phi\rangle, \tag{7.48}$$

where we have $\hat{\psi}(x) \equiv \hat{a}_x$ are the quantum fields. As in the coordinate space path integral, the starting point to represent the evolution operator is the standard Trotter decomposition:

$$K(x_f, t; x_i) = \langle 0|\hat{\psi}(x_f) e^{-\frac{i}{\hbar} t\hat{H}} \hat{\psi}^\dagger(x_i)|0\rangle = \langle 0| \; \hat{\psi}(x_f) \; \left(e^{-\frac{i}{\hbar}\Delta t \hat{H}}\right)^{N_t} \; \hat{\psi}^\dagger(x_i)|0\rangle. \tag{7.49}$$

where $\Delta t = t/N_t$.

Next, we introduce an infinite set of completeness relations between each matrix elements, based on the coherent states associated to each point in space:

$$\langle 0| \; \hat{\psi}(x_f) \; \left(e^{-\frac{i}{\hbar}\Delta t \hat{H}}\right)^{N_t} \; \hat{\psi}^\dagger(x_i)|0\rangle = \langle 0|\hat{\psi}(x_f) \prod_{l=1}^{N_t} \left(e^{-\frac{i}{\hbar}\Delta t \hat{H}}\right) \psi^\dagger(x_i)|0\rangle$$

$$= \int \prod_{k=1}^{N_t} \left[\frac{D\phi_k(x) D\phi^*_k(x)}{2\pi i} \; e^{-\int d^3x \phi^*_k(x)\phi_k(x)}\right] \; \langle 0|\hat{\psi}(x_f)|\phi_{N_t}\rangle \tag{7.50}$$

$$\cdot \langle\phi_k|e^{-\frac{i}{\hbar}\hat{H}\Delta t}|\phi_{k-1}\rangle \ldots \langle\phi_1|\hat{\psi}^\dagger(x_i) \; |0\rangle, \tag{7.51}$$

where the symbol $D\phi(x)$ collectively denotes a product of infinitely many integrals, one for each point in *space*. Each of such integrals is performed over the values of the complex number ϕ associated to a different point of space x. Hence $D\phi_k$ represents the functional measure an integral over the field configuration at the k-th time slice. $|\phi\rangle$ denotes the direct product of infinitely many coherent states, each one associated to a different point in space.

First, let us notice that.

$$\begin{aligned}\langle x_f|\phi\rangle = \langle 0|\hat{\psi}(x_f)|\phi\rangle &= \langle 0|\hat{\psi}(x_f)\ e^{\int d^3x\phi(x)\hat{\psi}^\dagger(x)}|0\rangle = \phi(x_f) &(7.52)\\ \langle\phi|x_i\rangle = \langle\phi|\hat{\psi}^\dagger(x_i)|0\rangle &= \phi^*(x_i). &(7.53)\end{aligned}$$

Next, we use Eq. (7.46) and find

$$\langle\phi_{k+1}|e^{-\frac{i}{\hbar}\Delta t\hat{H}}|\phi_k\rangle = e^{\int d^3x\phi^*_{k+1}(x)\phi_k(x)-\frac{i}{\hbar}\Delta t\ H(\phi^*_{k+1},\phi_k)} \tag{7.54}$$

It is convenient to introduce the notion of time-dependent trajectory $\phi(t)$ to collectively represent the value of the field at all points at time t. In addition, we use the notation

$$\phi^*(x,t)\ \frac{\phi(x,t)-\phi(x,t-\Delta t)}{\Delta t} \equiv \phi^*(x,t)\frac{\partial}{\partial t}\phi(x,t) \tag{7.55}$$

Consistently with our order-Δt discretization, we also perform the identification:

$$H[\phi^*(t),\phi(t-\Delta t)] \to H[\phi^*(t),\phi(t)] \tag{7.56}$$

With such a notation, the exponent collecting all the phases obtained from all elementary propagators can be written as follows:

$$\begin{aligned}&\int d^3x\phi^*(x,t)\phi(x,t) + \frac{i}{\hbar}\int_0^t dt'\int d^3x\left[\ i\hbar\ \phi^*(x,t')\frac{\partial}{\partial t'}\phi(x,t') - H[\phi^*(x,t),\phi(x,t)]\right]\\ \equiv\ &\int d^3x\phi^*(x,t)\phi(x,t) + \frac{i}{\hbar}\mathcal{S}[\phi^*,\phi], \qquad (7.57)\end{aligned}$$

where

$$\mathcal{S}[\phi^*,\phi] \equiv \int_0^t dt'\int d^3x\left[\ i\hbar\ \phi^*(x,t')\frac{\partial}{\partial t'}\phi(x,t') - H[\phi^*(x,t),\phi(x,t)]\right] \tag{7.58}$$

is the so-called Schrödingier action functional, whose variation with respect to the field generates the Schrödinger equation of motion for the fields.

Combining these results in the functional integral notation, we find our final result for the Feynman propagator:

$$\langle x_f|e^{-\frac{i}{\hbar}\hat{H}\ t}|x_i\rangle \ = \ \mathcal{N}\int\mathcal{D}\phi^*\mathcal{D}\phi\ e^{\int d^3x\phi^*(x,t)\phi(x,t)}\phi(x_f,t)\ \phi^*(x_i,0)\ e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]} \tag{7.59}$$

In this expression we have used the functional measure $\mathcal{D}\phi$ to represent infinitely many integrals over the values of the field at each point in *space and in time* and we denoted with $\mathcal{N}$ the standard normalization factor. To get rid of this ugly normalisation factor, we multiply and divide by the vacuum-vacuum transition $\langle 0|e^{-tH}|0\rangle$ and use $\hat{H}|0\rangle \equiv 0$. Thus

$$\langle x_f|e^{-\frac{i}{\hbar}\hat{H}\ t}|x_i\rangle \ = \ \frac{\int\mathcal{D}\phi^*\mathcal{D}\phi\ e^{\mathcal{L}(t)}\phi(x_f,t)\ \phi^*(x_i,0)\ e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]}}{\int\mathcal{D}\phi^*\mathcal{D}\phi\ e^{\mathcal{L}(t)}\ e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]}} \tag{7.60}$$

where $\mathcal{L}(t) \equiv \int d^3x\phi^*(x,t)\phi(x,t)$. Thus the normalization factor is $\mathcal{N} = 1/Z$ with

$$Z = \int \mathcal{D}\phi^*\mathcal{D}\phi \; e^{\mathcal{L}(t)} \quad e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]}. \tag{7.61}$$

It is straightforward to generalise this calculations to include an arbitrary number of particles in the initial and final states. This simply amounts to introduce more fields in the prefecture entering in the numerator. For example, the two particle propagator reads:

$$\langle x_f y_f | e^{-\frac{i}{\hbar}\hat{H}\; t} | x_i y_i \rangle \quad = \quad \frac{\int \mathcal{D}\phi^*\mathcal{D}\phi \; e^{\mathcal{L}(t)}\phi(x_f,t)\; \phi^*(x_i,0)\; \phi(y_f,t)\; \phi^*(y_i,0)\; e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]}}{\int \mathcal{D}\phi^*\mathcal{D}\phi \; e^{\mathcal{L}(t)} \; e^{\frac{i}{\hbar}\mathcal{S}[\phi^*,\phi]}} \tag{7.62}$$

We note one significant difference between the coherent state functional integral (7.62) and the Feynman path integral (1.17) is in the dependence upon $\hbar$. In the Feynman case, an overall factor $\frac{i}{\hbar}$ appears as a constant multiplying the integral at the exponent, hence the stationary phase approximation yields directly the classical limit. On the other hand, in the coherent state path integral, the factor $\frac{i}{\hbar}$ appears also inside the Schrödinger action, so that the stationary phase approximation yields results which are distinct from the classical limit.

We conclude this section by deriving an alternative operator expression for the Feynman's path integral.

$$\begin{aligned}
\langle x_f | e^{-\frac{i}{\hbar}\hat{H}\; t} | x_i \rangle &= \langle 0 | \psi(x_f) e^{-\frac{i}{\hbar}\hat{H}t} \psi^\dagger(x_i) | 0 \rangle \\
&= \langle 0 | e^{+\frac{i}{\hbar}\hat{H}t}\; \psi(x_f) e^{-\frac{i}{\hbar}\hat{H}t} e^{\frac{i}{\hbar}\hat{H}0} \psi^\dagger(x_i) e^{-\frac{i}{\hbar}\hat{H}0} | 0 \rangle \\
&= \langle 0 | T[\hat{\psi}(x_f,t)\hat{\psi}^\dagger(x_i,0)] | 0 \rangle.
\end{aligned} \tag{7.63}$$

In this expression, $\hat{\psi}(x,t)$ denotes the field operator in the Heisemberg representation and $T[\ldots]$ represents the time ordering. Hence, we have shown that the field theoretic path integral is related to time ordered vacuum expectation values of field.

7.3 Statistical Fields

In the previous section we have seen how field-theoretic path integral naturally emerges in the discussion of quantum dynamics, when the propagator is evaluated using the coherent state basis of the Fock space. In statistical mechanics, field theoretic path integrals are often adopted to represent the partition functions of classical systems with infinitely many degrees of freedom. Such fields, however, do not represent quantum coherent quantum states but should only be regarded as an infinite collection of stochastic variables describing the local configurations of the system.

A standard example of statistical field theory is used to represent Landau's model for ferromagnetic phase transition (see e.g. [11]) . It can be shown that, when the temperature of a magnetizable material approaches the critical temperature of the ferromagnetic phase-transition, all correlation lengths diverge, decoupling from the microscopic scale a associated to the lattice spacing. Therefore, in this temperature regime is possible to define an effective theory with spacial resolution defined by a new length scale $\lambda \gg a$. At such a low-resolution, the configuration of the spins in the lattice can be effectively provided by a field of stochastic variables $\phi(x)$, which represent locally the result of averaging the lattice spin components in a small portion of volume λ^3, centred around point x. The canonical partition function of the (D-dimensional) Landau's effective theory is

$$Z[B] = \int \mathcal{D}\phi \; e^{-\beta H[\phi,B]} \tag{7.64}$$

where $H[\phi, B]$ is the Hamiltonian:

$$H[\phi, B] = \int d^D x \left[\frac{1}{2}(\nabla\phi(x))^2 + r_0(T)\phi^2(x) + u_0(T)\phi^4(x) \right] - \int d^D x \; \phi(x) B(x) \tag{7.65}$$

and $B(x)$ is the modulus of the external magnetic field, assumed to be pointing along a fixed direction $\hat{z}$. Other example of statistical field theory are applied to model the equilibrium properties of soft-condensed matter systems, such as e.g. liquid crystals or polyelectrolytes.

7.4 N-point Correlation Functions and the Generating Functional

Much of the physical information which can be extracted from quantum or statistical field theories is encoded in their $n-$point correlation functions. These Green's functions express the dynamical or statistical correlations between arbitrary subsets of the degrees of freedom in the system, as a result of the effect of interactions, thermal or quantum fluctuations and possibly of quantum statistics.

As a first example, let $\phi(x)$ be the statistical field which expresses the local spin polarization in Landau's theory. Then the combination of thermal averages

$$G_2(x, y) = \langle \phi(x)\phi(y) \rangle - \langle \phi(x) \rangle \langle \phi(y) \rangle \tag{7.66}$$

tells us how the local polarization at the points x and y are correlated at thermal equilibrium. In particular, we expect that in the paramagnetic phase such a correlation should drop exponentially to 0 with the distance between x and y, i.e. $G_2(x, y) \propto \exp(-\frac{|x-y|}{2\sigma})$, where σ is called the correlation length.

As a second example one can consider the correlation between the position of the endpoints of a linear polymer. In the next section, we shall show that, in the limit in which this molecule can be described by a "leckneck" of beads connected by harmonic springs (random chain), then $G_2(x, y) \propto \exp(-(x - y)^2/4\kappa^2)$, where κ is called the persistence length.

Finally, we have seen that in quantum field theory the time-dependent time-ordered Green's function

$$\Pi_2(\mathbf{x}, t; \mathbf{y}, t') = \langle 0 | T[\hat{\psi}(\mathbf{x}, t)\hat{\psi}^\dagger(\mathbf{y}, t')] | 0 \rangle \tag{7.67}$$

expresses the probability amplitude for a particle to be prepared at position $\mathbf{y}$ at time t' to be observed at $\mathbf{x}$ at time t'. In view of its path integral representation, this Feynman propagator can be interpreted as the correlation function of the quantum fields.

Of course, it is possible to define correlation functions involving an arbitrary number of fields, which encode more and more specific information about the systems' dynamical correlations. For example the quantum field theory Green's function

$$\Pi_2(\mathbf{x},t;\mathbf{y},t') = \langle 0|T[\hat{\psi}(\mathbf{x},t)\hat{\psi}(\mathbf{x}',t)\ \hat{\psi}^\dagger(\mathbf{y},t')\hat{\psi}^\dagger(\mathbf{y}',t')]|0\rangle \tag{7.68}$$

naturally occurs in the study bound-states and scattering problems.

