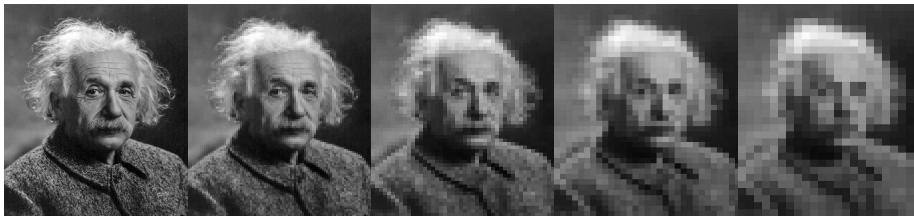# Theory of Coarse-Graining
Vol 1: Descriptions with a finite number of variables

Pep Español

January 14, 2017

"It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience"

A. Einstein in *"On the Method of Theoretical Physics"*, The Herbert Spencer Lecture, delivered at Oxford (10 June 1933); also published in Philosophy of Science, Vol. 1, No. 2 (April 1934), pp. 163-169., p. 165

Contents

Preface

In *Impression: soleil levant*, Claude Monet used a few colors and stains to capture the state of a seaport sunset by dismissing all the precise details of a realistic description of the Le Havre harbor.

The behaviour of the world around us is notoriously complex. Take, for example, two people discussing the aesthetical value of Monet's *Impression*. The reasoning and emotions felt by both individuals are the result of an extremely complex network of processes that involves, to name a few: the articulation of words, the propagation of sound in air, the interaction of sound with the tympanic membrane, the transduction of the oscillations of the membrane into electric impulses transmitted through nerves that reach the brain, the activation of millions of different neurons in the brain. Those in the neocortex are the responsible for abstract reasoning and language but deep down the brain, millions of processes take place simultaneously to keep the rest of the body functioning, from organs to tissues, to cells. Cells are like small specialized chemical factories where thousands of chemical reactions take place, while water crosses pores in cell membranes and ATP consuming proteins transport cellular material from one side to the other of the cell. Water, proteins, lipids, hydroxilapatite in bones, and carbohidrates like glucose make most of the constitutent of the body. The order required over the approximatelly 10^{28} atoms that pile up to form the bodies of these two people as they talk seems a miracle.

Human intelligence tries to grasp this vast complexity by focusing on very partial aspects of the world, and using highly simplifying assumptions about them. The method of **Coarse-Graining** is an attempt to capture part of the richness of the physical world in a simplified way, much in the spirit of an impressionist picture representing the real world. The term coarse-grain probably arises from the field of photography, where in

the early days, the size of the colloidal particles of the photosensitive silver emulsion in the photographic film would determine the final quality of the image. Coarser grains in the emulsion lead to poorer, although at some times more evocative, images. In modern digital photography, the equivalent to "size of the grains of silver" is the size of a pixel. A given image may be represented with different number of pixels of different sizes, leading to smaller information content with larger pixel sizes, as can be seen in the cover picture of Albert Einstein. By following this analogy of photographic resolution, a given physical system made of microscopically interacting particles may be described with different resolutions, each one characterized by a different set of variables. The Theory of Coarse-Graining aims at obtaining how these variables evolve in time in order to predict the evolution of the system.

The Theory of Coarse-Graining presented in this book is **a theory of theories**. It is a general framework in which many different theories of matter fit in. Thermodynamics, coarse-grained models for complex molecules, diffusion, hydrodynamics, elasticity, kinetic theory, chemical reactions, etc. are all theories about the behaviour of matter that have the same structure in terms of reversible and irreversible terms, conserve energy and increase entropy, thus respecting the First and Second Laws of Thermodynamics. The key concept of the theory is the separation of time scales or, equivalently, the notion of quasi-equilibrium, and all the mentioned theories rely on this notion. The realization of the unifying principles behind all these theories is a remarkable scientific achievement. Once these principles are grasped, along with the underlying assumptions and approximations that are required for the structure to hold, it is relatively easy to generate new theories applicable to systems not considered before.

Theory of theories

The aim of these two volumes is to present the structure of the Theory of Coarse-Graining and to show how fruitful it is, by describing a relatively large number of applications of the theory to particular descriptions of different systems. In Volume 1, we will consider descriptions involving a finite number of CG variables. In Volume 2 we will address systems described with fields. In Vol 1 we are concerned with fluctuations while in Vol 2 only the average behaviour of fields will be considered.

The Theory of Coarse-Graining is also referred to as Non-Equilibrium Statistical Mechanics and sometimes it is referred to as the Mori-Zwanzig formalism. The theory is presented by using the technique of projection operators. It should be clear, though, that the theory can be derived without any projection operator at all. The essential physics is not in the projector, but in the way the statistical information is represented approximately, through the so called quasi-equilibrium ensemble also known as relevant ensemble. The projection operators are a fancy recourse to get the CG equations in a relatively simple way. There are several different projectors in the literature. By following Grabert's textbook, we choose the Kawasaki-Gunton projector, as it is the most general one, including particular versions of the projector used by Robertson, Mori, and Zwanzig.

Projection operators

In this book, we limit ourselves to the realm of Classical Mechanics. The emphasis is on "ordinary everyday systems" (or in modern terms, we focus on Soft Matter) for which there is no need for relativistic nor quantum mechanics, because the velocities of the particles are small enough compared with the speed of light, and the masses of the particles and temperatures involved are large enough for quantum effects to be neglected. The secret reason for not dealing with Quantum Mechanics is that I do not yet understand it well enough to talk about it.

Classical Mech

Deductive approach

The approach taken in this book is deductive. In this sense it has a distinct Statistical Mechanics flavour, instead of a Thermodynamic one. In Statistical Mechanics, one takes for granted the microscopic laws governing the evolution of microscopic degrees of freedom and derives the dynamic equations for macroscopic information. A Thermodynamicist would proceed by simply postulating the macroscopic equations with due respect to symmetries, and general principles (like the First and Second Laws). However, much more can be learned if we do not forget what are the underlying building blocks of matter. In particular, we may benefit from computer simulations of the microscale to infer properties of the macroscale that would not be apparent in a purely thermodynamicist point of view. Despite of this deductive approach, there are several steps in the derivation that require approximations of very difficult rigorous justification. We do not claim mathematical rigor and are sufficiently happy with a physicist style that favors intuitive reasoning over ϵ, δ mathematical proofs.

Conforming with this deductive approach, we aim at producing a "recipe for coarse-graining", this is, a well-defined number of steps that, no matter the system one considers, lead to its macroscopic description. This "coarse-graining for dummies" approach is headed to success *only* if the user has a deep understanding of the "Physics" that is going on in the system, in particular if she is able to determine what variables obey the separation of time scales assumption. By proceeding in this way, one can clearly separate the two steps needed when addressing the problem of the non-equilibrium dynamics of a system: A first step where intuitions about the system are generated much like an *artist* dreams her piece of art, followed by a second step which is the construction of the equations, where one takes the noble role of an *artisan*. Both steps, each one taken by a different side of the brain are necessary for successfully solving the problem.

No predictions

As we are interested in the structure of different macroscopic and mesoscopic theories for matter, we limit ourselves to the construction of the coarse-grained equations of motion, but we do not solve them, nor explore its dynamic behaviour. Therefore, we do not make predictions with the theories obtained, nor validate the theory by comparing these predictions with experimental or simulation results. These crucial steps in the scientific method should be addressed, of course, but it is not the focus of the present book. Many of the theories discussed are very old (like kinetic theory, elasticity, hydrodynamics) and their validity has been established after many years of scientific research, at least in certain limits. Other theories discussed in the book do deserve further attention in this respect. Because the dynamic equations are usually non-linear, the predictions can only be made by solving numerically the coarse-grained dynamic equations. In some instances, like in hydrodynamics, we may describe how to implement the numerical solution of these equations. The theory of coarse-graining can illuminate the process by which the discrete equations required by a computer in order to simulate a field theory can be also directly obtained from the microscopic dynamics.

Equil and Non-Equil StatMech

The presentation of the theory gives a unified approach of equilibrium and non-equilibrium statistical mechanics. The former is just a particular case of the latter, aimed at describing the very late state of the evolution of a system, where only the actual dynamic invariants of the system are needed for the description of the system. In order to deal with both "kinds" of statistical mechanics at the same time, we have decided to use dimensionless quantities for entropy and thermodynamic potentials. In particular, we do not include the Boltzmann constant k_B in the definition of the entropy.

The reason for including the k_B factor in the definition of the entropy is historical, not physical, as it cames from trying to match results obtained from the entropy and from the ideal gas equation of state containing the ideal gas constant. We should stress that the theory presented is not limited to situations near equilibrium. As it will become clear, violent non-equilibrium phenomena like shocks and turbulence can be described with the equations derived with the notion of quasi-equilibrium, the founding concept of the theory of coarse-graining.