Solving a quantum or statistical field theory ultimately amounts to being in condition to compute arbitrary n-point Green's functions. In the following sections we shall discuss techniques to achieve this goal through ad-hoc approximations. For sake of definiteness, we shall carry out this discussion focusing on a specific statistical field theory consisting of a single real statistical scalar field in D dimension[1] described by the following Hamiltonian

$$H = \int d^Dx \left(-\frac{1}{2}\phi(x)\nabla^2\phi(x) + \frac{m^2}{2}\phi^2(x)\right) + V[\phi(x)], \tag{7.69}$$

where $V[\phi]$ is the so-called interaction potential for the field. A particularly interesting case is one in which $V[\phi]$ is a local polynomial potential, i.e. it is in the form

$$V[\phi] = \sum_{\alpha=1}^{n} \frac{g_\alpha}{\alpha!}\int d^Dx\ \phi^\alpha(x), \tag{7.70}$$

where α is an integer number greater than 3 and g_α are real coupling constants. The factorial pre-factor $\frac{1}{\alpha!}$ is introduced for convenience. Indeed, as we shall shortly see, a theory with a ϕ^α-type interaction will generate a perturbation theory in which leading-order corrections consist of $\alpha!$ different terms.

In discussing the calculation of N-point correlation function in field theory it is very convenient to introduce a mathematical object called *generating functional*, defined as follows

$$Z[j] = \int \mathcal{D}\phi\ e^{-\beta H[\phi]-\int d^Dx j(x)\phi(x)} \tag{7.71}$$

In this expression $j(x)$ is an arbitrary external field — which will be called the *source field*. In some particular cases, -source fields may have a direct physical interpretation. For example in Landau theory, they represent the external magnetic field $B(x)$. In general, the usefulness of introducing the generating functional with external sources is that arbitrary N-point correlation functions can be obtained by multiple appropriate (functional) differentiation of the logarithm of the generating function. For example, the two point correlation function (7.6.1) is given by:

$$G_2(x,y) \quad = \quad \lim_{j(x)\to 0}\left[\frac{\delta^2 \log Z[j]}{\beta^2\delta j(x)\delta j(y)} - \left(\frac{\delta \log Z[j]}{\delta j(x)}\right)^2\right] \tag{7.72}$$

---

[1] The extension to coherent fields is straightforward and further applications of these methods to relativistic vector or Dirac fields can be found in all introductory quantum field theories textbooks.

Hence, in completely general terms, the problem of solving a field theory is ultimately equivalent to computing the generating functional with an arbitrary source term.

7.5 Gaussian Path Integrals

In this section, we shall show that when the Hamiltonian of a field theory is at most quadratic in the fields it is possible to analytically evaluate the path integral. This result provides the basis to derive perturbation theory and Feynman diagrams, which will be discussed in section 7.6.

7.5.1 Wick Theorem

We assume that the exponent entering the definition of generating functional of$Z[j]$ contains only terms which are linear and quadratic in the fields. Namely, we consider the a general structure of the Hamiltonian in the form:

$$H_0[\phi] = \frac{1}{2}\int d^Dx\phi(x)\hat{W}(x)\phi(x) \tag{7.73}$$

where $\hat{W}(x)$ may be a differential operator. Choosing a system of units in which $\beta = 1$ generating functional reads:

$$Z[J] = \int \mathcal{D}\phi e^{-\frac{1}{2}\int d^Dx\phi(x)\hat{W}(x)\phi(x) - \int d^DxJ(x)\phi(x)} \tag{7.74}$$

For example, a paramagnetic system may be described by the quadratic Hamiltonian H_0::

$$H_0[\phi] = \int d^Dx\phi(x)\left(-\frac{1}{2}\nabla^2 + \frac{1}{2}m^2\right)\phi(x) \tag{7.75}$$

The trick to conveniently evaluate the path integral is to perform a shift of the field variables using the following re-definition:

$$\phi(x) \to \phi'(x) = \phi(x) - \int d^Dy\Delta_0(x-y)J(y), \tag{7.76}$$

where $\Delta_0(x)$ is the Green's function of the operator which enters in the quadratic part of the Hamiltonian, i.e.

$$\hat{W}(x)\Delta_0(x-y) = \delta(x-y). \tag{7.77}$$

In the specific example provided by the Hamiltonian (7.75) one can evaluate $\Delta_0(x-y)$ by transforming this equation into Fourier space. Then one immediately reduces to quadrature the problem of computing $\Delta_0(x-y)$:

$$\Delta_0(x-y) = \langle\phi(x)\phi(y)\rangle_0 = \int \frac{d^Dk}{(2\pi)^D}\,\frac{e^{-i(x-y)\cdot k}}{k^2+m^2}, \tag{7.78}$$

The result of this Fourier transform obviously depends on the number of dimensions. In particular, for $D = 1, 3, 4$ one finds

$$D = 1 : \qquad \Delta_0(x) = \sqrt{\frac{\pi}{2}}\, e^{-m|x|} \tag{7.79}$$

$$D = 3 : \qquad \Delta_0(x) = \frac{e^{-m|x|}}{4\pi|x|} \tag{7.80}$$

$$D = 4 : \qquad \Delta_0(x) = \frac{m}{4\pi^2|x|} K_1(m|x|). \tag{7.81}$$

Where $K_1(z)$ is a modified Bessel function of the first kind. The expression for $D = 2$ involves a cumbersome combination of special functions and will not be reported here.

The change of variable (7.76) has obviously a unitary Jacobian. Upon dropping the redundant superscript $'$ the generating functions in the new variables reads:

$$Z_0[j] = Z_0[j = 0] \times e^{\frac{1}{2}\int d^Dx \int d^Dy \; j(x) \; \Delta_0(x-y) \; j(y)}. \tag{7.82}$$

Note that the pre-factor $Z_0[j = 0]$ is an overall constant which does not influence thermal averages, hence can be dropped.

An important observation is that the Green's function $\Delta_0(x-y)$ can be identified with the two-point function:

$$\Delta_0(x - y) = G_2(x - y) = \langle \phi(x)\phi(y) \rangle \tag{7.83}$$

This can be shown directly applying the functional derivatives in Eq. (7.72) and then setting the external currents to 0. Note also that, for a quadratic Hamiltonian and vanishing external current $j = 0$, the field averages vanish by symmetry, i.e. $\langle \phi(x) \rangle = 0$.

Having obtained the explicit expression for the generating functional in the free theory, it is straightforward to evaluate all $n-$point correlation functions, by performing multiple functional differentiation and then setting the source to 0. Any time the functional differentiation acts on the exponent, it brings down a new power of the external source in the pre-factor. For example

$$\begin{aligned} &\frac{\delta}{\delta j(z)} e^{\frac{1}{2}\int d^Dx \int d^Dy \; j(x) \; \Delta_0(x-y) \; j(y)} \\ = &\int d^Dy \Delta_0(z - y) \; j(y) e^{-\frac{1}{2}\int d^Dx \int d^Dy \; j(x) \; \Delta_0(x-y) \; j(y)} \end{aligned} \tag{7.84}$$

The next time a functional derivative is performed, it can act on the pre-factor or on the exponent. In the former case it lowers the order of the source polynomial in the pre-factor, while in the second case increases. Therefore the functional differentiation leads to an expression in the general form:

$$P[j] e^{\frac{1}{2}\int d^Dx \int d^Dy j(x) \Delta_0(x-y) j(y)} \tag{7.85}$$

where $P_n[j]$ is a polynomial of order n in the sources, whose coefficients contain product of Green's functions $\Delta_0(x - y)$. Now we recall that at the end of the calculation all the sources must be set to zero. Thus only the zero-th order term in

$$\Delta_0(x_1,x_5)\Delta_0(x_2,x_6)\Delta_0(x_3,x_8)\Delta_0(x_4,x_7) \Longleftrightarrow$$

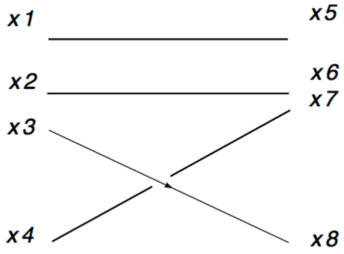

Figure 7.1: Graphical representation of a Wick contraction involved in a 8-point correlation function

.

$P_n[j]$ will survive. Based on this observation, it is immediate to proof the following fundamental result which is usually referred to as Wick theorem:

$$\begin{aligned}\langle\phi(x_1)\ \phi(x_2)\ldots\phi(x_n)\rangle_0 &= \Delta_0(x_1-x_2)\ \Delta_0(x_3-x_4)\ \ldots\ \Delta_0(x_{n-1}-x_n)\\ &\quad +(\text{ all possible permutations of space indexes})\end{aligned} \tag{7.86}$$

A graphical representation of the Wick contraction appearing in a 8-point function is given in Fig. 7.1.

Some comments on this result are in order. In the quantum field theory literature, the propagator $\Delta_0(|x_1-x_2|)$ is sometimes referred to as a *contraction* of $\phi(x_1)$ and $\phi(x_2)$, With such a terming, the Wick theorem can be formulated as " the n-point correlation function is obtained by summing over all possible products of contractions". We also observe that correlations functions with an odd number of fields are identically vanishing.

Feynman introduced an extremely useful graphical notation to represent the result of Wick theorem for generic n-point correlation function. In Feynman diagrams, each contraction is (propagator $\Delta_0(x,y)$ denoted with a line connecting the endpoints x and y. Then each permutation in the sum over contraction is represented by a different way of connecting by straight lines the n-points of the correlation function. A simple example is given in Fig. 1, where we show one particular contraction appearing in the evaluation of the 8-point correlation function $G_8(x_1,\ldots,x_8)$.

7.6 Perturbation Theory

In the previous section we have discussed the exact solution of Gaussian path integrals, and the associated Wick theorem. We recall that Gaussian path integral are associated to quadratic Hamiltonians (or actions in quantum theories), which describe systems in which the dynamical degrees of freedom are not coupled to one another, but only (at most) to an external field (the source).

The usefulness of the results concerning Gaussian path integrals go well beyond the discussion of free systems. Indeed, free fields provide the basis to develop a perturbation theory which is applicable any the Hamiltonian contains terms which are of higher order polynomials in the fields, yet the corresponding coupling constant can be regarded as a small parameter.

The key step towards developing such a perturbation theory is to expand the exponent containing the interaction terms in Taylor series of the coupling. This way, one reduces the problem of computing an arbitrary $n-$point correlation function in the interacting theory to that of evaluating a number of $m-$point correlation functions with $m \geq n$, in the *free theory*. Then, Wick theorem and Feynman diagrams can be used to evaluate each of such correlation functions

We now discuss in detail a simple example in which this program is carried out in detail and use this to infer Feynman rules, i.e. a set of procedures which enables one to avoid complicated and error-prone calculations, and obtain directly the result, from a set of surprisingly simple graphical rules.

7.6.1 Example: Calculation of a Two-Point Correlation Function in ϕ^4 Theory.

Consider an interacting theory with a potential in the form

$$V[\phi] = \frac{\lambda}{4!}\int d^D x \phi^4(x). \tag{7.87}$$

Our task is to evaluate the two point correlation function

$$G_2(x,y) = \langle \phi(x)\phi(0)\rangle = \frac{\int \mathcal{D}\phi\ \phi(x)\phi(0)\ e^{-\beta H[\phi]}}{\int \mathcal{D}\phi\ e^{-\beta H[\phi]}} \tag{7.88}$$

up to order λ. To this end, we separately analyze the leading terms in the Taylor expansion for small λ of both the numerator and denominator of (7.88).