A very large number of books exist that deal directly or indirectly with the subject of the present one, some of them are referenced in the Reference List. My particular thinking style has enjoyed enormously the books "Thermodynamics" by Callen with its emphasis in clarifying the structure of Thermodynamics, "Non-Equilibrium Thermodynamics" by De Groot and Mazur addressing the myriad of different applications of Linear Irreversible Thermodynamics and, of course, "Projection operators techniques in Nonequilibrium Statistical Mechanics" by Grabert, where I learned almost everything I know about the micro to macro connection. I have also been strongly influenced by the GENERIC approach of Öttinger and Grmela since my Aha! experience reading the original papers. The book "Beyond Equilibrium Thermodynamics" by Öttinger describes extensively the GENERIC framework. Given the stature of these works, I have pondered about the usefulness of one additional book on the subject of Non-Equilibrium Statistical Mechanics. Yet, it is my belief that a comprehensive presentation of the theory of coarse-graining may help many students and researches to get the fundamental and basic ideas that are behind so many different theories for the macroscopic behaviour of systems with many degrees of freedom. Although most part of the material of this book is well-known to specialists, the perspective offered should help in understanding when one should expect a CG model to work or fail. In any case, I felt personally compelled to write this book by my desire to fully understand the subject. In the very process of writing I have learned much more than I was expecting. It is my great hope that the reader will learn also something new.

[References to prior work] [As compared with Grabert's textbook from where many results have been drawn, we have payed special attention to the role of dynamic invariants and, in particular, of the total energy. This is due to the more recent understanding of the GENERIC structure of the equations.]

[New result: I think that the transition probability result regarding irreversibility is new.]

[Vol1 deals with fluctuations, Vol 2 deals with averages. Averages are only useful when fluctuations are small, in such a way that the equation for the average is "representative" of what any particular realization of the system is doing.]

Influential books

Notation and conventions

We follow a number of conventions and notation along this book. Words in **bolface** denote important concepts introduced for the first time. Words "in quotes" denote concepts that are not well-defined, although may convey valuable intuition.

Functions of microstates z in phase space, that is of coordinates and momenta of the particles of the sytem are denoted with a hat as in $\hat{A}(z)$. We follow this convention except for probability densities, as in $\rho_t(z)$. Operators acting on functions of phase space are denoted with calligraphic symbols, as in $i\mathcal{L}, \mathcal{P}$. Vectors and matrices are not denoted with boldfaces, except when they are in the physical space of $\mathbb{R}^3$. We use a number of superscripts for matrices. M^T is the transpose of M, $M^S = (M + M^T)/2$ is the symmetric part of the matrix M while $M^A = (M - M^T)/2$ is the antisymmetric part. We follow Einsteins summation convention in which, for example the product of two matrices in components is written as

$$(AB)_{\mu\nu} = A_{\mu\nu}B_{\nu\sigma} = \sum_{\nu} A_{\mu\nu}B_{\nu\sigma} \qquad (0.0.1)$$

Operators on phase functions act on everything on its right, except if delimiters are placed. For example, the action of the Liouville operator (a first order differential operator) on two phase functions is

$$i\mathscr{L}\hat{A}\hat{B} = \hat{B}i\mathscr{L}\hat{A} + \hat{A}i\mathscr{L}\hat{B} \qquad (0.0.2)$$

while

$$\left(i\mathscr{L}\hat{A}\right)\hat{B} = \hat{B}i\mathscr{L}\hat{A} \qquad (0.0.3)$$

The derivative of a composite function is expressed as

$$\frac{\partial}{\partial x}F(g(x)) = \frac{\partial F}{\partial g}(g(x))\frac{\partial g}{\partial x}(x) \qquad (0.0.4)$$

Acknowledgments

Part I

Theory

1

The Microscopic Level of Classical Mechanics

1.1   Classical versus Quantum mechanics

According to the Standard Model, which is the widely accepted model for the constituents of matter at a fundamental level, the physical world is made of elementary particles (electrons, photons, quarks, etc.) that interact with the four basic forces of Nature (strong, weak, electromagnetic, and gravitation). Under the extreme conditions found in some places of the Universe, like the centers of galaxies or in particle accelerators laboratories, where particles are created and destroyed, one needs to describe Nature with Quantum Field Theories, where different particles are the quantum excitations of fields. However, at low energies where the speed of the particles involved is much smaller than the speed of light, the basic quantities are "particles" (protons, neutrons, electrons) that do not change its nature or number during the observation time. At these low energies, Nature is fairly well described with Quantum Mechanics. The basic forces that governs the dynamics of these particles is electromagnetism and gravitation. The strong forces between quarks confine them into protons and neutrons and the residual forces between quarks, make protons and neutrons to attract each other and form stable nuclei. The nuclei are globally charged positively and attract electrons to form atoms. As long as we do not consider situations like nuclear reaction where the nuclei change their nature we may forget about strong and weak forces. For the behaviour of systems here on the Earth surface, in low energy labs, all we need to know to get a fairly good description of matter behaviour is that it is made of charged massive nuclei and electrons, forming atoms that interact with gravity and electromagnetism. The atoms may arrange, as a result of their specific electric charge distribution, in a huge variety of forms forming different types of chemical bonds (ionic, covalent, van der Waals, metallic, etc.). In this way, atoms may group to form molecules, crystals, metals, etc. The description of chemical bonding requires the use of Quantum Mechanics. The state of a quantum system is described with the wavefunction, a complex function defined on the space of coordinates of all the particles of the system. The dynamics of the wavefunction is governed by Schrodinger's

equation. The behaviour of quantum particles is very different from that of classical particles. They are subject to the Heisenberg uncertainty principle and, therefore, the notion of a trajectory is missing in the quantum world (except in some interpretations like Bohmian Quantum Mechanics [**?**].) The fact that the mass of the electrons is much smaller than the mass of the nuclei allows one to decouple the dynamics of electrons and nuclei. The Schrodinger equation for electrons and nuclei can be solved in two steps within the Born-Oppenheimer approximation. First one assumes that the nuclei do not move, their positions are just fixed parameters in Schrodinger equation for electron in the "external Coulomb field" of the nuclei. Then one find the eigenvalues of the electronic Hamiltonian. These eigenvalues depend parametrically on the positions of the nuclei. The lowest energy eigenvalue corresponds to the ground state of the system, and receives the name of potential energy surface (PES). The next step in the Born-Oppenheimer approximation is to solve the Schrodinger equation for the nuclei, under the "potential energy" given by the PES.

Born-Oppenheimer

In the Born-Oppenheimer approximation, one needs to solve Schrodinger's equation for the nuclei, which is a partial differential equation defined in a multidimensional space. From a computational point of view, the solution of such an equation is very costly. If the physical volume occupied by the nuclei is covered by grid with M nodes per space dimension, the computational cost scales exponentially as M^{3N} with the number N of nuclei, making any direct solution unfeasible. Special techniques like the Hartree-Fock method and the Density Functional Theory have been devised in order to treat molecular and condensed matter capturing quantum molecular effects with accuracy. Still, the methods are limited to relatively small number of particles. It is for this reason that finding classical approximation to the motion of nuclei through approximate force fields (where the PES is a good candidate, if available) for the interacting nuclei has been, and still is, a subject of intensive research. It is expected that at sufficiently large temperatures, for sufficiently large systems, a classical approximation is sufficiently accurate and we will restrict in this book to Classical Mechanics for the evolution of the atoms.

The question of how and when a system needs to be described with Quantum Mechanics is usually answered with the following rule of thumb: Assume that the particles of the system have their momenta distributed according to a Maxwellian distribution corresponding to the temperature of the system. This implies, from the Heisenberg principle, that they will be described with wavefunctions with a width of the order of $\Delta x \sim \hbar/\sqrt{mk_BT}$. If the forces felt by the particle do not change on the length scale of the width of the wavepacket, then the particle will behave in a classical way. This rule shows that, typically, one needs to incorporate quantum effects when the temperature of the system is very low or when the mass of the particle is very small. At room temperatures, for example, electrons always need to be treated quantum mechanically.

The classical limit

1.2   Separating the world in two

An important decision when studying the behaviour of matter is how to draw the boundary between what constitutes "our system of interest" and what is not. In other words, we need to specify what atoms of the universe are our system. The remaining atoms of the universe are excluded from the description but its effect may still be present. When the portion of the universe that we have selected as our system in fact does not interact with the rest of the universe or, more precisely, its interactions are so weak that they can be neglected as compared with the interactions within the system itself, the system is said to be an **isolated system**. The assumption that our system is isolated allows us to formulate in simpler terms the theory of Statistical Mechanics and, therefore, one criteria to choose (or prepare) the system is that it should be sufficiently isolated from the rest of the Universe.

In a laboratory, however, it is very difficult to have isolated systems, because we need to *contain* the portion of matter under study either with physical containers or with external fields. A "physical container" is the portion of the universe that interacts with our system with short ranged interactions, whereas an "external field" is created by a collection of particles far away from our system that interact with long range interactions (like gravity or electric and magnetic fields).

Physical containers are typically systems composed by a number of particles much larger than the number of particles of the system that they contain. In this way, they affect the system but are largely unaffected by the system. If this was not the case, we need to include the dynamics of the container in the description. A rubber balloon containing a portion of gas is a primary example. In this book, physical containers will be dealt with as if the total system composed by the physical container and the system of interest that it contains, is itself an isolated system. Then the effect of the container, that interacts with the system with short ranged forces, will be typically as producing **boundary forces** on the system of interest. These boundary forces are usually, but not always, expressed as **boundary conditions** to be imposed on partial differential equations. In Vol. 2 we will consider the derivation of boundary conditions describing the interaction of fluids and solid containers.

On the other hand, long range interactions with remote external particles will be dealt in this book as **external fields**. Typically, it is assumed that the fields themselves (or rather the particles creating them) are so minutely affected by the interactions with our system that this effect is neglected. Therefore, external fields affect the system but are not affected by the system. By including external fields that may, in general, depend on time, the Hamiltonian becomes time-dependent and the resulting description becomes slightly more complicated. We will study the effect of external fields in Chapter 4.

We should emphasize that we aim at describing *systems that do not exchange particles with the rest of the Universe.* In other words, our system is *defined* by a given number of particles and once we know what are the particles that make up our system, we may formulate a Hamiltonian once and forever. In this way, we preclude the possibility of our system *exchanging* particles with the rest of the Universe. This may appear as a limitation when discussing problems of phase equilibria, for which a description in which particles are exchanged between systems seems appropriate. While from a phenomenological point of view "our system" may be described by a box in which particles can enter and exit,

a corresponding description of this situation in microscopic terms, i.e. in terms of phase spaces, is extremely awkward. We prefer to have "our system" defined in terms of the particles that make it and not by the region of space that they occupy. Of course the particles may move in space an go to regions outside the limits of the box. We will describe situations of phase equilibrium and phase change not through systems with changing number of particles, but rather in terms of fixed particle numbers that may occupy different regions of space. The appropriate way to deal with these phase change phenomena is through density and concentration fields as we will see in Chap. [??].

1.3   Classical Mechanics and Hamilton's equations

In this book, we will deal with macroscopic systems that can be appropriately described with Classical Mechanics. Classical Mechanics is a theory of *point particles* with definite positions $\mathbf{r}_i$ and velocities $\mathbf{v}_i$. These quantities need to be referred with respect to an inertial reference frame in order for the Newton's Laws to be valid. We will often idealize and refer to atoms and even molecules as point particles, even though they may be composite objects. These particles interact through forces. The functional forms of the forces between the point particles is known only in an approximate way, through model force fields. We will always assume that these force fields are given.

There are many different formulations of the laws of Classical Mechanics which are all equivalent to the original formulation set forth by Newton. Newton's Laws give rise to equations of motion for the positions of classical objects that are differential equations of second order. Their solution requires the knowledge of two initial conditions, given by the positions and velocities of all the particles at the initial time in order to predict the future evolution of the system. As stressed originally by Gibbs, the Hamiltonian description of Classical Mechanics is particularly suited to the formulation of Statistical Mechanics because one of the distinguishing features of Hamilton's equations is that they are first order differential equations. Therefore, they only need one initial condition and this simplifies the specification of statistical considerations. In Hamilton's formulation, the microscopic state $z = (q, p)^T$ of a system of N point particles is given by the collection of all positions $q = (\mathbf{q}_1, \cdots, \mathbf{q}_N)^T$ and momenta $p = (\mathbf{p}_1, \cdots, \mathbf{p}_N)^T$ of the particles. The notation emphasizes that z is a column vector, where T is the transpose. The microscopic state of the system evolves according to Hamilton's equations

$$\dot{\mathbf{q}}_i = \frac{\partial \hat{H}}{\partial \mathbf{p}_i} \tag{1.3.1}$$

$$\dot{\mathbf{p}}_i = -\frac{\partial \hat{H}}{\partial \mathbf{q}_i} \tag{1.3.2}$$

where $\hat{H}(z)$ is the Hamiltonian function assumed, until Chapter 4, to be explicitly independent of time. Typically the Hamiltonian has the form

$$\hat{H}(z) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + \hat{U}(\mathbf{q}_1, \cdots, \mathbf{q}_N) + \sum_{i=1}^{N} V(\mathbf{q}_i) \tag{1.3.3}$$

where the first sum is the kinetic energy of the system, the potential of interaction between particles is $\hat{U}(\mathbf{q}_1, \cdots, \mathbf{q}_N)$ and $V(\mathbf{r})$ is a time-independent external potential. We will always assume that the Hamiltonian is a bounded function from below. Because we can choose an arbitrary constant for the potential function without changing the dynamics, we will assume that $\hat{H}(z) \geq 0$ for all microstates z. The condition of bounded Hamiltonian is a requisite for the existence of a proper equilibrium state [**?**].

Hamilton's equations can be written in compact form as

$$\dot{z}_t = J \cdot \frac{\partial \hat{H}}{\partial z}(z_t) \equiv \hat{v}(z_t) \tag{1.3.4}$$

where we have defined in the last equation the **Hamiltonian vector field** $\hat{v}(z)$ and J is the so called **symplectic matrix** which has the form

$$\begin{pmatrix} 0 & +1_{3N} \\ -1_{3N} & 0 \end{pmatrix}. \tag{1.3.5}$$

where 1_{3N} is the identity matrix of dimension $3N \times 3N$. In the form (1.3.4) Hamilton's equation have the intuitive meaning that in order to determine the "next" value of the microstate, you only have to follow the indications given by the velocity field at the current value of the microstate. In more mathematical parlance, the trajectory z_t in phase space is an integral curve of the velocity field.

As we have already mentioned, Hamilton's equations are first order differential equations that require the knowledge of an initial condition $z_0 = (q_0, p_0)^T$. The ulterior evolution of the microstate, denoted by z_t is a trajectory in the $6N$ dimensional space of all microscopic states z known as **the phase space** Γ of the system. We may introduce the Hamiltonian **flow map** as the function $\hat{\Phi}_t$ that takes the initial microstate z_0 at time $t = 0$ into the final z_t at time t, this is

$$z_t = \hat{\Phi}_t(z_0) \tag{1.3.6}$$

In terms of the flow map, Hamilton's equations can also be written as

$$\frac{d}{dt}\hat{\Phi}_t = \hat{v}(\hat{\Phi}_t) \tag{1.3.7}$$

Note that, conceptually, this is a functional relation between two functions in phase space $\frac{d}{dt}\hat{\Phi}_t$ and $\hat{v}(\hat{\Phi}_t)$, while (1.3.4) is a differential equation for the trajectory z_t.

[**Use a consistent notation for the flow map, by using the initial and final times.**]

The flow map defines a one to one mapping between initial and final microstates. The inverse of the flow map satisfies

$$\left[\hat{\Phi}_t\right]^{-1}(z_t) = z_0 \tag{1.3.8}$$

and it is given by

$$\left[\hat{\Phi}_t\right]^{-1} = \hat{\Phi}_{-t} \tag{1.3.9}$$

Therefore, if the flow map takes a microstate z_0 into z_1, then the inverse flow map takes the microstate z_1 and by running backwards in time, gives z_0.

1.4   Operators and the Hilbert space of functions in phase space

It proves convenient to go to a more abstract level and present the dynamics of microstates in the language of linear operators. This language is familiar from Quantum Mechanics. The esential motivation for doing this is that Hamilton's equations are a set of non-linear ordinary differential equations, and by going to this operator description, we obtain linear equations that are amenable of theoretical treatment. This comes with a price, though, because the resulting equations are now *partial* differential equations. The idea is to consider the collection of all complex functions $\hat{F}(z)$ in phase space as a Hilbert vector space. As mentioned earlier, through this book we will denote with a hat any function of the microscopic state of the system. Each function $\hat{F}$ is regarded as a vector in this infinite-dimensional vector space. It can also be denoted with Dirac notation as $|F\rangle$. One particular important (vector) phase function is $\hat{z}$ that takes any microstate z into $\hat{z}(z) = z$, this is, $\hat{z}$ is the identity function in phase space. We carefuly distinguish with this notation the function $\hat{z}$, which is a vector of the Hilbert space, from the argument z of the function that belongs to phase space. We can also denote this vector function as $|z\rangle$. A Hilbert space is a linear vector field with a notion of a distance. The scalar product in the Hilbert space of phase function is defined as

$$\langle G|F\rangle = \int dz \hat{G}^*(z)\hat{F}(z) \tag{1.4.1}$$

where $\hat{G}^*(z)$ is the complex conjugate of $\hat{G}(z)$.

A crucial operator in the dynamics of the microstates is the **Liouville operator** $i\mathcal{L}$. Note that in this book operators are generally written with calligraphic symbols. The Liouville operator has the explicit form

$$i\mathcal{L} \equiv \sum_i^N \left(\frac{\partial \hat{H}}{\partial \mathbf{p}_i} \frac{\partial}{\partial \mathbf{q}_i} - \frac{\partial \hat{H}}{\partial \mathbf{q}_i} \frac{\partial}{\partial \mathbf{p}_i} \right) \tag{1.4.2}$$

and, therefore, it may be written in several equivalent forms. When it acts on an arbitrary function $\hat{F}(z)$ the above expression is also equal to

$$i\mathcal{L}\hat{F} = -\left(\frac{\partial \hat{H}}{\partial z}\right)^T \cdot J \cdot \frac{\partial \hat{F}}{\partial z} = \hat{v}^T \cdot \frac{\partial \hat{F}}{\partial z} = \frac{\partial}{\partial z}^T \cdot \left(\hat{v}\hat{F}\right) = -\{\hat{H}, \hat{F}\} \tag{1.4.3}$$

In the latest form we have introduced the **Poisson bracket** of two arbitrary functions

$\hat{F}, \hat{G}$ is

$$\{\hat{F}, \hat{G}\} \equiv \left(\frac{\partial \hat{F}}{\partial z}\right)^T \cdot J \cdot \frac{\partial \hat{G}}{\partial z} = \sum_i^N \left[\frac{\partial \hat{F}}{\partial \mathbf{q}_i}\frac{\partial \hat{G}}{\partial \mathbf{p}_i} - \frac{\partial \hat{F}}{\partial \mathbf{p}_i}\frac{\partial \hat{G}}{\partial \mathbf{q}_i}\right] \tag{1.4.4}$$

which is just another function in phase space.