The numerator is:

$$\begin{aligned} N_2(x,y) &= \int \mathcal{D}\phi\ \phi(x)\phi(y)\ e^{-\beta H_0[\phi]} + \int \mathcal{D}\phi\ \phi(x)\phi(y)\ \left(-\frac{\lambda}{4!}\int d^D z \phi^4(z)\right)\ e^{-\beta H_0[\phi]} \\ &+ \frac{\lambda^2}{(4!)^2}\int \mathcal{D}\phi\ \phi(x)\phi(y)\ \int d^D z\phi^4(z)\ \int d^D w \phi^4(w)\ e^{-\beta H_0[\phi]} + \dots \end{aligned} \tag{7.89}$$

Upon multiplying and diving for the partition function of the free theory we find:

$$\begin{aligned} N_2(x,y) &= Z_0\Bigg(\langle\ \phi(x)\phi(0)\rangle_0 - \frac{\lambda}{4!}\int d^D z\langle\ \phi(x)\phi(y)\phi^4(z)\ \rangle_0 \\ &+ \frac{\lambda^2}{(4!)^2}\int d^D z\int d^D w \langle \phi(x)\phi(y)\ \phi^4(z)\phi^4(w)\rangle_0 + \dots\Bigg) \end{aligned} \tag{7.90}$$

Similarly, the denominator reads

$$D_2 = Z_0\left(1 - \frac{\lambda}{4!}\int d^D z\langle\ \phi^4(z)\ \rangle_0 + \frac{\lambda^2}{(4!)^4}\int d^D z\int d^D w\langle\ \phi^4(z)\ \phi^4(w)\rangle_0 \dots\right) \tag{7.91}$$

Each correlation function in the numerator and in the denominator can be evaluated using Wick theorem. The graphical representation of the sum over all contractions

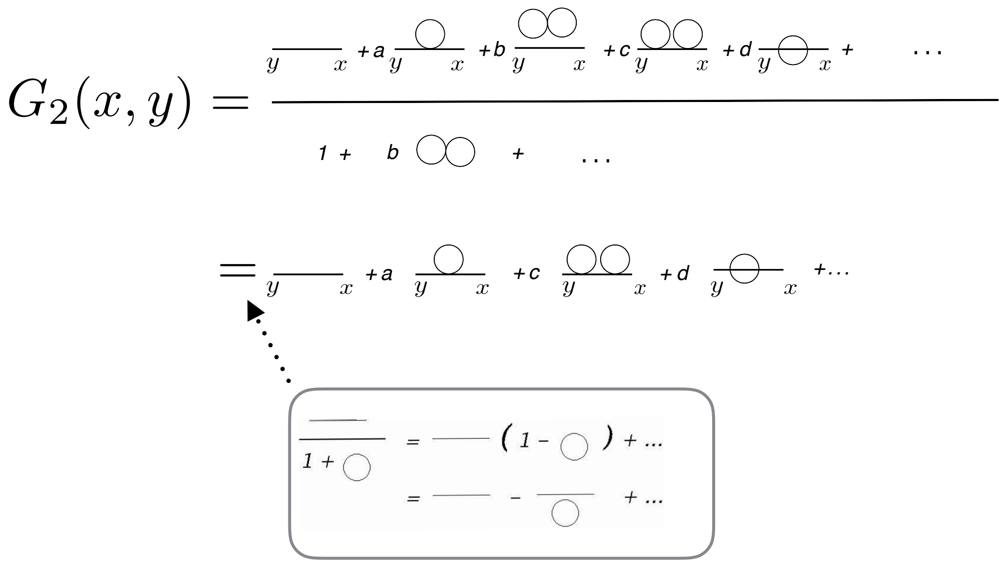

Figure 7.2: Graphical representation of the contribution of the numerator and denominator to the two-point correlation function in the ϕ^4 theory to order λ^2. The lower inset shows an example of cancellation of disconnected diagrams

in the numerator and denominator is given in Fig. 1.1. In this diagram we are adopting a notation in which an integral over the position of each vertex attached to a closed loop is implicitly assumed and $a, b, c, \ldots$ are integer numbers called symmetry factors. These coefficients count the numbers of different contractions which give raise to topologically equivalent diagrams. In general, evaluating symmetry factors is a tedious task, as it requires performing all Wick contractions. In some cases (such as e.g for the present ϕ^4 scalar field theory of for Quantum Electrodynamics) simple graphical rules can be established yield directly the symmetry factor and avoid having to perform all Wick contractions explicitly (see e.g. [11, 12, 13]). Another important observation is that all topologically disconnected diagrams in the numerator cancel exactly with the diagrams in the denominator, order-by-order in perturbation theory. This feature is explained graphically in Fig. 1.8 and is a general feature of the perturbative expansion in field theory .

To zero-th order in the coupling, we trivially recover the unperturbed result:

$$G_2(x, y) = \Delta_0(x - y). \tag{7.92}$$

To order λ, only one diagram contributes and gives:

$$G_2(x, y) = \Delta_0(x - y) - \frac{\lambda}{2} \int d^D z \Delta_0(x - z)\ \Delta_0(z - y)\ \Delta_0(z - z) \tag{7.93}$$

Notice that, already a this level, a divergence occur since $\lim_{\epsilon \to 0} \Delta_0(\epsilon) = \infty$. This an example of the famous UV divergences which appear in the perturbative expansion of all statistical and quantum field theories. They emerge because the locality of the interaction equally couples of arbitrary momenta and can be cured by the renormalization procedure. For excellent introductory discussion of modern renormalisation theory and on the physical origin of UV divergences we recommend the reader to consult Lepage's lectures on "How to renormaliza the Schrödinger Equation" [16].

From this example we can read-off some features which are preserved in the perturbative calculation of arbitrary n-point correlation function and are usually referred as Feynman Rules, which are tremendously useful. Indeed, they enable to derive the results of otherwise lengthy and error-prone calculations, simply by drawing a limited number of so-called Feynman diagrams. Each of these graphs can be translated into an equation and represents a family of Wick contractions.

Feynman rules can be formulated both in configuration space and momentum space. The former are most commonly adopted in statistical field theory, while the latter are the standard choice to tackle scattering problems in quantum field theory. The rules in momentum space can be obtained from those in coordinate space by analysing the result of Fourier transforming the corresponding equations. A detailed discussion and derivation of Feynman rules can be found in essentially all introductory quantum and statistical field theory textbooks, such as e.g. [11] and [13].

Feynman rules for scalar field theory in coordinate space

1. Draw all topologically different connected Feynman graphs, linking the end-points of the correlation function.

2. Introduce a number of internal vertexes defined by the order of the perturbative expansion, i.e. to the desired order of coupling g_α. Each vertex is assigned a factor $-g_\alpha$ where g_α is the coupling constant

3. impose the conservation of the total momentum flowing through the vertex.

4. every line joining two points x_1 and x_2 is assigned a propagator $\Delta_0(x_1 - x_2)$.

5. Integrate over all internal points z_i: $\int d^D z_i$

6. To every graph assign the corresponding *symmetry factor*.

Feynman rules for scalar field theory in momentum space

1. Draw all topologically different connected Feynman graphs, linking the end-points of the correlation function.

2. Introduce a number of internal vertexes defined by the order of the perturbative expansion, i.e. to the desired order of coupling g_α. Each vertex is assigned a factor $-g_\alpha$ where g_α is the coupling constant.

3. To every line assign a propagator $\tilde{\Delta}_0(k) \equiv \int dy e^{ik\cdot x} \Delta_0(x, 0)$.

4. Integrate over the momentum flowing inside each loop.

5. Each graph is assigned its symmetry factor

7.7 Non-Perturbative Methods in Field Theory

In this section, we consider problems in which a perturbative approach is not tenable. As a prototypical case, we shall mostly focus on the ferro-magnetic transition in Landau theory. Alternative problems include bound state calculations in Euclidean (i.e. Wick rotated) quantum field theory.

We refer the reader interested in the phenomenology of ferromagnetic transitions to specialized books on the subject [2]. One of the essential features of second order phase transitions is that correlations function diverge at the critical point. This striking phenomenon implies that an exponentially large number of degrees of freedom (representing lattice spins) are simultaneously interacting. This situation clearly defies the condition of operation of perturbation theory. In Landau theory, this problem is circumvented by invoking a mean-field approximation which can be deduced using functional integral methods.

The Landau's approach was originally derived by directly postulating an expression for the Gibb's free-energy as a function of the magnetization M, $\Gamma[M]$. Here, we follow the approach described in Le Bellac's Quantum and Statistical Field Theory book and consider a more microscopic standpoint. Specifically, we start by defining the field-theoretic Hamiltonian

$$H[\phi] = \int d^D x \phi(x)(-\frac{1}{2}\nabla^2 + \frac{r_0}{2})\phi(x) + \frac{u_0}{4!}\phi^4(x) \tag{7.94}$$

In this expression the field $\phi(x)$ represent the local spin density and is used to represent the thermodynamics of the ferro-magnetic material only near the critical condition. Then, the structure of the Hamiltonian (7.94) can be inferred from Renormalization Group theory effective field theory arguments. Namely, near the critical point, the magnetization $M(x) = \langle\phi(x)\rangle$ is small, thus terms containing higher power polynomials of fields s also suppressed. In addition, for $T \sim T_c$ correlation lengths decouple, which that the contribution of higher derivative terms in the effective Hamiltonian are sub-leading. Finally, in this effective theory we are allowing the coefficients r_0 and u_0 to "run" with the temperature, which effectively represents a renormalization scale.

Adopting units in which $\beta = 1$, the generating functional $Z[B]$ for this theory reads:

$$Z[B] = \int \mathcal{D}\phi \; e^{-H[\phi]+\int d^D x \phi(x) B(x)}, \tag{7.95}$$

Notice that, in expression (7.95), the unphysical external source (usually denoted with $J(x)$) has been replaced a physical external magnetic field $B(x)$. Thus, unlike in the our previous discussion concerning the path integral derivation of Wick theorem, here we are also interested in the most general case in which the source field is finite. As usual, the Helmotz free-energy in units of $1/\beta$ is given by

$$F[B] = -\log Z[B]. \tag{7.96}$$

---

[2] An excellent treatment (which closely resembles the one we are going to give here) can be found e.g. in Le Bellac's book ("Quantum and Statistical Field Theory")

7.7.1 Saddle-Point Approximation

To infer the emergence of a phase transition, we need to perform a change in thermodynamic coordinates, i.e. to implement a Legendre transformation from the Helmotz free-energy $F[B]$ to the Gibbs free-energy $\Gamma[M]$, which explicitly depends on the local magnetization field, defined as

$$M(x) = \langle \phi(x) \rangle = -\frac{\delta}{\delta B(x)} F[B]. \tag{7.97}$$

The Gibbs' Free energy is defined as the Legendre transform

$$\Gamma(M) = \left(F[B] + \int d^D x M(x) B(x) \right)_{B=B(M)} \tag{7.98}$$

Note that the magnetic field and the magnetization are conjugate variables, i.e. the connection between B and M is provided by

$$B = \left. \frac{\partial \Gamma}{\partial M} \right|_T \tag{7.99}$$

To approximatively evaluate the Legendre transform leading to the Gibb's free energy, we first evaluate the path integral defining the partition function $Z[B]$ in saddle-point approximation. To this end, we seek for stationary points of the exponent in the integrant of $Z[B]$:

$$\frac{\delta}{\delta \phi(x)} \left(H[\phi] - \int d^D y \phi(y) B(y) \right) = 0 \tag{7.100}$$

Let $\bar{\phi}_B(x)$ be the solution of such an equation for a given choice of the external field B, i.e.

$$\left. \frac{\delta}{\delta \phi(x)} H[\phi] \right|_{\phi = \bar{\phi}_B} = B(x) \tag{7.101}$$

In saddle-point approximation we have

$$Z[B] = e^{-F[B]} \sim \mathcal{N} \, e^{-H[\bar{\phi}_B] + \int B(x) \bar{\phi}_B(x)} \tag{7.102}$$

In order to evaluate the Gibb's free energy we need a connection between $\bar{\phi}_B$ and the magnetization M. To this end, we evaluate:

$$\begin{aligned} M(x) &= -\frac{\delta}{\delta B(x)} \log Z \\ &\simeq -\int d^D y \, \frac{\delta H[\phi_B]}{\delta \phi_B(y)} \, \frac{\delta \bar{\phi}_B(y)}{\delta B(x)} + \bar{\phi}_B(x) \\ &\quad + \int d^D y B(y) \frac{\delta}{\delta B(y)} \bar{\phi}_B(y) \end{aligned} \tag{7.103}$$

Then, using Eq. (7.101)

$$M(x) = -\frac{\delta}{\delta B(x)} \log Z[B] = \bar{\phi}_B(x) \tag{7.104}$$

Then, using Eq. (7.104) into Eq. (7.98) we find

$$\begin{aligned}\Gamma(M) &\sim \left(-\log Z + \int d^D x M(x)B(x)\right)_{B=B(M)} \\ &= \left(H[\bar{\phi}_B] - \int d^D x B(x)\bar{\phi}_B(x) + \int d^D x M(x)B(x))\right)_{B=B(M)} \\ &= H[M] \end{aligned} \tag{7.105}$$

Thus, to saddle-point level, the Gibb's free-energy is obtained by replacing ϕ with M into the functional form of the Hamiltonian. This is, however, only true at the saddle-point level of approximation. In Landau theory, this connection is exploited to explore the phases of the ferromagnet by assuming that, at the critical temperature r_0 changes sign. In particular, below the critical temperature $r_0(T) < 0$. Assuming a uniform system enables to remove spatial dependences. As a result, the system's magnetization is the one which minimizes the Gibb's free-energy:

$$\Gamma[M] = (-\frac{r_0}{2}M^2 + \frac{u_0}{4!}M^4) \tag{7.106}$$

The magnetization is the obtained by looking for stable stationary point configuration. For negative values of r_0 we find also solutions for $M \neq 0$, representing the ordered phase.

We emphasize that, in general, these equation have been derived in the presence of an unspecified external magnetic field. To read-off the connection between the magnetic field and the magnetization, we use Eq. (7.99):

$$B = \frac{\delta\Gamma}{\delta M} = -r_0 M + \frac{u_0}{6}M^3 \tag{7.107}$$

which is a standard result of mean-field approximation applied to the Ising Model. This is no coincidence, indeed to saddle-point level one retrieves only the field configuration $\bar{\phi}(x) = \langle\phi(x)\rangle$ which is the mean-value of the field. A similar situation was encountered when applying the stationary-phase approximation to the Feynman path integral, yielding the classical path, which (by Erhenfest theorem) evolves according to the expectation value of position operator. In general, a saddle-point (or stationary-phase) approximations are always equivalent to a mean-field limit.

7.7.2 Beyond Mean-field: The Ginzburg Correction

Landau theory completely neglects the effects of fluctuations around the mean-field. Here, we address the problem of perturbatively accounting for the effects of small fluctuations around the mean-field. We emphasize the fact that this does not correspond to performing a *bona-fide* perturbative calculation, because the perturbation is evaluated around a non-trivial mean-field, which was obtained nonperturbatively.

To evaluate leading-order corrections to Landau's mean-field approximation, we funcitonally expand the functional at the exponent in the partition function to quadratic order around the saddle-point. Namely, introducing for convenience

$$W[\phi, B] \equiv H[\phi] - \int d^D y \phi(y)B(y) \tag{7.108}$$

and $\psi(x) = \phi - \bar{\phi}_B(x)$, where $\phi_B(x)$ is a solution of the saddle-point equation (7.100) which can be written as

$$\frac{\delta}{\delta\phi(x)} W[\phi, B]\bigg|_{\phi=\phi_B(x)} = 0 \tag{7.109}$$

$$W[\phi] \simeq W[\bar{\phi}_B] + \frac{1}{2}\int d^D x \int d^D y \psi(x) \left.\frac{\delta^2 W[\phi]}{\delta\phi(x)\delta\phi(y)}\right|_{\phi=\bar{\phi}_B} \psi(y) + \dots \tag{7.110}$$

The linear term in the expansion disappears precisely because ϕ_0 is a saddle-point field configuration.

After changing integration variable from ϕ to ψ and retaining this quadratic form at the exponent, the partition function reads

$$Z[B] \simeq \mathcal{N} \int \mathcal{D}\psi \; e^{-\frac{1}{2}\int d^D x \int d^D y \psi(x) \left.\frac{\delta^2 W[\phi]}{\delta\phi(x)\delta\phi(y)}\right|_{\phi=\bar{\phi}_B} \psi(y)} \tag{7.111}$$

We note that the path integral is quadratic, thus can be evaluated analytically. The operator entering the field bilinear term is:

$$\hat{D}[\bar{\phi}_B] \equiv \left.\frac{\delta^2 W[\phi]}{\delta\phi(x)\delta\phi(y)}\right|_{\phi=\bar{\phi}_B} = (-\nabla^2 + r_0 + \frac{u_0}{2}\bar{\phi}_B^2(x)) \tag{7.112}$$

We stress that the nonperturbative character of this calculation is intrinsic in the fact that this operator depends on the nonperturbative mean-field $\bar{\phi}_B$.

To evaluate the Gaussian Integral (7.114) we recall the following general result which holds in any number of dimension:

$$\int \prod_k dx_k \; \exp[-\sum_{ij} x_i O_{ij} x_j] = (\det O)^{-1/2} = \exp[-\frac{1}{2}\mathrm{Tr}\log O] \tag{7.113}$$

where O is a Hermitian matrix. After extending to infinite dimensions, this result can be applied to the quadratic form (7.114):

$$\begin{aligned} Z[B] &\simeq \mathcal{N}\; e^{-W[\bar{\phi}_B]} \int \mathcal{D}\psi \; e^{-\frac{1}{2}\int d^D x \int d^D y \psi(x) \left.\frac{\delta^2 W[\phi]}{\delta\phi(x)\delta\phi(y)}\right|_{\phi=\bar{\phi}_B} \psi(y)} \\ &= \frac{e^{-W[\bar{\phi}_B]}}{\sqrt{\det \hat{D}[\bar{\phi}_B]}} \\ &= e^{-W_{eff}[\bar{\phi}_B]}, \end{aligned} \tag{7.114}$$

where

$$W_{eff}[\bar{\phi}_B] = W[\bar{\phi}_B] + \frac{1}{2}\mathrm{Tr}\log(\hat{D}[\bar{\phi}_B]) \tag{7.115}$$

$W_{eff}[\bar{\phi}_B]$ represents an effective Hamiltonian which includes the effect of fluctuations around the mean-field.

Eq. (7.115) is a formal expression, since the determinant of an infinite dimensional matrix different from identity is in general infinite. To obtain a finite result one needs to go through the regularization and renormalization procedure. To this end, first we note that the logarithm of an operator is defined as its Taylor series

$$\log D = \sum_k (-1)^{k-1} \frac{(\hat{D} - \hat{1})^k}{k} \tag{7.116}$$

Exploiting the linearity of the trace, $\mathrm{Tr} \log D$ can be evaluated by resumming terms individually. In particular, the key ingredient is

$$\mathrm{Tr}[\hat{D}[M]] = \int \frac{d^D p}{(2\pi)^D} (p^2 + r_0^2 + \frac{u_0}{2} M^2) \tag{7.117}$$

Then after resumming all terms, we obtain:

$$\mathrm{Tr}[\log \hat{D}[M]] = \int \frac{d^D p}{(2\pi)^D} \log(p^2 + r_0^2 + \frac{u_0}{2} M^2) \tag{7.118}$$

$$W_{eff}[M] = W[M] + \frac{1}{2} \int \frac{d^D p}{(2\pi)^D} \log(p^2 + r_0^2 + \frac{u_0}{2} M^2) \tag{7.119}$$

The integral appearing in the right-hand-side needs to be regularized, according to the standard renormalization procedure.

In quantum field theory, where the role of the Hamiltonian is played by the action, the Ginzburg correction is usually referred to as the one-look expression of the effective action.

Chapter 8

Wilson's Approach to Renormalization Group

At the beginning of this course (see Prologue) we introduced the concept of Renormalization Group and rigorous effective theory. A pivotal notion at the basis of the very notion of rigorous effective theory is the idea that the parameters in front of the counter-terms in an effective theory parametrize our ignorance about the short-distance (UV) physics which is not resolved by the effective theory.

A natural consequence of this concept is the observation that the effective coefficients must "run" with the renormalization scale, i.e. with the energy or momentum scale Λ at which they are determined by matching the predictions of the effective theory against experiment or against calculations performed in a more microscopic theory. As long as Λ is chosen in the gap between UV and IR scales, by slightly changing it one can gives rise to a whole family of "equivalent" effective theories, i.e. theories which represent the same IR physics, but which different in how some physics near the renormalization scale is represented. Namely, the result of lowering Λ is that of "re-shuffling" some physical content from the IR sector (explicitly represented in the Hilbert space) to the UV sector (implicitly parametrized).

In this chapter[1], we deal with the problem of deriving the set of partial differential equations which can be use predict how the set of effective coefficients $\{C_i(\Lambda)\}_i$ vary with the renormalization scale. The result is a dynamical system which relates the same physical system at different renormalization points. Trajectories in this dynamical system connect families of equivalent theories and define what is called the Renormalization Group flow. Stability points in this dynamical system plays an important role, as we shall see. In essence, they represent special set of parameters which are universal, in the sense that they describe the IR limit of many different effective theory. But we shall come to discuss this point as we move on.

In the following, we provide a very pragmatic, "hands-on" brief excursion into Wilson's renormalization theory, without attempting to be neither very general, nor too rigorous. We shall specifically choose a quartic Hamiltonian in the form of the one used in Landau-Ginzburg theory. However, to keep technical difficulties to a minimum, we shall consider the specific case in which the quartic coupling u_0 is a small parameter, thus correlation functions can be evaluated by perturba-

---

[1] This chapter follows rather closely the discussion in Chapter 12 of Peskin and Schröder's Book "An Introduction to Quantum Field Theory"

tion theory on the quadratic Hamiltoinian. To mark the difference with Landau-Ginzburg Hamiltonian (which is used for the non-perturbative problem of understanding second-order phase-transition) let us here re-write it with a e notation which is closest to that of the (perturbative) quantum field theory literature,

$$H[\phi] = \frac{1}{2}\int d^3x \phi(x)(-\nabla^2 + M^2)\phi(x) + \frac{\lambda}{4!}\phi^4(x) \tag{8.1}$$

The partition function is defined as usual

$$Z_\Lambda = \int \mathcal{D}\phi \; e^{-H[\phi]} \tag{8.2}$$

In this expression, the sub-fix Λ has been introduced to remind us that this partition function is defined up to a finite UV cut-off scale.

8.0.1 Decimation Procedure

The key idea of Wilson's renormalization approach is to analyze the contribution to the contribution to the partition functions' path integral in momentum space. Due to the presence of a UV cut-off Λ, the Fourier transform of field $\tilde{\phi}(k)$ is non-vanishing only within a sphere with a radius Λ, i.e.

$$\tilde{\phi}(k) = 0 \qquad |k| \geq \Lambda \tag{8.3}$$

$$\tag{8.4}$$

All the physics associated to modes with $|k| \geq \Lambda$ is implicitly accounted for by the effective coefficients.

In the regularized theory, the fields $\phi(x)$ can be defined as the inverse Fourier transform of the $\tilde{\phi}(k)$ fields. Indeed, the contribution from harder modes are systematically excluded by the presence of the cut-off.

As a first step in Wilson's renormalization program, we perform the decimation procedure, which consists in explicitly performing the path integral over a thin shell of momentum modes, near the edge of the sphere, i.e. for momenta satisfying the inequalities $b\Lambda < |k| < \Lambda$ where b is a number close to 1 defined in the interval $[0, 1]$.

Namely, we split the fields in two as follows:

$$\begin{aligned}
\phi(x) &= \int_0^{|\Lambda|} \frac{d^3k}{(2\pi)^3}\, e^{ikx}\, \tilde{\phi}(k) \equiv \phi_>(x) + \phi_<(x) && (8.5)\\
\phi_<(x) &= \int_0^{b|\Lambda|} \frac{d^3k}{(2\pi)^3}\, e^{ikx}\, \tilde{\phi}(k) && (8.6)\\
\phi_>(x) &= \int_{b|\Lambda|}^{|\Lambda|} \frac{d^3k}{(2\pi)^3}\, e^{ikx}\, \tilde{\phi}(k) && (8.7)
\end{aligned}$$

We shall refer to the $\phi_>$ and $\phi_<$ as the hard and soft field, respectively.

Since the Fourier transformation is unitary the path integral measure can be split as follows:

$$\int \mathcal{D}\phi(x) \to \int \mathcal{D}\phi_>(x)\mathcal{D}\phi_<(x) \tag{8.8}$$

We then split the contribution of these fields components into the Hamiltonian as follows

$$H[\phi] = H[\phi_< + \phi_>] = H[\phi_<] + \Delta H[\phi_<, \phi_>] \tag{8.9}$$

where $\Delta H[\phi_<, \phi_>] \equiv H[\phi_> + \phi_>] - H[\phi_<]$ and $H[\phi_<]$ is the functional form of the Hamiltonian expressed as a function of the soft field, only. Let us the partition function introducing the splitting into $H_<[\phi_<]$ and $\Delta H[\phi_<, \phi_>]$ explicitly (assuming Einstein's summation convention):

$$\begin{aligned} Z_\Lambda &= \int \mathcal{D}\phi_> \int \mathcal{D}\phi_< \; e^{-\int d^3x (\nabla_i\phi_> + \nabla_i\phi_<)^2 + \frac{M}{2}(\phi_> + \phi_<)^2 + \frac{\lambda}{4!}(\phi_< + \phi_>)^4} && (8.10) \\ &= \int \mathcal{D}\phi_< e^{-H[\phi_<]} \int \mathcal{D}\phi_> e^{-\int d^3x \frac{1}{2}(\nabla_i\phi_>)^2 + \frac{1}{2}M\phi_>^2} \\ & \quad e^{-\lambda \int d^3x \left(\frac{1}{6}\phi_<^3\phi_> + \frac{1}{4}\phi_<^2\phi_>^2 + \frac{1}{4!}\phi_<\phi_>^3\right)} && (8.11) \end{aligned}$$

In this expression, the soft and hard components of the original field $\phi(x)$ are regarded as two independent fields and we have explicitly calculated their interaction. We stress the fact that none of the interactions contains derivatives. This is because, in this theory, derivative coupling could only come from the kinetic energy term. However, as we have previously emphasized, the kinetic energy operator is diagonal in momentum representation, thus does not mix momentum components. The goal of the decimation procedure is to explicitly performing the integral over the hard fields, within a perturbative approach. The result is an effective theory defined in terms of the soft fields only:

$$Z = \int \mathcal{D}\phi_< e^{-H_{eff}[\phi_<]} \tag{8.12}$$

where

$$\begin{aligned} e^{-H_{eff}[\phi_<]} &= e^{-(H[\phi_<] + V_{eff}[\phi_<])} \\ e^{-V_{eff}[\phi_<]} &= Z_>[\phi_<] = \int \mathcal{D}\phi_> e^{-\int d^3x \frac{1}{2}(\nabla_i\phi_>)^2 + \frac{1}{2}M\phi_>^2 + \lambda \;\int d^3x \left(\frac{1}{6}\phi_<^3\phi_> + \frac{1}{4}\phi_<^2\phi_>^2 + \frac{1}{4!}\phi_<\phi_>^3\right)} \end{aligned} \tag{8.13}$$

We note that the effective interaction $V_{eff}[\phi_<]$ is basically the Helmotz free-energy associated with the hard fields, in which the soft fields play the role of an external source term. To evaluate it we recall that, in the perturbative expansion, the partition function can be expressed as a sum of loop diagrams (see e.g. the graphs in the "denominator" of Fig. 6.2). In addition, the result taking the logarithm to relate $V_{eff}[\phi_<]$ to $Z_>[\phi_<]$ is that of removing all the topologically disconnected loop graphs, order-by-order in perturbation theory. In conclusion, to evaluate $V_{eff}[\phi_<]$ in perturbation theory, we need to sum over to a given order all the disconnected graphs generated by the Hamiltonian

$$\Delta H = \int d^3x \frac{1}{2}(\nabla_i\phi_>)^2 + \frac{1}{2}M\phi_>^2 + \lambda \;\int d^3x \left(\frac{1}{6}\phi_<^3\phi_> + \frac{1}{4}\phi_<^2\phi_>^2 + \frac{1}{4!}\phi_<\phi_>^3\right) \tag{8.14}$$

in which $\phi_<$ plays the role of an external source field.