The Liouville operator (1.4.2) is a first order differential operator. We will assume that it satisfies

$$\int_\Gamma dz i \mathcal{L} \hat{F}(z) = 0 \tag{1.4.5}$$

where the integral is taken over all the available phase space. By using the divergence theorem, we have

$$\int_\Gamma dz i \mathcal{L} \hat{F}(z) = \int_\Gamma dz \frac{\partial}{\partial z}^T \cdot \hat{v}(z) \hat{F}(z) = \int_{\partial \Gamma} dS^T \cdot \hat{v}(z) \hat{F}(z) \tag{1.4.6}$$

where $\partial \Gamma$ is the boundary of phase space Γ and dS is a surface vector element. What is the boundary of phase space? The phase space Γ is the space of all the possible values that the positions and momenta of the particles in the system may take. The momenta of the particles can take any value from $(-\infty, +\infty)$ and in the direction of momenta, the boundary is "at infinity". For the positions we usually assume that the particles are confined in a volume V due to some external confining potential or live in a periodic system. Therefore, if the function $\hat{F}(z)$ satisfies

$$\lim_{\mathbf{p}_i \to \pm \infty} \hat{F}(z) = 0 \qquad\qquad \lim_{\mathbf{q}_i \in \partial V} \hat{F}(z) = 0 \tag{1.4.7}$$

where ∂V is the boundary of the the volume, or $\hat{F}(z)$ is periodic in the positions, then we have that (1.4.5) is true. We will always assume that the phase functions that we consider "vanish at the boundary of phase space $\partial \Gamma$". For this type of functions the identity (1.4.5) is always true.

Eq. (1.4.5) implies that for any two phase functions $\hat{F}(z), \hat{G}(z)$ "vanishing at the boundary $\partial \Gamma$", the Liouville operator satisfies

$$\int dz \hat{F}(z) i \mathcal{L} \hat{G}(z) = -\int dz \hat{G}(z) i \mathcal{L} \hat{F}(z) \tag{1.4.8}$$

The adjoint $\mathcal{A}^\dagger$ of an operator $\mathcal{A}$ is defined as

$$\langle G | \mathcal{A} F \rangle = \langle F | \mathcal{A}^\dagger G \rangle \tag{1.4.9}$$

Then Liouville operator is self-adjoint, usually expressed in the form $\mathcal{L}^\dagger = \mathcal{L}$, this is

$$\langle A | \mathcal{L} B \rangle = \langle B | \mathcal{L} A \rangle \tag{1.4.10}$$

The reason to introduce the complex number i in the definition of the Liouville operator

is that in this way the Liouville operator is Hermitian.

Another metric that is usually introduced to define the Hilbert space of functions in phase space is

$$\langle G|F\rangle^{\text{eq}} = \int dz \rho^{\text{eq}}(z)\hat{G}^*(z)\hat{F}(z) \tag{1.4.11}$$

where $\rho^{\text{eq}}(z)$ is a distribution function that satisfies $i\mathcal{L}\rho^{\text{eq}}(z) = 0$. Note that with respect to this metric, the Liouville operator is also self-adjoint

$$\int dz \rho^{\text{eq}}(z)\hat{F}^*(z)i\mathcal{L}\hat{G}(z) = \int dz \rho^{\text{eq}}(z)\hat{G}^*(z)i\mathcal{L}\hat{F}(z) \tag{1.4.12}$$

The importance of the Liouville operator stems from the following observation. The action of the Liouville operator on the identity function $\hat{z}$ is just the Hamiltonian vector field

$$i\mathcal{L}\hat{z} = \hat{v} \tag{1.4.13}$$

Therefore, a complicated nonlinear function $\hat{v}$ is transformed into the action of a linear operator on the identity function. "Non-linear" is difficult, "linear" and "identity" are simple, and this suggest that we can do more things, at least in theory, with the language of the Liouville operator. With the use of the identity (1.4.13), Hamilton's equations (1.3.4) can be written in terms of the Liouville operator in the form

$$\frac{d}{dt}z_t = i\mathcal{L}\hat{z}(z_t) \tag{1.4.14}$$

We may use this form of Hamilton's equation to consider a Taylor expansion of the trajectory z_t around $t = 0$, this is

$$z_t = z_0 + \frac{dz_t}{dt}(0)t + \frac{1}{2!}\frac{d^2 z_t}{dt^2}(0)t^2 + \cdots \tag{1.4.15}$$

We will need all the high order time derivatives, which are given by (1.4.14) and

$$\frac{d^2 z_t}{dt^2} = \frac{d}{dt}i\mathcal{L}\hat{z}(z_t) = \hat{v}^T(z_t)\cdot\frac{\partial i\mathcal{L}\hat{z}}{\partial z}(z_t) = (i\mathcal{L})^2\hat{z}(z_t)$$

$$\vdots$$

$$\frac{d^n z_t}{dt^n} = (i\mathcal{L})^n\hat{z}(z_t) \tag{1.4.16}$$

Therefore, by evaluating all these time derivatives at $t = 0$, (1.4.15) becomes

$$z_t = \hat{z}(z_0) + i\mathcal{L}\hat{z}(z_0)t + \frac{1}{2!}(i\mathcal{L})^2\hat{z}(z_0) + \cdots = \exp\{i\mathcal{L}t\}\hat{z}(z_0) \tag{1.4.17}$$

where we have introduced the exponential operator through the Taylor series

$$\exp\{i\mathcal{L}t\} \equiv \mathcal{I} + i\mathcal{L}t + \frac{1}{2!}(i\mathcal{L}t)^2 + \frac{1}{3!}(i\mathcal{L}t)^3 + \cdots . \tag{1.4.18}$$

where $\mathcal{I}$ is the identity operator. Note that with the use of the Liouville operator, we have obtained the solution of Hamilton's equations! It should be clear, though, that this is just a *formal* solution because the action of the operator is not known explicitly. The mathematical justification of the existence of the exponential of the Liouville operator, i.e. the convergence of the Taylor series (1.4.18) is not trivial. In particular, we know that the trajectory of the microscopic state is highly chaotic and subtle issues may arise as a consequence. Nevertheless, the expression of the solution in terms of the exponential operator captures the essential properties of Hamiltonian dynamics (like unitarity, symplectic, reversibility, etc.). In, particular, the exponential solution offers a practical method to solve Hamilton's equations numerically through the so called symplectic algorithms as discussed beautifully in Tuckerman's textbook [**?**].

Sometimes the notation is oversimplified by writing (1.4.17) in the form

$$z_t = \exp\{i\mathcal{L}t\}z_0 \tag{1.4.19}$$

a notation that we will adopt with the understanding that what we mean is (1.4.17), which in words is "apply the operator $\exp\{i\mathcal{L}t\}$ on the *function* $\hat{z}$, giving as a result another function $\exp\{i\mathcal{L}t\}\hat{z}$, and *then* evaluate this function at the value z_0". The formal solution (1.4.17) also shows that the flow map introduced in (1.3.6) is given by

$$\hat{\Phi}_t = \exp\{i\mathcal{L}t\}\hat{z} \tag{1.4.20}$$

The time derivative of both sides of this equation shows

$$\frac{d}{dt}\hat{\Phi}_t = i\mathcal{L}\hat{\Phi}_t \tag{1.4.21}$$

as can be verified by differentiating the exponential operator (1.4.18) term by term with respect to t.

Because of (1.4.17), the exponential of the Liouville operator receives the name of **evolution operator** and gets denoted with

$$\mathcal{U}_t \equiv \exp\{i\mathcal{L}t\} \tag{1.4.22}$$

This operator satisfies a semigroup property

$$\begin{aligned} \mathcal{U}_0 &= \mathcal{I} \\ \mathcal{U}_t\mathcal{U}_{t'} &= \mathcal{U}_{t+t'} \end{aligned} \tag{1.4.23}$$

It also satisfies

$$\frac{d}{dt}\mathcal{U}_t = i\mathcal{L}\mathcal{U}_t \tag{1.4.24}$$

Because the Liouville operator is Hermitian, the evolution operator $\mathcal{U}_t = \exp\{i\mathcal{L}t\}$ is unitary, i.e. it satisfies

$$\mathcal{U}_t^\dagger \mathcal{U}_t = \mathcal{I} \qquad (1.4.25)$$

This is simple to show because

$$\mathcal{U}_t^\dagger = \exp\{-i\mathcal{L}^\dagger t\} = \exp\{-i\mathcal{L}t\}$$
$$\mathcal{U}_t^\dagger \mathcal{U}_t = \exp\{-i\mathcal{L}t\}\exp\{i\mathcal{L}t\} = \mathcal{I} \qquad (1.4.26)$$

Note that the operator inverse of the evolution operator $\mathcal{U}_t$ is $\mathcal{U}_{-t}$, which is just a reflection of the time reversibility of Hamilton's equations.