The key step towards building a perturbative theory is that of computing the unperturbed propagator, i.e. the Green's function of the operator defining the field bilinear, which enters the quadratic part of the Hamiltonian. According to Wick theorem, this propagator can be used to evaluate the two-point correlation function of the hard fields, in the unperturbed theory.

$$\langle \tilde{\phi}_>(k)\tilde{\phi}_<(p)\rangle_{0>} \equiv \frac{\int \mathcal{D}\tilde{\phi}_> \; \tilde{\phi}_>(k)\tilde{\phi}_>(p)e^{-H_0[\tilde{\phi}_>]}}{\int \mathcal{D}\tilde{\phi}_> \; e^{-H_0[\tilde{\phi}_>]}} \tag{8.15}$$

where

$$H_0[\tilde{\phi}_>] = \int_{b\Lambda}^{\Lambda} \frac{d^3k}{(2\pi)^3} \tilde{\phi}^*(k) k^2 \phi(k) \tag{8.16}$$

is the Fourier space expression for the quadratic part of the $H[\phi]$ Hamiltonian. We stress, however, that unlike in the original theory, this expression contains both an IR and UV cut-off, provided by $b\Lambda$ and Λ, respectively. Consequently, the propagator in momentum space reads:

$$\langle \tilde{\phi}_>(k)\tilde{\phi}_<(p)\rangle_{0>} = \frac{(2\pi)^3}{k^2} \; \delta(k+p) \; \Theta(k), \tag{8.17}$$

where

$$\Theta(k) = \begin{cases} 1 & \text{if} \quad b\Lambda < |k| < \Lambda \\ 0 & \text{otherwise.} \end{cases} \tag{8.18}$$

This propagator can be used to evaluate connected loop diagrams which are generated by the ΔH Hamiltonian, order by order in perturbation theory. Here, we consider only the leading-order perturbative correction. We do not evaluate all diagrams which appear at this order, but only the one that appears after expanding to leading order the interaction vertex $\tilde{\phi}^2_>\tilde{\phi}^2_<$, as an illustrative example.

$$\begin{aligned} \int d^3x \frac{\lambda}{4}\phi^2_<(x)\langle \phi_>(x)\phi_>(x)\rangle_{>0} &= \frac{\mu}{2}\int \frac{d^3k}{(2\pi)^3}\tilde{\phi}_<(k)\tilde{\phi}_<(-k) \\ &= \frac{\mu}{2}\int d^3x \frac{\lambda}{4}\phi^2_<(x), \end{aligned} \tag{8.19}$$

where

$$\mu = \frac{\lambda}{2}\int_{b\Lambda}^{\Lambda} \frac{d^3q}{(2\pi)^3}\frac{1}{q^2} \tag{8.20}$$

This calculation shows that to leading order in the perturbative expansion the integration over the hard fields generates an effective correction to the mass (i.e. quadratic) term in the theory of the soft-modes. Thus, removing some hard momentum component fields from the path integral effectively leads to "renormalizing" the mass term. Physics is being shuffled from explicit components in Hilbert space to effective parameters. Students are encouraged to evaluate other diagrams which follow from different terms. In particular, to convince themselves that new vertexes can appear in the effective theory for $\phi_<$ which are not present in the original (i.e.

"bare") Hamiltonian. Including all such corrections, the effective Hamiltonian can be written as follows:

$$H_{eff}[\phi_<] = \int d^3x \left[\frac{1}{2}(1+\Delta Z)(\nabla_i\phi_<)^2 + \frac{1}{2}(M+\Delta M)\phi_<^2 \right.$$
$$\left. \frac{1}{4}(\lambda-\Delta\lambda)\phi_<^4 + \Delta C(\nabla_i\phi_<)^4 + \Delta D\phi_<^6 + ...\right] \tag{8.21}$$

In this expression ΔC and ΔD parametrize the new vertexes generated by the integration over the hard fields, while $\Delta Z, \Delta\lambda$, and ΔM are corrections to terms which are present also in the bare Hamiltonian. The dots represents terms which appear only to the next order of perturbation theory and possibly include new vertexes. Ultimately, this procedure generates all vertexes compatible with the overall symmetry of the bare Hamiltonian, as we anticipated in our preliminary discussion on RG (see Prologue). We stress again that each of these coefficients can be evaluated order-by-order in perturbation theory and that, due to the simultaneous presence of an IR and UV cut-off, all calculations are divergence free.

In the end, up to higher corrections in perturbation theory, we have obtained two equivalent expression of the original partition function:

$$Z_\Lambda \simeq Z_{b\Lambda} \tag{8.22}$$

The subfix Λ and $b\Lambda$ have been introduce to emphasize the fact that the two expression for the partition function are defined using two different UV cut-off scales, Λ a $b\Lambda$ respectively.

8.1 Rescaling Procedure

The partition functions Z_Λ and $Z_{b\Lambda}$ are supposed to represent the same physics, up to higher order perturbative corrections. On the other hand, they have different cut-offs. To make a more careful comparison between these two theories, it is convenient to rescale the momentum and length scale in order to match the UV cut-off of the two theories. To this end, we perform the rescaling transformation:

$$k' = k/b \qquad x' = bx \tag{8.23}$$

Note that the new variable k' is now bonded by the condition $|k'| < \Lambda$. After rescaling, the partition function:

$$Z_{b\Lambda} = \int \mathcal{D}\phi\ e^{-H_{eff}[\phi]} \to Z'_\Lambda = \int \mathcal{D}\phi\ e^{-H'_{eff}[\phi]} \tag{8.24}$$

where

$$H_{eff}[\phi_<] = \int d^3x b^{-3} \left[\frac{1}{2}(1+\Delta Z)\ b^2\ (\nabla_i\phi_<)^2 + \frac{1}{2}(M+\Delta M)\phi_<^2 \right.$$
$$\left. \frac{1}{4}(\lambda-\Delta\lambda)\phi_<^4 + \Delta C\ b^4\ (\nabla_i\phi_<)^4 + \Delta D\phi_<^6 + ...\right] \tag{8.25}$$

Then, we rescale the fields in order for the kinetic energy part to look like in the bare Hamiltonian:

$$\phi'(x) = [\frac{1}{b}(1+\Delta Z)^{1/2}\phi(x) \tag{8.26}$$

Then, the decimated expression for the partition function finally reads:

$$Z'_\Lambda = \int \mathcal{D}\phi \; e^{-\mathcal{H}[\phi]} \tag{8.27}$$

$$\begin{aligned} \mathcal{H}_{eff}[\phi_<] = \int d^3x \Big[& \frac{1}{2}(\nabla_i\phi_<)^2 + \frac{1}{2}M^{2'}\phi_<^2 \\ & \frac{1}{4}\lambda'\phi_<^4 + \Delta C' \; (\nabla_i\phi_<)^4 + D'\phi_<^6 + ... \Big] \end{aligned} \tag{8.28}$$

where the new effective coefficients are defined as follows:

$$M^{'2} = (M^2 + \Delta M^2)(1+\Delta Z)^{-1}b^{-2} \tag{8.29}$$

$$\lambda' = (\lambda + \Delta\lambda)(1+\Delta Z)^{-2}b^{-1} \tag{8.30}$$

$$C' = (C + \Delta C)(1+\Delta Z)^{-2}b^{3} \tag{8.31}$$

$$D' = (D + \Delta D)(1+\Delta Z)^{-3}b^{-3} \tag{8.32}$$

$$.. \quad ... \tag{8.33}$$

where the dots represent additional coefficient which may appear at higher order in the perturbative expansion.

Through the decimation and rescaling procedure, we have defined a continuous transformation in the space of Hamiltonians[2]. This transformation is expected to leave the physics unchanged, at least to the specific order in perturbation theory adopted to deriving the corrections $\Delta Z, \Delta M^2, \Delta C, ...$ This means that we can choose to compute correlations function using any of the theories related by a renormalization group transformation of the Hamiltonian, and we expect to obtain the same answer, up to small perturbative corrections.

8.2 Relevant, Irrelevant and Marginal Operators

The effective parameters in the theories related by RG transformation may be very different. To analyze more closely how these parameters vary under the RG transformation let us here consider a theory defined in the vicinity of the origin in parameter space, i.e.

$$(M^2, \lambda, C, D, \ldots) = (0.0.0, 0.\ldots) \tag{8.34}$$

This choice is done to ensure the applicability of the perturbative treatment. The corresponding Hamiltonian is then simply

$$H_0 = \frac{1}{2}\int d^3(\nabla_i\phi)^2 \tag{8.35}$$

In the vicinity of the point (8.34) we can ignore terms $\Delta M^2, \Delta\Lambda, \ldots$ and consider only the leading (linear) terms, leading to the very simple transformation law:

$$M^{'2} = M^2 b^{-2}, \quad \lambda' = \lambda b^{-1}, \quad C' = Cb^3, \quad D' = D \tag{8.36}$$

---

[2] For historic reasons, this set of transformation is usually referred to as the renormalization group, although it is not really a group, since it lacks the inverse element.

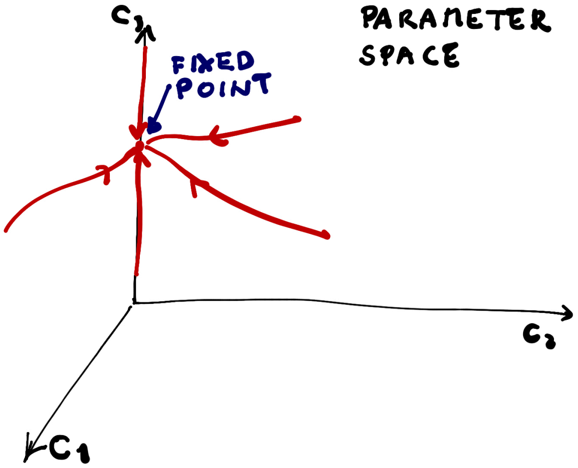

Figure 8.1: Graphical representation of a fixed point in RG flow

.

Now, we recall that $b < 1$, which implies that some of these coefficients grow under RG flow, while other are suppressed. Finally, we note that the D coefficient is left invariant (these results specifically apply for 3 dimensions, in different dimensions the relationships may The terms in the Hamiltonian whose coefficient grow under RG flow are call *relevant operators*, while those whose coefficient decrease are referred to as *irrelevant*. Finally, the terms with coefficients that remain unchanged under RG are referred to as *marginal*.

In general the relevance, irrelevance or marginality of an operator can be inferred by simple dimensional analysis. Namely, in D dimensions, the coefficient of an operator with N powers of ϕ and M derivatives transforms as

$$C' = b^{N(D/2-1)+M-D} C_{N,M} \tag{8.37}$$

The evolution of the effective parameters under RG can be regarded as dynamical system in the multi-dimensional space of effective theory's parameters. Stability points in this dynamical states identify particular sets of parameters, which are approached by RG flow starting from very different point of parameter space (see Fig. 7.1). These are called *fixed-points* and the set of theories evolving to the same fixed-point form an universality class.

Chapter 9

Polymer Field Theory

The statistical physics of macromolecular systems involving polymers of different type and possibly subject to geometrical or topological constraints has been a very active field or research since pioneering work of Edwards in the 60's. This is perhaps in view of the countless implications of this subject: indeed, not only polymers are the main constituents of glues and plastic, but also life itself is based on polymeric biomolecules: RNA, DNA and proteins.

In the 70's De Gennes and collaborators developed one of the most beautiful theory of modern physics, which established and exploited highly non-trivial formal connections between classical statistical models of polymer with quantum field theories of several types. This way, many general properties of polymeric systems could be investigated using powerful analytic tools such as path integrals, along with renormalisation group and scaling concepts. This nobel-winning work has opened the way to an entire new branch of theoretical physics, known as polymer field theory.

The purpose of this chapter is to introduce the very basic elements of this theory and in particular to illustrate the usefulness of field theory techniques in this context. For a more exhaustive recent review on this subject see, e.g. [17]. For application of polymer field theory to biomolecules and in particular proteins, see e.g. [**?**].

9.1 Gaussian Chains

The canonical partition function for a single chain in solution can be written as follows[1]:

$$Z = \mathcal{N} \prod_{k=1}^{N} d^3\mathbf{r}_k \prod_{l=1}^{N-1} g(\mathbf{r}_{l+!}, \mathbf{r}_l)\; e^{-\beta H[R]} \tag{9.1}$$

In this equation, $R = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ is a point in the polymer configuration space, while the function $g(\mathbf{r}_{l+!}, \mathbf{r})$ introduces constraints which ensure the chain topology, and $H(R)$ contains all other interactions, for example those responsible for excluded volume, hydrophobic effective attraction, or hydrophilic effective repulsion. The

---

[1] In the statistical mechanics of polymers, one is usually not interested in the kinetic energy. Therefore, we assume that the kinetic contribution to the canonical partition function factorizes and cancels out in the statistical averages. What we refer to the system's Hamiltoninan is in fact a potential energy.

simplest model for a single polymeric chain is one in which the beads are connected by harmonic springs, while all other interactions are neglected.

In the spirit of effective theory, to derive this model by analyzing the Fourier transform of $g(\mathbf{r}_{l+!}, \mathbf{r}_l)$:

$$\tilde{g}(\mathbf{k}) = \int d\mathbf{y}\ e^{i\mathbf{k}\cdot\mathbf{y}}\ g(\mathbf{y}, \mathbf{0}) \tag{9.2}$$

We now rewrite $\tilde{g}(\mathbf{k})$ in terms of a function $f(\mathbf{k})$ defined by:

$$\tilde{g}(\mathbf{k}) \equiv e^{-f(\mathbf{k})} \tag{9.3}$$

Finally, we assume to study the polymer thermodynamics at law resolution, where $f(\mathbf{k})$ can be expanded in Taylor series:

$$\tilde{g}(\mathbf{k}) \equiv\ e^{-(c_0 + c_2\ \mathbf{k}^2 + \ldots)} \tag{9.4}$$

where the odd-terms have been removed assuming the rotational invariance of the system and $c_2 > 0$.

After transforming back to configuration space, we obtain the so-called Gaussian chain model, defined by the partition function

$$\begin{aligned} Z &= \mathcal{N} \prod_{k=1}^{N} d^3\mathbf{r}_k\ e^{-\beta H_0[R]} \qquad \text{with} && (9.5)\\ H_0 &= \frac{\kappa}{2} \sum_{i=1}^{N-1} (\mathbf{r}_{i+1} - \mathbf{r}_i)^2. && (9.6) \end{aligned}$$

From this point on, we shall implicitly assume that energy is measured in units $1/\beta$, i.e. set $\beta = 1$. In the limit in which the number of beads in the chain becomes infiinite, we can turn the sum of harmonic springs in the exponent into an integral.

$$H_0 = \int_0^N ds\ \frac{\alpha}{2} \dot{\mathbf{r}}^2(s) \tag{9.7}$$

where s is a continuous curvilinear abscissa which replaces the monomer index. The continuous version of the Gaussian chain model is sometimes referred to as the Edwards-Helfand Hamiltonian.

The average distance b between consecutive monomers depends on the spring constant and on the chain dimensionality. Therefore, it is customary to eliminate α in terms of these two parameters and obtain:

$$H_0 = \frac{d}{b^2} \int_0^N ds\ \frac{1}{2} \dot{\mathbf{r}}^2(s) \tag{9.8}$$

In the following, we focus on the case of chains in 3 dimensions, so the partition function reads:

$$Z = \mathcal{N} \int \mathcal{D}\mathbf{r}\ e^{-\frac{3}{b^2} \int_0^N ds\ \frac{1}{2} \dot{\mathbf{r}}^2(s)} = \int d\mathbf{r}_f \int d\mathbf{r}_0\ G(\mathbf{r}_f, N|\mathbf{r}_0) \tag{9.9}$$

where $G(\mathbf{r}_f, N|\mathbf{r}_i)$ reads

$$G(\mathbf{r}_f, N|\mathbf{r}_0) = \mathcal{N} \int_{\mathbf{r}_0}^{\mathbf{r}_f} \mathcal{D}r \; e^{-\frac{3}{b^2}\int_0^N ds \; \frac{1}{2}\dot{r}^2(s)} \tag{9.10}$$

and can be interpreted as the conditional probability of having the chain endpoint at position $\mathbf{r}_f$, provided that the other endpoint is at $\mathbf{r}_0$.

Looking at the path integral expression of $G(\mathbf{r}_f, N|\mathbf{r}_0)$ one immediately recognizes a formal analogy with the imaginary time Feynman propagator of a free quantum particle. Therefore, $G(\mathbf{r}, N|\mathbf{r}_0)$ is Green's function of the free Schrödinger operator:

$$\left[\frac{\partial}{\partial s} - \left(\frac{b^2}{6}\right)\; \nabla^2_{\mathbf{r}}\right] \; G(\mathbf{r}, s|\mathbf{r}_0, 0) = i\delta(s)\delta(\mathbf{r}_f - \mathbf{r}_0) \tag{9.11}$$

The solution of this equation is given by:

$$G(\mathbf{r}, s|\mathbf{r}_0, 0) = \left(\frac{3}{2\pi \; b^2 N}\right)^{3/2} \; e^{-3\frac{(\mathbf{r}_f - \mathbf{r}_0)^2}{2b^2 N}} \tag{9.12}$$

In the next chapter we shall discover yet another formal mapping, this time between the imaginary-time quantum propagation and Brownian diffusion. On the basis of such a duality, the Green's function $G(\mathbf{r}, s|\mathbf{r}_0, 0)$ can be viewed as the conditional probability of a freely randomly diffusing particle (cfr. Eq. 3.60). For this reason, this simple model is sometimes referred to as a Brownian chain.

This analogy can be exploited to compute the mean square distance between the chain's endpoint. By using the random walk Green's function (9.12) we obtain:

$$\langle(\mathbf{r_f} - \mathbf{r}_0)^2\rangle = \frac{1}{3} b^2 N \tag{9.13}$$

Which reflects famous Einstein's law for random walk. Thus, this simple model yields the scaling law $\langle(\mathbf{r_f} - \mathbf{r}_0)^2\rangle \propto N^\alpha$ with $\alpha = 1/2$. A typical problem of polymer physics consists in determining the scaling exponent α for more complicated models. In the following we discuss a field theoretic solution of one of these models, in which a local pairwise interaction is introduced in order to account for excluded volume.

9.2 Self-Avoiding Chain

Let us now consider a more realistic model in which the beads in the chain have a short-range interactions which are repulsive at short-distance (preventing any pair of them to be at the very same location) and possibly attractive as some intermediate distance. This request can be implemented by adding new terms to the Gaussian chain Hamiltonian:

$$H = \frac{3}{b^2}\int_0^N ds \; \frac{1}{2}\dot{\mathbf{r}}^2(s) + \frac{1}{2}\int_0^N ds' \int_0^N ds \; V_2[s, s'; \mathbf{r}(s), \mathbf{r}(s')] + ... \tag{9.14}$$

where ... denote possible three-body correlations. Here we consider the simplest choice for $V_2(\mathbf{r}, \mathbf{r}')$, which is of course a delta-function:

$$V_2(\mathbf{r}, \mathbf{r}') \equiv v\delta(\mathbf{r}_1 - \mathbf{r}_2). \tag{9.15}$$

We emphasise that the partition function contains a nonlocal pair-wise interactions. In the dual picture based on the mapping with quantum propagation in imaginary time, such a non-locality is equivalent to a retardation effect. Indeed, shall now see that, in complete analogy with quantum field theory, locality of interactions can be recovered introducing auxiliary fields which "mediate" the interaction.

The most instructive way to do so is by introducing auxiliary fields into the path integral and using the techniques to evaluate Gaussian path integrals discussed in the previous chapter. We begin by introducing a density field into the path integral defining the partition function [2] by means of an appropriate functional delta:

$$\begin{aligned} Z &= \mathcal{N}\int \mathcal{D}\rho \int \mathcal{D}\mathbf{r}\; e^{-\frac{3}{b^2}\int_0^N ds\;\frac{1}{2}\dot{\mathbf{r}}^2(s)-\frac{v}{2}\int_0^N ds\int_0^N ds'\int d^3\mathbf{x}\int d^3\mathbf{y}\delta(\mathbf{x}-\mathbf{r}(s))\;\delta(\mathbf{y}-\mathbf{r}(s'))\;\delta(\mathbf{x}-\mathbf{y})} \\ &\quad \delta\left[\rho(\mathbf{x})-\int_0^N ds\delta(\mathbf{x}-\mathbf{r}(s))\right] \\ &= \mathcal{N}\int \mathcal{D}\rho \int \mathcal{D}\mathbf{r}\; e^{-\frac{3}{b^2}\int_0^N ds\;\frac{1}{2}\dot{\mathbf{r}}^2(s)-\frac{v}{2}\int d^3\mathbf{x}\rho^2(\mathbf{x})}\;\delta\left[\rho(\mathbf{x})-\int_0^N ds\delta(\mathbf{x}-\mathbf{r}(s))\right] \end{aligned} \tag{9.16}$$

Next, the functional delta is represented in terms of its Fourier integral, by introducing one more functional integral over the Fourier-conjugate field $\phi(\mathbf{x})$:

$$\delta\left[\rho(\mathbf{x})-\int_0^N ds\;\delta(\mathbf{r}(s)-\mathbf{x})\right]=\int \mathcal{D}\phi e^{i\int d\mathbf{x}\;\phi(\mathbf{x})\rho(\mathbf{x})-i\int_0^N ds\phi(\mathbf{r}(s))} \tag{9.17}$$

As before, the partition function is completely determined by the two-point correlation function $G(\mathbf{r},N|\mathbf{r}_0)$, which now reads

$$G(\mathbf{r},N|\mathbf{r}_0)=\int \mathcal{D}\rho\int \mathcal{D}\phi\; e^{-\int d^3\mathbf{x}\left[\frac{v}{2}\rho(\mathbf{x})^2-i\phi(\mathbf{x})\rho(\mathbf{x})\right]}\int_{\mathbf{r}_0}^{\mathbf{r}} \mathcal{D}\mathbf{r}\; e^{-\frac{3}{b^2}\int_0^N ds\;\left(\frac{1}{2}\dot{\mathbf{r}}^2(s)+i\phi(\mathbf{r}(s))\right)} \tag{9.18}$$

The integral over the $\rho(\mathbf{x})$ field is now Gaussian and can be carried out analytically using the Green's function technique developed in the previous chapter. Up to an irrelevant normalization factor we obtain:

$$G(\mathbf{r},N|\mathbf{r}_0)=\int \mathcal{D}\phi\; e^{-\frac{1}{2v}\int d^3\mathbf{x}\phi(\mathbf{x})^2}\int_{\mathbf{r}_0}^{\mathbf{r}} \mathcal{D}\mathbf{r}\; e^{-\frac{3}{b^2}\int_0^N ds\;\left(\frac{1}{2}\dot{\mathbf{r}}^2(s)+i\phi(\mathbf{r}(s))\right)} \tag{9.19}$$

The term

$$G_{[\phi]}(\mathbf{r},N|\mathbf{r}_0)=\int_{\mathbf{r}_0}^{\mathbf{r}} \mathcal{D}\mathbf{r}e^{-\frac{3}{b^2}\int_0^N ds\;\left(\frac{1}{2}\dot{\mathbf{r}}^2(s)+i\phi(\mathbf{r}(s))\right)} \tag{9.20}$$

can be interpreted as an imaginary-time Feynman's path integral describing a quantum particle interacting with a complex external potential $i\phi(\mathbf{x})$. Such a Green's

---

[2] From this point on we shall no longer keep track of the irrelevant normalization multiplicative factor $\mathcal{N}$ in the expression of the partition function.

function corresponds to diffusion of a particle in an imaginary (random) external potential ϕ, i.e.

$$\left[\frac{\partial}{\partial s} - \left(\frac{b^2}{6}\right)\ \nabla_x^2 + i\phi(x)\right]\ G_{[\phi]}(x,s|r_i,0) = \delta(s)\delta(x-r_i) \tag{9.21}$$

Therefore, so far we have shown that the excluded volume effects can be *exactly* represented by a stochastic average over an external field $\phi(\mathbf{x})$.