In addition, the unitarity of the evolution operator is also a reflection of the incompressibility of phase space. In fact, the matrix of the evolution operator $\mathcal{U}_t$ in an orthonormal basis $\langle \phi_\alpha |$ has unit determinant. The evolved phase function $F(\mathcal{U}_t z)$ can be expressed as in (1.4.32)

$$\hat{F}(\mathcal{U}_t z) = \exp\{i\mathcal{L}t\}\hat{F}(z) \qquad (1.4.27)$$

or in Dirac notation

$$|F\rangle_t = \mathcal{U}_t\,|F\rangle \qquad (1.4.28)$$

If we consider two evolved phase functions and consider its scalar product, we will have

$${}_t\langle G|F\rangle_t = \left\langle G\mathcal{U}_t^\dagger |\mathcal{U}_t F\right\rangle = \langle G|F\rangle \qquad (1.4.29)$$

where in the last identity we have used that $\mathcal{U}_t$ is unitary. But translating back the notation this identity is just

$$\int dz \hat{G}^*(\mathcal{U}_t z)\hat{F}(\mathcal{U}_t z) = \int dz \hat{G}^*(z)\hat{F}(z) \qquad (1.4.30)$$

which can only be true if the Jacobian of the transformation $\mathcal{U}_t z$ is unity. We see, therefore, that the unitary evolution reflects also phase space volume conservation.

Any function in phase space $\hat{X}(z)$ acquires a time dependence once evaluated on the time dependent microstate, this is $\hat{X}(z_t)$. Phase functions evolve because the microstates evolve in time. The time derivative of this function is

$$\frac{d}{dt}\hat{X}(z_t) = \left(\frac{\partial \hat{X}}{\partial z}(z_t)\right)^T \cdot \frac{d}{dt}z_t = \left(\frac{\partial \hat{X}}{\partial z}(z_t)\right)^T \cdot J \cdot \frac{\partial \hat{H}}{\partial z}(z_t) = i\mathcal{L}\hat{X}(z_t) \qquad (1.4.31)$$

This identity allows to refer to the Liouville operator as a "time derivative" operator. The differential equation (1.4.31) has as formal solution

$$\hat{X}(z_t) = \mathcal{U}_t\hat{X}(z_0) \qquad (1.4.32)$$

as can be seen by using (1.4.24). The way to read this equation is as follows. The

operator $\mathcal{U}_t$ acts on the space of the phase functions. Therefore, $\mathcal{U}_t\hat{X}$ is another phase function and $\mathcal{U}_t\hat{X}(z_0)$ is this phase function evaluated at the initial condition z_0. A useful mnemotechnical rule for (1.4.32) is

$$\exp\{i\mathcal{L}t\}\hat{X}(z_0) = \hat{X}(\exp\{i\mathcal{L}t\}z_0) \tag{1.4.33}$$

So the evolution operator "enters inside the function".

An alternative way of understanding (1.4.32) is

$$\hat{X} \circ \hat{\Phi}_0^t = \exp\{i\mathcal{L}t\}\hat{X} = \mathcal{U}_0^t\hat{X} \tag{1.4.34}$$

When we evaluate this function at z_0 we obtain (1.4.32). Eq. (1.4.34) shows that the composition of an arbitrary phase function $\hat{X}$ with the flow map $\hat{\Phi}_t$ (which is itself a phase function parameterized with t) is given by the action of the operator $\exp\{i\mathcal{L}t\}$ on the function $\hat{X}$.

The evolution operator satisfies

$$\exp\{i\mathcal{L}t\}\left(\hat{f}(z)\hat{g}(z)\right) = \left(\exp\{i\mathcal{L}t\}\hat{f}(z)\right)(\exp\{i\mathcal{L}t\}\hat{g}(z)) \tag{1.4.35}$$

This is simple to prove by introducing $\hat{h}(z) = \hat{f}(z)\hat{g}(z)$ as a single function of phase space, for which

$$\exp\{i\mathcal{L}t\}\left(\hat{f}(z)\hat{g}(z)\right) = \exp\{i\mathcal{L}\}\hat{h}(z) = \hat{h}(T_tz) = \hat{f}(T_tz)\hat{g}(T_z)$$

$$= \left(\exp\{i\mathcal{L}t\}\hat{f}(z)\right)(\exp\{i\mathcal{L}t\}\hat{g}(z)) \tag{1.4.36}$$

1.5   Properties of the Hamiltonian flow

1.5.1   The Hamiltonian flow is incompressible

We may think of the flow map (1.3.6) as a coordinate transformation from z to z_t. The Jacobian of this coordinate transformation is, by definition, the following determinant

$$\mathcal{J}_t(z) \equiv \det\left(J_t(z)\right) \tag{1.5.1}$$

where the Jacobian matrix is

$$J_t(z) \equiv \frac{\partial\hat{\Phi}_t(z)}{\partial z} \tag{1.5.2}$$

The time derivative of the Jacobian can be computed by using the identity

$$\mathcal{J}_t(z) = \det\left(J_t(z)\right) = \exp\{\text{Tr}[\ln J_t(z)]\} \tag{1.5.3}$$

where the logarithm of the Jacobian matrix is defined in terms of the Taylor series of the logarithm. One way to prove the above identity is by diagonalizing the Jacobian matrix with an orthogonal matrix. Because the determinant and trace operations are invariant

under such a transformation, we find that

$$\exp\{\mathrm{Tr}[\ln J_t(z)]\} = \exp\left\{\sum_k \ln \lambda_k\right\} = \prod_k \lambda_k \tag{1.5.4}$$

where λ_k are the eigenvalues of J. The last identity is just the determinant of the Jacobian matrix. If we now take the time derivative of the Jacobian, we obtain

$$\begin{aligned}
\frac{d}{dt}\mathcal{J}_t(z) &= \exp\{\mathrm{Tr}[\ln J_t(z)]\}\frac{d}{dt}\mathrm{Tr}[\ln J_t(z)]\} \\
&= \exp\{\mathrm{Tr}[\ln J_t(z)]\}\mathrm{Tr}\left[\left(\frac{d}{dt}J_t(z)\right)J_t^{-1}(z)\right] \\
&= \mathcal{J}_t(z)\mathrm{Tr}\left[\left(\frac{d}{dt}J_t(z)\right)J_t^{-1}(z)\right]
\end{aligned} \tag{1.5.5}$$

Consider the time derivative of the Jacobian matrix

$$\begin{aligned}
\frac{d}{dt}J_t(z) &= \frac{d}{dt}\frac{\partial}{\partial z}\cdot z_t(z) = \frac{\partial}{\partial z}\cdot\frac{d}{dt}z_t(z) = \frac{\partial}{\partial z}\cdot J\cdot\frac{\partial\hat{H}}{\partial z}(z_t(z)) \\
&= J : \frac{\partial^2\hat{H}}{\partial z\partial z}(z_t(z))\frac{\partial z_t(z)}{\partial z} = J : \frac{\partial^2\hat{H}}{\partial z\partial z}(z_t(z))J_t(z)
\end{aligned} \tag{1.5.6}$$

Therefore, by inserting this result into (1.5.7) we obtain

$$\frac{d}{dt}\mathcal{J}_t(z) = \mathcal{J}_t(z)\mathrm{Tr}\left[J : \frac{\partial^2\hat{H}}{\partial z\partial z}(z_t(z))\right] = 0 \tag{1.5.7}$$

because the trace of the full contraction (denoted by the colon :) of a symmetric and antisymmetric matrices vanishes. Because the Jacobian is constant and at $t = 0$ it takes the value 1, then it will always be equal to 1. This result is a special case of the theorem of the integral invariants of Poincaré [**?**]).

Another way to see that the Hamiltonian flow is incompressible is by looking at the Hamiltonian velocity introduced in (1.3.4). This velocity field satisfies

$$\frac{\partial}{\partial z}\cdot\hat{v}(z) = \frac{\partial}{\partial z}\cdot\left(J\cdot\frac{\partial\hat{H}}{\partial z}\right) = J : \frac{\partial^2\hat{H}}{\partial z\partial z} = 0 \tag{1.5.8}$$

The zero comes from the fact that we are contracting a fully antisymmetric matrix J with a fully symmetric one, the Hessian of the Hamiltonian. Eq. (1.5.8) tells us that the flow velocity has null divergence in phase space. This is another reflection of the fact that the flow in phase space is incompressible, or that volumes are conserved by the Hamiltonian dynamics.

1.5.2   Stretching in phase space

Typical Hamiltonian systems describing molecular systems have trajectories in phase space that display the phenomenon of **chaos**. This term refers to the property that if we start with initial conditions that are very similar, the trajectories starting at these initial conditions separate from each other exponentially in time, as we will see below. Therefore, while Hamilton's equations are deterministic, in practice its predictive power is rather limited because any small uncertainty in initial conditions readily explodes and renders the prediction very inaccurate. Although this seems to be an unfortunate feature it is, paradoxically, what makes the statistical methods applicable and what, together the large number of degrees of freedom in any macroscopic system, makes Statistical Mechanics a predictive theory.