Using the coherent state path integral representation introduced in the previous chapter it can be written in path integral form as[3] :

$$\begin{aligned} G(\mathbf{r},N|\mathbf{r}_0) &= \int \mathcal{D}\phi e^{-\frac{1}{2v}\int d^3\mathbf{x}\phi^2(\mathbf{x})} \int \mathcal{D}\psi^*\mathcal{D}\psi\ \psi(\mathbf{r},N)\ \psi^*(\mathbf{r}_0,0) \\ &\quad e^{-\int_0^N ds \int d^3\mathbf{x}\ \psi^*\left(\frac{\partial}{\partial s} - \frac{b^2}{6}\ \nabla^2 + i\phi(\mathbf{x})\ \right)\ \psi} \end{aligned} \tag{9.22}$$

Finally, we integrate over the Gaussian $\phi(\mathbf{x})$ field, which leads to the following final expression:

$$G(\mathbf{r},N|\mathbf{r}_0) = \int \mathcal{D}\psi^*\mathcal{D}\psi\ \psi(\mathbf{r},N)\ \psi^*(\mathbf{r}_0,0)\ e^{-\int_0^N ds \int d^3\mathbf{x}\ \left[\psi^*\left(\frac{\partial}{\partial s} - \frac{b^2}{6}\ \nabla^2\ \right)\ \psi - \frac{v}{2}\int_0^N ds' |\psi(\mathbf{x},s)^2|\psi(\mathbf{x},s')|^2\right]} \tag{9.23}$$

Thus, we have obtained the expression for the partition function of a self-avoiding chain as a propagator in a non-relativistic ϕ^4 theory.

9.2.1 Breakdown of the perturbative approach

The most natural starting point to analyse the effect of excluded volume effects is to solve the theory in the perturbative limit of small coupling v. Explicit calculation of $\langle \mathbf{R}^2 \rangle$ in this limit yields (for an arbitray number of dimensions d):

$$\langle \mathbf{R}^2 \rangle = Nb^2\ \left[1 + \frac{4}{3}(vb^2N)^{(4-d)/2} + \ldots\right] \tag{9.24}$$

Calculations show that perturbative corrections always come with powers of $v\sqrt{N} \ll 1$. Since $vN = v\sqrt{N}\ \sqrt{N} \to \infty$, perturbative approach breaks down for any dimension smaller than 4, including the physically relevant $d = 3$ case. The breaking of perturbation theory reflects the fact that the perturbative expansion parameter is not v itself, rather the combination $v\sqrt{N}$ which is never small, in the continuous limit model we are considering.

---

[3] The careful reader might have noticed that in this coherent path integral expression we have dropped the overall boundary-term phase factor $e^{i\mathcal{L}(t)}$ which is due to the over-completeness of the coherent state basis (cfr. Eq. (1.85)). The justification for this is the following. Let us imagine to expande this exponent to an arbitrary order in powers of the fields. To all orders in perturbation theory Wick contractions with this term will give rise to diagrams in which at least one line propagates backwards in time. However, in non-relativistic quantum mechanics (both in real and imaginary time, such diagrams vanish because of causality. For the same reason we have dropped the path integral associated to the normalisation of the propagator (sum over loop diagrams)

Therefore we need to tackle the case in which excluded volume effects are not small, and evaluate the propagator (9.23) within a fully non-perturbative approximation scheme. A standard approach consists to perform the saddle-point expansion, which is equivalent to a mean-field approximation. Imposing the stationarity of the exponent with respect to variations of the ψ^* we obtain a time-dependent Gross-Pitaevski equation:

$$\left(\frac{\partial}{\partial s} - \frac{b^2}{6}\ \nabla^2 + v|\psi(\mathbf{x},s)|^2\right)\psi(\mathbf{x},s) = 0 \tag{9.25}$$

This self-consistent equation can be solved with standard iterative schemes. Variation with respect to ϕ yields the complex conjugate of this equation.

9.3 Grand-Canonical Ensemble of Many Chains

In the previous section we have discussed the thermodynamics of a single chain in solution. Let us now generalise to the case in which the system consists of a fixed number n of polymers. The partition function for this system is

$$Z(n) = \frac{e^{-\frac{Nn}{2}V[0]}}{n!(\lambda_T^3)} \prod_{j=1}^{n} \int \mathcal{D}\mathbf{r}\ e^{-\sum_{k=1}^{n} \frac{d}{b^2} \int_0^N ds\ \frac{1}{2}\dot{\mathbf{r}}_k^2(s) + \sum_{k,l=1}^{n} \frac{1}{2} \int_0^N ds' \int_0^N ds\ V[\mathbf{r}_k(s) - \mathbf{r}_j(s')]} \tag{9.26}$$

where the exponent in the pre-factor has been introduced for convenience.

We can now apply the bosonization scheme deviced for one chain. The result of introducing auxiliary fields through appropriate functional deltas and then integrating out Gaussian fields is equivalent to exploiting the so-called Hubbard-Stratonovich transformation:

$$\begin{aligned} &\exp\left[-\frac{v}{2}\sum_{kl}\int_0^N ds \int_0^N ds' \delta\left(\mathbf{r}_k(s) - \mathbf{r}_k(s')\right)\right] \\ &= \mathcal{N}\int \mathcal{D}\phi \exp\left[-\frac{1}{2v}\int d^3\mathbf{x}\ \phi^2(\mathbf{x}) - i\int_0^N ds\ \phi(\mathbf{r}(s))\right] \end{aligned} \tag{9.27}$$

The n polymer partition function can then be written as:

$$Z(n) = Z_0 \int \mathcal{D}\phi \exp\left[-\frac{1}{2v}\int d^3\mathbf{x}\ \phi^2(\mathbf{x})\right] Q^n[i\phi], \tag{9.28}$$

where $Q[i\phi]$ is the single chain partition function in the background of the external field $i\phi$ and

$$Z_0 = \frac{1}{n!}\left(\frac{e^{\frac{N}{2}V(0)}}{\lambda^{3N}}\right)^n \tag{9.29}$$

Now we can finally consider the grand-canonical partition function. For sake of clarity we now restore the explicit temperature dependence:

$$\Omega(\mu,\beta) = \sum_{n=0}^{\infty} e^{\beta\mu n} Z_0 \int \mathcal{D}\phi \exp\left[-\frac{1}{2v}\int dx\phi^2\right] Q^n[i\phi] \tag{9.30}$$

$$= \sum_{n=0}^{\infty} \int \mathcal{D}\phi \exp\left[-\frac{1}{2v}\int dx\phi^2\right] \frac{(z'\ Q[i\phi])^n}{n!}, \tag{9.31}$$

$$= \int \mathcal{D}\phi \exp\left[-\frac{1}{2v}\int dx\phi^2\right] \exp[z'Q] \tag{9.32}$$

where μ is the chemical potential and z' is the so-called chain activity

$$z' = \frac{e^{\beta\mu+\frac{N}{2}V(0)}}{\lambda^{3N}} \tag{9.33}$$

Let us now rewrite the grand-canonical partition function

$$\sum_{n=0} \frac{(z'\ Q[i\phi])^n}{n!} = \sum_{n=0} \frac{\left[\ z'\ \int d\mathbf{r}_i \int d\mathbf{r}_f G_{[i\phi]}(\mathbf{r}_f, N|\mathbf{r}_i, 0)\right]^n}{n!} \tag{9.34}$$

Recalling that

$$G_{[\phi]}(\mathbf{r}_f, N|\mathbf{r}_i) = \langle 0|\psi(\mathbf{r}_f, N)\psi^\dagger(\mathbf{r}_i, 0)|0\rangle_{i\phi} \tag{9.35}$$

we find

$$\sum_{n=0} \frac{(z'\ Q[i\phi])^n}{n!} = \frac{\int \mathcal{D}\psi^*\mathcal{D}\psi e^{-H_o(i\phi)+\sqrt{z'}\int d^3\mathbf{r}(\psi^*(\mathbf{r},0)+\psi(\mathbf{r},N))}}{\int \mathcal{D}\psi^*\mathcal{D}\psi e^{-H_o(i\phi)}} \tag{9.36}$$

The latter equality can verified by inspection by expanding to all orders the second exponent: any term with odd number of fields disappears, while all terms with an even number give raise to powers of correlations functions...

Finally, completing the Gaussian integral over the field ϕ we find our final expression:

$$\Omega(\mu) = \int \mathcal{D}\psi^* \int \mathcal{D}\psi\ \ \exp(-\mathcal{H}[\psi^*, \psi]) \tag{9.37}$$

where the functional $\mathcal{H}$ reads

$$\begin{aligned} \mathcal{H} &= \int ds \int d^3\mathbf{x} \left[\psi^*(\mathbf{x}, s)\ \left(\partial_s - \frac{b^2}{6}\nabla^2\right)\ \psi(\mathbf{x}, s) + \frac{v}{2}\ |\psi(\mathbf{x}, s)|^4\right] \\ &\quad -\sqrt{z'}\int d^3\mathbf{r}(\psi^*(\mathbf{r}, 0) + \psi(\mathbf{r}, N)) \end{aligned} \tag{9.38}$$

Part III

Appendix

Appendix A

Basic Chemistry of Biopolymers

A.1 Introduction

Typical molecules of biological importance contain a number of atoms which can range form 10^3 — for some small poly-peptides—- to 10^4 for typical globular proteins, up to 10^7 or larger for RNA and DNA strands, membranes and so on. Such large molecules play a fundamental role in all living organisms. Hence, understanding their dynamics would have potentially infinite biomedical implications.

From a physical point of view, RNA and proteins are (bio)-polymers, i.e. molecules made by the several sub-units called monomers, which are linked together to form a chain. In the case of proteins and RNA at room temperature, the chain folds to form a well-defined three-dimensional structure, called the native state.

The monomers forming the chain of such macromolecues are kept together and linked to one another by chemical bonds which can never be spontaneously broken, at room temperature. In addition, there are long range Coulombic correlations, and short range steric effects, hydrogen bonds, sulfur bridges, and so on.

Like most other strongly correlated systems, biopolymers exhibit critical behavior, i.e. collective phenomena which drastically change the physical properties of the system and take place around a well defined temperature For example, above the so-called unfolding temperature, proteins and RNA swell and assume coil configurations. it is impropriate however to speak about a true phase transition, since macromolecules are not infinite systems. In fact, biopolymers are typically considered as prototypes of mesoscopic systems. This adjective implies that the effect associated to their finite size are generally important and cannot be neglected.

The presence of strong correlations and significant finite-size effects make the theoretical modeling of macromolecules a challenging task. In fact, the strong correlations prevent from using perturbative approaches, while finite size effects make it impossible to rely on the thermodynamical limit. As a consequence of such difficulties, modern approaches to the study of the physics of macromolecules must rely on computers. In this context, the continuous developments of new theoretical tools in statistical mechanics, the progress in our understanding of the intra-atomic interactions are deeply intertwined with the development of new algorithms and new computing machines. This combination makes this a field very intriguing and keeps it in continuous evolution.

This course provides a brief introduction to the fundamental tools of statistical mechanics and of computational physics which are used to model the dynamics and thermodynamics of macromolecules. The rest of this chapter contains an elementary introduction of the biochemical and biophysical aspects of macro-molecules, and in particular of proteins and RNA, which will be the main subject of our study. In the subsequent chapters will introduce elements of equilibrium and out-of-equilibrium statistical mechanics and discuss how the relevant equations can be put on a computer and solved. The last chapter is devoted to the presentation some example of applications of numerical simulations to model physical aspects of biopolymers.

A.2 Interactions inside Macromolecules

As a first step, let us analyze the interactions between the various atoms present in a macromolecular system. At the microscopic level there are only Coulombic interactions between electrons an nuclei. However, such a microscopic approach is at present out of reach as it would require to solve a Schrodinger Eq. with a prohibitively large number of electrons. Hence, chemists and physicists have instead analyzed interactions at a more macroscopic level: they introduce semi-empirical interactions which can then be included in energy minimizations, molecular dynamics, Monte Carlo calculations, etc... Along these lines, one may distinguish two main types of interactions:

A.2.1 Bonded

These are the covalent bonds between the atoms of the molecule. their energy range is from 50 kCal/mole to 150 kCal/mole. They correspond typically to 100 $k_B T$ at room temperature, and therefore cannot be broken by thermal fluctuations.

A.2.2 Non Bonded

These are non-covalent interactions between the atoms in the molecule. Their energy range is from 1 to 5 kCal/mole. They are thermally excited at room temperature, and are thus responsible for the dynamics and observed dynamical and thermodynamical properties of macromolecules.

There are several types of non-bonded interactions:

- *Coulomb*:

  The electronic structure of an atom inside a group of other atoms in a molecule is distorted with respect to the electronic distribution of the same in vacuum. As a result, while atoms in vacuum are neutral, in the empiric force fields they are assigned partial charges q_j (smaller than one electronic charge). Consequently atoms interact through Coulomb interaction

  $$v_C(\mathbf{r}) = \frac{q_i q_j}{\epsilon} \frac{1}{r} \tag{A.1}$$

  The question of the relative dielectric constant is a matter of debate and depends on the specific model. For two atoms in bulk water the dielectric constant is about 80 times larger than in vacuum.