The measure of stretching of a flow is given by the Liapunov exponents of the flow, which are defined as follows. Consider a small vector $\Delta z = z_1 - z_0$ in phase space, that goes from the microstate z_0 to z_1. The vector has length $|\Delta z|$. After a time t both points evolve according to the Hamiltonian flow and the vector is transformed as

$$\Delta z(t) = \Phi_t(z_1) - \Phi_t(z_0) = \Phi_t(z_0 + \Delta z) - \Phi_t(z_0) \simeq \nabla \Phi_t(z_0) \cdot \Delta z \qquad (1.5.9)$$

The gradient of the flow determines the stretching of little vectors. Note that the eigenvalue problem

$$\nabla \Phi_t(z_0) w_i = g_i w_i \qquad (1.5.10)$$

has $6N$ eigenvalues g_i and corresponding eigenvectors w_i which depend, in general, on t, z_0. We see, therefore that small vectors in the direction of the eigenvectors, evolve according to

$$\Delta z_i(t) \simeq g_i \Delta z_i \qquad (1.5.11)$$

therefore, the eigenvalues of the gradient of the flow determine the rate at which small vectors in the eigendirections grow. The Liapunov exponents are defined as

$$\sigma_i = \lim_{t \to \infty} \frac{1}{t} \ln |g_i| \qquad (1.5.12)$$

and they depend on z_0. Because the flow is incompressible, we have that $\det |\nabla \Phi_t(z_0)| = 1$. Because the determinant is given as the product of the eigenvalues, we have $\prod_i^{6N} g_i = 1$, implying $\prod_i^{6N} |g_i| = 1$ and, therefore, $\sum_i \sigma_i(z_0) = 0$. Because not all $\sigma_i(z_0)$ can be zero simultaneously (except if the Hamiltonian system is integrable, i.e. having 6N conservation laws), there should be positive and negative Liapunov exponents. This implies that for sufficiently long times a little vector in the eigenvector of positive Liapunov exponent

$$\Delta z_i(t) \simeq e^{\sigma_i t} \Delta z_i \qquad (1.5.13)$$

and the vector in this direction grows exponentially. Therefore, a little volume in phase space spanned by $6N$ vectors evolves and distorts in such a way that the volume is
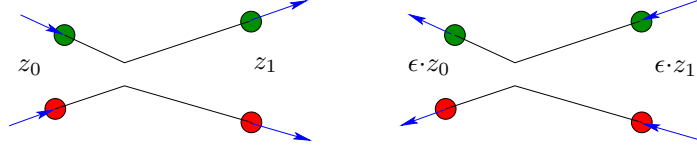
Figure 1.1: Two particles (red and green) interact according to Hamilton's equation of motion. In the left panel the final state after a time t is $z_1 = \hat{\Phi}_t(z_0)$ (which means that $z_0 = \hat{\Phi}_{-t}(z_1)$). In the right panel the initial state is $\epsilon{\cdot}z_1$ (with reversed velocities with respect to z_1). According to time reversibility, the final state should be $\epsilon{\cdot}z_0 = \hat{\Phi}_t(\epsilon{\cdot}z_1)$. Indeed, we have $\epsilon{\cdot}\hat{\Phi}_{-t}(z_1) = \hat{\Phi}_t(\epsilon{\cdot}z_1)$, which is just (1.5.20).

conserved but some directions grow exponentially while others shrink also exponentially. This produces a highly filamentous structure in phase space that, in many aspects, resembles the turbulent flow of an ordinary fluid.

1.5.3 Time reversibility

Hamilton's equations are time reversible equations. In order to grasp the meaning of this statement, it is convenient to introduce the **time reversal operator** $\mathcal{T}$. Its action on an arbitrary function $\hat{F}(z)$ is

$$\mathcal{T}\hat{F}(z) = \hat{F}(\epsilon{\cdot}z) \tag{1.5.14}$$

where ϵ is the matrix

$$\begin{pmatrix} +1_{3N} & 0 \\ 0 & -1_{3N} \end{pmatrix}. \tag{1.5.15}$$

Therefore, the action of the time reversal operator on a phase function is to reverse the sign of the momenta. This operator is an involution, i.e. it satisfies

$$\mathcal{T}^2 = 1 \tag{1.5.16}$$

The Hamiltonian is invariant under time reversal

$$\mathcal{T}\hat{H}(z) = \hat{H}(\epsilon{\cdot}z) = \hat{H}(z) \tag{1.5.17}$$

which implies that the Hamiltonian vector field transforms according to

$$\frac{\partial}{\partial z}\hat{H}(z) = \frac{\partial}{\partial z}\hat{H}(\epsilon{\cdot}z) = \epsilon{\cdot}\frac{\partial \hat{H}}{\partial z}(\epsilon{\cdot}z) \tag{1.5.18}$$

Note that the matrices ϵ, J anticommute

$$\epsilon{\cdot}J = -J{\cdot}\epsilon \tag{1.5.19}$$

This fact has a profound consequence. In fact, we show that the flow map, this is, the solution of Hamilton's equation satisfy the **time reversibility condition**

$$\hat{\Phi}_t(\epsilon \cdot z) = \epsilon \cdot \hat{\Phi}_t^{t_0}(z) \tag{1.5.20}$$

[Check the superscript] In words, what this means is that if we change the sign of the momenta of a microstate z and evolve it in time from t_0 to t, this is the same as evolving the microstate backwards in time from t to t_0, and then reversing the sign of the resulting momenta. The proof of (1.5.20) starts by noticing that it is an identity at $t = 0$. By taking the time derivative

$$\epsilon \cdot \frac{d}{dt} \hat{\Phi}_t(\epsilon \cdot z) = \epsilon \cdot J \cdot \frac{\partial \hat{H}}{\partial z}(\hat{\Phi}_t(\epsilon \cdot z)) = -J \cdot \epsilon \cdot \frac{\partial \hat{H}}{\partial z}(\hat{\Phi}_t(\epsilon \cdot z)) = \left. -J \cdot \frac{\partial \hat{H}}{\partial z'}(\hat{\Phi}_t(z')) \right|_{z' = \epsilon \cdot z}$$

$$= -\frac{d}{dt} \hat{\Phi}_t(\epsilon \cdot z) = \frac{d}{dt} \hat{\Phi}_t^{t_0}(\epsilon \cdot z) \tag{1.5.21}$$

where in the last identity we have used (**??**). We see, therefore, that the time derivative of the left hand side of (1.5.20) coincides with the time derivative of the right hand side. As the expression coincides at $t = t_0$, (1.5.20) is an identity.

1.5.4   Symmetries and conserved quantities

A fundamental theorem due to Emmy Noether ensures that for every symmetry of the Hamiltonian there is a **dynamic invariant**. The trajectory traced by a microstate as it evolves is restricted to live in a submanifold of the full phase space, because of the existence of dynamic invariants. A dynamical invariant $\hat{\mathcal{I}}(z)$ is any dynamical function that does not change in time, this is,

$$\frac{d}{dt} \hat{\mathcal{I}}(z_t) = 0 \tag{1.5.22}$$

The trajectory z_t is, therefore, restricted to be in the submanifold $\mathcal{I}(z) = I_0$ where $I_0 = \hat{\mathcal{I}}(z_0)$ is the value of the dynamical invariants at the initial time. Note that dynamical invariants, by definition satisfy $\mathcal{I}(\mathcal{U}_t z) = \hat{\mathcal{I}}(z)$. Therefore, by taking the time derivative, we conclude that

$$0 = \frac{d}{dt} \mathcal{I}(\mathcal{U}_t z) = i \mathcal{L} \hat{\mathcal{I}}(\mathcal{U}_t z) \tag{1.5.23}$$

and the action of the Liouville operator on dynamic invariants gives zero.

We will not discuss in detail the general Noether's theorem **[cite book]**, but we consider some particular symmetries and invariances. Consider the translation operator $T_{\mathbf{c}}$ that has the following action on a function of space

$$T_{\mathbf{c}} f(\mathbf{r}) = f(\mathbf{r} + \mathbf{c}) \tag{1.5.24}$$

where $\mathbf{c}$ is the constant translation vector. The Taylor series representation of the right

hand side allows one to write the operator in the form

$$T_{\mathbf{c}} = \exp\left\{\mathbf{c}\cdot\frac{\partial}{\partial\mathbf{r}}\right\} = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\mathbf{c}\cdot\frac{\partial}{\partial\mathbf{r}}\right)^n \tag{1.5.25}$$

We may now construct the translation operator in phase space

$$\mathcal{T}_{\mathbf{c}}f(\mathbf{q}_1,\cdots,\mathbf{q}_N,\mathbf{p}_1,\cdots,\mathbf{p}_N) = f(\mathbf{q}_1+\mathbf{c},\cdots,\mathbf{q}_N+\mathbf{c},\mathbf{p}_1,\cdots,\mathbf{p}_N) \tag{1.5.26}$$

that can be written in the form

$$\mathcal{T}_{\mathbf{c}} = e^{-\mathbf{c}\cdot\{\hat{\mathbf{P}},\cdots\}} \tag{1.5.27}$$

where we have introduced the following differential operator

$$\{\hat{\mathbf{P}},\cdots\} = \sum_{i}^{N}\left[\frac{\partial\hat{\mathbf{P}}}{\partial\mathbf{q}_i}\frac{\partial}{\partial\mathbf{p}_i} - \frac{\partial\hat{\mathbf{P}}}{\partial\mathbf{p}_i}\frac{\partial}{\partial\mathbf{q}_i}\right] \tag{1.5.28}$$

In this way, total momentum is said to be the generator of the translation symmetry. Note that if the Hamiltonian of the system is translation invariant, this is

$$\mathcal{T}_{\mathbf{c}}\hat{H}(z) = \hat{H}(z) \tag{1.5.29}$$

this implies that

$$e^{-\mathbf{c}\cdot\{\hat{\mathbf{P}},\cdots\}}\hat{H}(z) = \hat{H}(z) \tag{1.5.30}$$

which can only be true for all $\mathbf{c}$ if

$$\{\hat{\mathbf{P}},\hat{H}\} = 0 \tag{1.5.31}$$

But this expression is nothing but the time derivative of the total momentum, according to Hamilton's equations. Therefore, if the Hamiltonian is invariant under arbitrary translations, total momentum is conserved, and viceversa. This example shows that the quantity that is conserved is the generator of the symmetry.

Consider now rotations. Rotations can be generated by the following operator

$$\begin{aligned}T_{\boldsymbol{\theta}} &= \exp\left\{-\boldsymbol{\theta}\cdot\left(\mathbf{r}\times\frac{\partial}{\partial\mathbf{r}}\right)\right\}\\ &= \exp\left\{(\boldsymbol{\theta}\times\mathbf{r})\cdot\frac{\partial}{\partial\mathbf{r}}\right\}\end{aligned} \tag{1.5.32}$$

which, when applied to a function $f(\mathbf{r})$ have the following effect

$$T_{\boldsymbol{\theta}}f(\mathbf{r}) = f(\mathcal{R}_{\boldsymbol{\theta}}\mathbf{r}) \tag{1.5.33}$$

where the rotation matrix is

$$\mathcal{R}_{\boldsymbol{\theta}} = e^{[\boldsymbol{\theta}]_\times} \tag{1.5.34}$$

and the cross product matrix $[\boldsymbol{\theta}]_\times$ is

$$[\boldsymbol{\theta}]_\times = \begin{pmatrix} 0 & -\theta_z & \theta_y \\ \theta_z & 0 & -\theta_x \\ -\theta_y & \theta_x & 0 \end{pmatrix} \tag{1.5.35}$$

Here, $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$ is the vector of rotation, defining the axis of rotation and whose magnitude gives the rotated angle around the axis.

In order to proof (1.5.33) consider the identity

$$\exp\left\{ (\boldsymbol{\theta} \times \mathbf{r}) \cdot \frac{\partial}{\partial \mathbf{r}} \right\} = \left[\exp\left\{ \left(\frac{\boldsymbol{\theta}}{N} \times \mathbf{r} \right) \cdot \frac{\partial}{\partial \mathbf{r}} \right\} \right]^N \tag{1.5.36}$$

For sufficiently large N we may approximate

$$\exp\left\{ (\boldsymbol{\theta} \times \mathbf{r}) \cdot \frac{\partial}{\partial \mathbf{r}} \right\} \simeq \left[1 + \left(\frac{\boldsymbol{\theta}}{N} \times \mathbf{r} \right) \cdot \frac{\partial}{\partial \mathbf{r}} \right]^N \tag{1.5.37}$$

Consider the effect of one of the products on a function $f(\mathbf{r})$

$$\left[1 + \left(\frac{\boldsymbol{\theta}}{N} \times \mathbf{r} \right) \cdot \frac{\partial}{\partial \mathbf{r}} \right] f(\mathbf{r}) = f(\mathbf{r}) + \left(\frac{\boldsymbol{\theta}}{N} \times \mathbf{r} \right) \cdot \frac{\partial f}{\partial \mathbf{r}}(\mathbf{r}) \simeq f\left(\left[\mathbf{r} + \frac{\boldsymbol{\theta}}{N} \times \mathbf{r} \right] \right)$$

$$\simeq f\left(\exp\left\{ \frac{1}{N}[\boldsymbol{\theta}]_\times \right\} \mathbf{r} \right) = f\left(\mathcal{R}_{\frac{\boldsymbol{\theta}}{N}} \mathbf{r} \right) \tag{1.5.38}$$

and the action of one of the elements of the product is to generate an infinitesimal rotation in the argument of the function in the direction of the vector $\boldsymbol{\theta}$. The action of the N terms of the product gives, naturally, to the full rotation (1.5.33).

As in the case of rotations, we now promote the rotation operator (1.5.33) to an operator acting on phase space functions. Now the role of the linear momentum is played by angular momentum

$$\hat{\mathbf{L}} = \sum_i \mathbf{q}_i \times \mathbf{p}_i \tag{1.5.39}$$

The Poisson bracket of the angular momentum gives the following operator

$$\{\hat{\mathbf{L}}, \cdots\} = \sum_i \left[\frac{\partial \hat{\mathbf{L}}}{\partial \mathbf{q}_i} \cdot \frac{\partial}{\partial \mathbf{p}_i} - \frac{\partial \mathbf{L}}{\partial \mathbf{p}_i} \cdot \frac{\partial}{\partial \mathbf{q}_i^\alpha} \right] = -\sum_i \left[\mathbf{p}_i \times \frac{\partial}{\partial \mathbf{p}_i} + \mathbf{r}_i \times \frac{\partial}{\partial \mathbf{r}_i} \right] \tag{1.5.40}$$

The exponentiation of this operator is another operator that acts on functions defined

on phase space

$$\mathcal{T}_{\boldsymbol{\theta}} = \exp\left\{-\boldsymbol{\theta}\cdot\{\hat{\mathbf{L}}, \cdots\}\right\} \tag{1.5.41}$$

By following the same argument that lead to (1.5.33), the effect of the operator (1.5.41) on a phase space function is

$$\mathcal{T}_{\boldsymbol{\theta}}\hat{F}(\mathbf{q}_1, \mathbf{p}_1, \cdots \mathbf{q}_N, \mathbf{p}_N) = \hat{F}(\mathcal{R}_{\boldsymbol{\theta}}\mathbf{q}_1, \mathcal{R}_{\boldsymbol{\theta}}\mathbf{p}_1, \cdots \mathcal{R}_{\boldsymbol{\theta}}\mathbf{q}_N, \mathcal{R}_{\boldsymbol{\theta}}\mathbf{p}_N) \tag{1.5.42}$$

that is, it rotates all the arguments of the phase function (positions and momenta) around the axis given by the vector $\boldsymbol{\theta}$.

Now, if the Hamiltonian is invariant under rotations, it implies that

$$\mathcal{T}_{\boldsymbol{\theta}}\hat{H}(\mathbf{q}, \mathbf{p}) = \hat{H}(\mathbf{q}, \mathbf{p}) \tag{1.5.43}$$

or, by using (1.5.41)

$$\exp\left\{-\boldsymbol{\theta}\cdot\{\hat{\mathbf{L}}, \cdots\}\right\}\hat{H}(\mathbf{q}, \mathbf{p}) = \hat{H}(\mathbf{q}, \mathbf{p}) \tag{1.5.44}$$

which is only possible if $\{\hat{\mathbf{L}}, H\} = 0$, meaning that the angular momentum is conserved.

In summary, if the Hamiltonian (1.3.3) does not explicitly depends on time, then it generates a dynamics where the Hamiltonian is conserved. If there is no external potential that would break traslation invariance, and the interaction potential depends only on the relative distances between particles, then total linear and angular momentum will also be conserved.

1.6 Uncertainty of initial conditions and Liouville's theorem

Hamilton's equation are first order differential equations governing the deterministic evolution of the system *provided* that the initial condition z is given. However, it is in general impossible to know the precise value of the initial values of the positions and momenta of all the particles in our system. Usually, a system is prepared under identical *macroscopic* conditions that do not allow to fix the value of the positions and momenta of every single particle in the system. A collection of identical systems prepared in an identical, but macroscopic, manner at the initial time will have, in general, different initial microscopic states. For this reason, the best we can do is to express our knowledge about the initial microstate of the system in probabilistic terms by introducing a probability density $\rho_0(z)$ that the system has the microstate z as initial condition. The probability distribution in phase space is usually referred to as an **ensemble**. Even though the evolution of z_t is deterministic, the uncertainty about initial conditions converts the evolution in phase space into a stochastic process. The probability distribution function at a subsequent time is denoted by $\rho_t(z)$ and it describes our uncertainty on the actual microscopic state due to the fact that we could not control the actual microscopic state at the initial time.

What is the functional form of the probability distribution $\rho_0(z)$ at the initial time? This depends very much on what is the macroscopic information that we have about the system at the initial time. In Sec 2.5 we will obtain the initial ensemble with the help of the Principle of Maximum Entropy.

The Liouville theorem

The time-dependent ensemble $\rho_t(z)$ obeys the Liouville's equation. This equation is just an expression of the fact that Hamilton's equations conserve the volume in phase space. The fact that the Jacobian of the evolution is always equal to one has an important consequence on the evolution of the probability density in phase space, as we now show.