- *Van der Waals*:

  This interaction accounts for the strong steric repulsion at short distances, and the dipolar attraction at larger distances. It is usually represented by a Lennard-Jones 6-12 potential of the form:

  $$v_{LJ}(\mathbf{r}) = 4\epsilon \left(A\frac{\sigma^{12}}{r^{12}} - B\frac{\sigma^{6}}{r^{6}} \right) \tag{A.2}$$

- *Hydrogen bonds*:

  The hydrogen bond between two atoms is the result of having one hydrogen atom... The interaction responsible for the formation of H-bonds can be introduced explicitly by a 6-10 potential similar to the Lennard-Jones potential, but it is now quite accepted that H-bonds are just a result of the combination of Coulomb and Van der Waals interactions.

- *Interactions with water:*

  Water is a dipolar molecule, and thus has strong interactions with charged or dipolar groups. Since proteins are active in an aqueous environment, the interaction with water must be taken into account. This is the origin of the hydrophobic effect: hydrophobic amino-acids such as *phenylalanine, triptophan, valine, and leucine* are usually buried inside the globule, whereas polar or charged groups such as *lysine, arginine* are located at the surface of the globule, in contact with water. The molecular basis of the hydrophobic effect is still a subject of investigation. Nevertheless, it is generally believed to be entropic in origin, with an important role played by the entropy change of the solvent.

  The contribution to the total free energy of the folding coming from the packing of the hydrophobic groups is, in general, positive i.e. unfavorable, both from the enthalpic and entropic point of view. On the other hand, it is believed that water molecules are locally ordered around non-polar residues, leading to cage-like structures, associated to a network of hydrogen bonds. The burying of non-polar residues inside the globule reduces the amount of local order of water molecules around the protein, hence it leads to entropy production. This contribution is believed to be the dominant one in the total solvation free energy.

A.3 Fast and Slow Degrees of Freedom

In view of the above classification of the interactions, one can distinguish two types of degrees of freedom in macromolecule:

- *Hard degrees of freedom*: these are the covalent bonds (linking covalently two atoms along the chain), the valence angles (angle between two covalent bonds) and the peptide bond. These degrees of freedom are very rigid at room temperature, since, as we shall see later, their deformation requires energies much higher than the thermal energy kT .

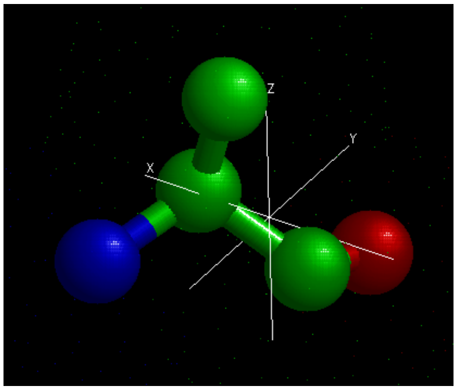

Figure A.1: The ball-and-stick representation of a simple amino acid (GLY)

- *Soft degrees of freedom:* they are essentially the torsion angles along the backbone chain (usually denoted with ϕ and ψ), and of the side chains (usually denoted with $\omega_1, \omega_2, ...$ Their energy scale is such that they can easily ßuctuate at room temperature (see Fig. A.3).

The characteristic oscillation frequencies ω_i of the hard degrees of freedom is such that, at room temperature, $\hbar\omega_i$ is comparable with (actually larger than) the thermal energy $k_B T$. Hence, the dynamics of such set of degrees of freedom is expected to be sensitive to quantum effects. On the other hand, the soft degrees of freedom have typical energies much smaller than the thermal energy and therefore, for them, a classical description is appropriate.

In first approximation, it seems natural to consider all the bonded interactions as frozen, hence implying that the primary structure is quenched so that only the non-bonded terms in the inteaction have to be taken into account.

A.4 Proteins

Proteins are biological molecules, present in any living organism. Their biological function includes catalysis (enzymes), transport of ions (hemoglobin, chlorophyll, etc...), muscle contraction, ... They also are present in virus shells, prions, etc.

The chemical reactions which are relevant for the biological activity of a protein take place in the so-called "active site" of the molecule, which is usually formed by a small subset of its atoms.

Proteins belong to the group of ıbiopolymers, which also comprise nucleic acids (DNA, RNA) and polysaccharides. In general biopolymers are heteropolymers, i.e. they are made out of different species of monomers. In particular, in proteins, the monomers are the 20 different amino-acids.

A.4.1 Amino-acids

The chemical formula of an amino-acid can be written as:

$$NH_2 - C_\alpha HR - COOH \tag{A.3}$$

The NH_2 part is called the *amine* group, the $COOH$ part is the *acidic* group[1].

Each amino acid is characterized by its residue R. The list of amino acids (along with their common three-letter label) is: Alanine (ALA), Isoleucine, Leucine (LEU), Methionine, Phenylalanine, Proline (PRO), Tryptophan (TRP), Valine (VAL), Asparagine (ASP), Cysteine (CYS), Glutamine (GLU), Glycine (GLY), Serine (SER), Threonine (THR), Tyrosine (TYR), Arginine (ARG), Histidine (His), Lysine (LYS), Aspartic acid, Glutamic acid.

For instance, the chemical formula of ALA is:

$$NH_2 - C_\alpha H - \mathbf{CH_3} - COOH \tag{A.4}$$

and that of tryptophan is:

$$NH_2 - C_\alpha H - \mathbf{CH_2} - \mathbf{C} - \mathbf{CH} - \mathbf{NH} - \mathbf{C_6H_4} - COOH. \tag{A.5}$$

In both formulas the bolded atoms are those forming the residue. The smallest residue is glycine, which is just a single H atom, and the largest is tryptophan, which contains ten heavy (i.e. non-hydrogen) atoms.

The typical size of a protein ranges from approximately 100 amino acids for small proteins to 500 for long immuno-globulins.

Amino acids can be classified according with their electrostatic properties:

- Neutral, non-polar:
  Alanine, Isoleucine, Leucine, Methionine, Phenylalanine, Proline, Tryptophan, Va- line.
- Neutral, polar:
  Asparagine, Cysteine, Glutamine, Glycine, Serine, Threonine, Tyrosine.
- Positively charged:
  Arginine, Histidine, Lysine.
- Negatively charged:
  Aspartic acid, Glutamic acid.

Charged residues are always hydrophilic, because of the large polarizability of water, while non-polar neutral residues are hydrophobic.

A.4.2 Polycondensation and peptidic bonds

The chemical reaction leading to the binding of the amino-acids into a protein is called *poly-condensation*, which can be schematically written as:

$$\begin{aligned} &NH_2 - C_\alpha HR_1 - COOH + NH_2 - C_\alpha HR_2 - COOH \rightarrow \\ &\rightarrow NH_2 - C - HR_1 - CONH - C_\alpha HR_2 - COOH + H_2O \end{aligned} \tag{A.6}$$

The poly-condensation produces a Òpeptide bondÓ $CONH$, represented in Fig. A.2. Due to electronic hybridization, this bond is strongly planar.

The repetition of this process produces a polypeptide chain. Proteins are therefore polypeptide chains whose chemical composition is entirely characterized by the sequence of residues $R_1, R_2, ..., R_N$. The sequence of C, C_α and N atoms connected by peptide bonds in the polypeptide chain is called the *backbone* of the chain.

---

[1] With the exception of proline, which has an *imine* group instead

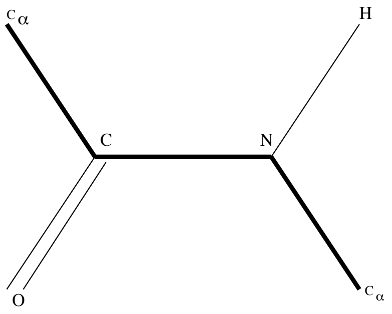

Figure A.2: Peptide Bond. The thick line denotes the back-bone

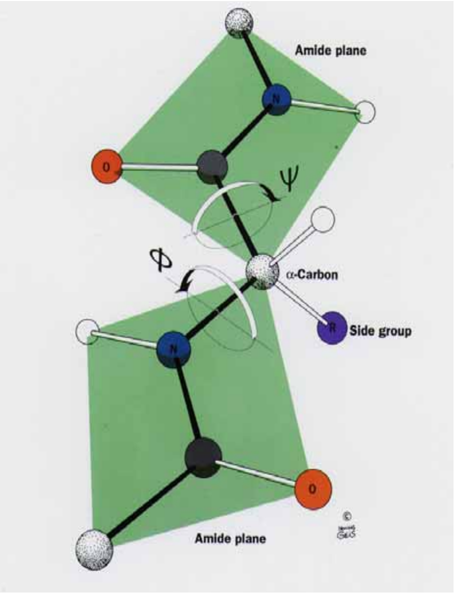

Figure A.3: The torsional angles Φ and Ψ along a peptide bond.

A.5 Nucleid Acids

Another important group of biopolymers are the nucleid acids, which play a fundamental role in cells and in viruses as they carry genetic information. The most common nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Along with them, there are also artificial nucleic acids, which include peptide nucleic acid (PNA), Morpholino and locked nucleic acid (LNA), as well as glycol nucleic acid (GNA) and threose nucleic acid (TNA). Each of these polymers is distinguished from naturally-occurring DNA or RNA by changes to the backbone of the molecule.

From a physical-chemical point of view, nucleic acid is a macromolecule composed of chains of monomeric nucleotides. Nucleotides are therefore the structural units of RNA and DNA and consist of three joined structures: a nitrogenous base, a sugar, and a phosphate group. The most common nucleotides can be divided into two groups —purines and pyrimidines— based on the structure of the nitrogenous base. The joined sugar is either ribose or deoxyribose.

Nucleotides can be synthesized with both purine and pyrimidine as bases. In DNA, the purine bases are *adenine* and *guanine*, while the pyrimidines are *thymine* and *cytosine*. RNA uses uracil rather than thymine (thymine is produced by adding a methyl to uracil). The nucleotide passes through numerous biochemical steps while being processed, adding and removing atoms through the use of numerous enzymes.

Bibliography

[1] P. Lepage, *How to Renormalize the Schrödinger equation* arXiv:nucl-th/9706029.

[2] R.P. Feynman, Rev. Mod. Phys. 20, 367 (1948)

[3] Kleinert, Hagen. Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets (4th ed.). Singapore: World Scientific (2004).

[4] L.S. Schulman. "Techniques and Applications of Path Integration". New York: John Wiley and Sons (1981)

[5] J.W. Negele and H. Orland, "Quantum Many-Particle Systems", Westview Press, 1997.

[6] E. Schneider, S. a Beccara and P.Faccioli, Phys. Rev. **B 88**, 085428 (2013).

[7] E. Schneider and P. Faccioli, Phys. Rev. **B** 89, 134305 (2014)

[8] E. Schneider, S. a Beccara, F. Mascherpa,, P. Faccioli, Phys. Rev. **B 94**, 014306 (2016).

[9] R. Elber and D. Shalloway, J. Chem. Phys. **112** 5539 (2000)

[10] S. Tanase-Nicola and J. Kurchan, J. Stat. Phys. **116**, 1201, (2004)

[11] M. Le Bellac, "Quantum and Statistical Field Theory", Oxford Science Publications, 1991.

[12] C. Itzykson and J.B. Zuber "Quantum Field Theory", McGraw Hill (New York, 1980)

[13] M.E. Peskin and D.V. Schroeder "An Introduction to Quantum Field Theory", Perseus Books (Cambridge, Massachusetts, 1995);

[14] F. Schwabl "Statistical Mechanics", Springer Verlag, 2006

[15] P. Hänggi, P. Talkner and M. Borkovec, Rev. Mod. Phys. 62, 251 (1990).

[16] P. Lepage, *How to Renormalize the Schrödinger equation* arXiv:nucl-th/9706029.

[17] T.A. Vilgis, Phys. Rep. 336, 167 (2000)

[18] C. Jarzynski, Phys. Rev. Lett. 78 2690 (1997).

[19] G. E. Crooks, Phys. Rev. **E** 60, 2721 (1999)

[20] F. Noe and E. Rosta, J. Chem. Phys. **151**, 190401 (2019).

[21] S. Orioli and P. Faccioli, J. Chem. Phys. **145**, 124120 (2016)

[22] J.P. Blaizot, D. De Boni, P. Faccioli and G. Garberoglio, Nucl. Phys. A 946, 49 (2016)

[23] W. E, W. Ren and E. Vanden Ejden, Chem. Phys. Lett. **413** 242 (2005).

[24] E. Facco, M. d'Errico, A. Rodriguez and A. Laio, Scientific Reports **7**, 12140 (2017).

[25] A. Rodriguez, A. Laio, Science **344**,1492 (2014).

[26] K. Sekimoto, Prog. Theor. Phys. Suppl. 130, 17 (1998).

[27]] K. Sekimoto, Stochastic Energetics, Lecture Notes inPhysics Vol. 799 (Springer, New York, 2010).