Let M be a region of not vanishing measure of Γ and $\mathcal{U}_t M$ the region resulting from the evolution of each point of M according to Hamilton's equations. It is obvious that the probability that the system is in the region M at $t = 0$ is *identical* to the probability of being at $\mathcal{U}_t M$ at $t = t$. For this reason,

$$\int_M \rho_0(z)dz = \int_{\mathcal{U}_t M} \rho_t(z)dz. \tag{1.6.1}$$

By performing the change of variables $z' = \mathcal{U}_{-t}z$ (with unit Jacobian) the integral in the left hand side becomes

$$\int_{\mathcal{U}_t M} \rho_t(z)dz = \int_M \rho_t(\mathcal{U}_t z)dz. \tag{1.6.2}$$

Eq. (4.2.42) and (4.2.43) imply

$$\int_M \rho_0(z)dz = \int_M \rho_t(\mathcal{U}_t z)dz. \tag{1.6.3}$$

This is true for any region M and, therefore, the integrand of the right hand side of (4.2.42) and the left hand side of (4.2.43) should be equal, i.e.

$$\rho_0(z) = \rho_t(\mathcal{U}_t z), \tag{1.6.4}$$

or, by a simple change of variables,

$$\rho_t(z) = \rho_0(\mathcal{U}_{-t} z). \tag{1.6.5}$$

This is the way the density in phase space evolves in time. It simply says that the probability density of microstate z at time t is the same as the one that had the initial condition of z at the initial time. We may of course find a differential equation for the probability density by just taking the time derivative on both sides of (1.6.5). This gives

$$\frac{d}{dt}\rho_t(\mathcal{U}_t z) = \frac{d}{dt}\rho_0(z) = 0. \tag{1.6.6}$$

that, again, expresses the fact that as we move with the flow in phase space, the probability density does not change. Further application of the chain rule leads to the Liouville's equation for the probability density in phase space,

$$\partial_t \rho(z,t) = -i\mathcal{L}\rho(z,t), \tag{1.6.7}$$

where we have used the notation $\partial_t = \frac{\partial}{\partial t}$ to denote the partial derivative with respect to time. It is obvious, by construction, that the formal solution of the Liouville equation (1.6.7) is given by (1.6.5).

   An alternative derivation of the Liouville's equation starts from the realization that the flow in phase space is incompressible, as reflected in (1.5.8). As any other probability density, the probability in phase space satisfies a continuity equation that reflects the fact that probability is locally conserved. The continuity equation is

$$\partial_t \rho_t(z) = -\frac{\partial}{\partial z} \cdot (\hat{v}(z)\rho_t(z)) \tag{1.6.8}$$

By using the incompressibility condition (1.5.8), we obtain again the Liouville's equation (1.6.7).

1.7   Probabilities of functions in phase space

One basic question in probability theory is the following. If we have a probability density over a continuous variable x, what is the probability density that a function $f(x)$ of this variable takes values y? For the time being assume that $x, f(x) \in \mathbb{R}$. The answer is found through the cumulative probability defined as

$$F(y) = \int dx \rho(x)\theta(y - f(x)) \tag{1.7.1}$$

where $\theta(y)$ is the Heaviside step function that takes the value 1 if $y > 1$ and zero otherwise. Here, $F(y)$ is obviously the probability that the function takes values smaller than y. Then the probability density is given as the derivative of the cumulative probability

$$P(y) = \frac{d}{dy}F(y) = \int dx \rho(x)\delta(f(x) - y) \tag{1.7.2}$$

where we have made use of the fact that the derivative of the step function is the Dirac delta function. The probability density $P(y)$ gives the probability that the function takes a value in the interval y_0, y_1 through the integral

$$\int_{y_0}^{y_1} dy P(y) = F(y_1) - F(y_0) \tag{1.7.3}$$

This probability is just the probability that the function takes values smaller than y_1 minus the probability that it takes values smaller than y_0. This argument is easily generalized to the multivariate case.

The relevance of this discussion in Statistical Mechanics stems from the fact that we often enquire about the probability that an observable, described with a function in phase space, has particular values. Assume that $\hat{A}(z)$ is a smooth integrable vector function in phase space and we ask for its probability density $P(a, t)$ given that we know that the microscopic probability density is $\rho_t(z)$. The connection is given by

$$P(a, t) = \int dz \rho_t(z)\delta(\hat{A}(z) - a) \tag{1.7.4}$$

This equation expresses the "macroscopic" probability $P(a, t)$ in terms of the "microscopic" probability $\rho_t(z)$. Through the solution (1.6.5) of Liouville equation, we may express this probability in terms of an average over initial conditions

$$P(a, t) = \int dz \rho_0(z)\delta(\hat{A}(\mathcal{U}_0^t z) - a) \tag{1.7.5}$$

where a change of variables $z' = \mathcal{U}_{-t}z$ (with unit Jacobian) has been done. The origin of the stochasticity at the level of the observables is clearly manifest in the uncertainty on initial conditions.

It is convenient to introduce, in addition to the probability density $\rho_t(z)$ in phase space, the joint probability density $\rho(z_0, t_0; z_1, t_1)$ that gives the probability (density) that the system is around z_0 at time t_0 and around z_1 at time t_1. Because the Hamiltonian dynamics is deterministic, this probability density is given by (for notational simplicity assume $t_0 = 0, t_1 = t$)

$$\rho(z_0, 0; z_1, t) = \rho(z_0, 0)\delta(z_1 - \mathcal{U}_t z_0) \tag{1.7.6}$$

The marginals of the joint probability are the one time probability density

$$\int dz_0 \rho(z_0, 0; z_1, t) = \rho(z_1, t)$$

$$\int dz_1 \rho(z_0, 0; z_1, t) = \rho(z_0, 0) \tag{1.7.7}$$

From the microscopic joint probability we can construct the joint probability density $P(a_0, 0; a, t)$ that a function $\hat{A}(z)$ takes values around a_0 at time $t = 0$ and values around a at time t. By analogy with (1.7.4) the connection between this joint probability and the microscopic joint probability is given by

$$P(a_0, 0; a, t) = \int dz_0 \int dz_1 \rho(z_0, 0; z_1, t) \delta(\hat{A}(z_0) - a_0) \delta(\hat{A}(z_1) - a) \tag{1.7.8}$$

By using the explicit form (1.7.6) we have

$$P(a_0, 0; a, t) = \int dz_0 \rho(z_0, 0) \delta(\hat{A}(z_0) - a_0) \delta(\hat{A}(\mathcal{U}_t z_0) - a) \tag{1.7.9}$$

The fundamental equations (1.7.5), (1.7.9) give the one and two time probabilities of any observable in terms of the microscopic dynamics and the initial ensemble. We will make extensive use of these expressions later on.

For future reference, we introduce the **conditional probability** of having a_1 at time t_1 given that the macrostate was a_0 at t_0 as

$$P(a_0, t_0 | a_1 t_1) = \frac{P(a_0 t_0, a_1 t_1)}{P_0(a_0)} \tag{1.7.10}$$

Note that, concerning notation, there is no agreement on how to denote the conditional probability of a_1 given a_0. Gardiner [**?**] and most mathematically oriented literature, for example, writes this conditional probability as $P(a_1 | a_0)$ while de Groot and Mazur [**?**] denote it by $P(a_0 | a_1)$. We choose the later notation as the "flow" of the events a_0, a_1 is in accordance with the direction of reading.

1.8   The tendency towards equilibrium

Let us consider now the final state predicted by the Liouville equation. A basic mathematical question is under which conditions the Liouville's equation (1.6.7), which is a first order partial differential equation, leads to a stationary solution with $\partial_t \rho(z, t) = 0$. That this is not generally the case can be seen by considering an initial distribution of the form $\rho(z, 0) = \delta(z - z_0)$ that expresses that we know with certainty that the initial state is z_0. In this case, we know that the solution is given by $\rho(z, t) = \delta(\mathcal{U}_t z_0 - z)$, this is, the distribution function remains peaked at the solution of Hamilton's equations. There is no *broadening* of the distribution function and the system does not reach a stationary state. However, if the dynamics generated by the Hamiltonian is highly unstable (i.e. chaotic), we may expect that any non-delta initial distribution will evolve with a sort

of filamentous structure that, eventually, lead to an effective broadening of the distribution and to an *effective* stationary microscopic probability $\rho^{\text{eq}}(z)$ called the **equilibrium ensemble**, this is

$$\lim_{t \to \infty} \rho(z, t) = \rho^{\text{eq}}(z) \tag{1.8.1}$$

Due to the highly filamentous structure of $\rho(z, t)$ at all times, strictly speaking, we should understand this limit in the *weak sense*, this is

$$\lim_{t \to \infty} \int dz \rho(z, t) \hat{f}(z) = \int dz \rho^{\text{eq}}(z) \hat{f}(z). \tag{1.8.2}$$

for any phase function $\hat{f}(z)$.

1.8.1   Mixing

One condition of that ensures that the limit (1.8.2) exists is the **mixing** condition of the Hamiltonian flow. As we will see, the mixing condition further predicts that the equilibrium ensemble is just a function of the dynamic invariants of the system, this is $\rho^{\text{eq}}(z) = \Phi(\hat{I}(z))$, ensuring trivially that it is a stationary solution of Liouville's equation. For simplicity, we assume that the only dynamic invariant of the Hamiltonian flow is the energy and consider that the system has a definite energy E and number of particles N. The submanifold $\hat{H}(z) = E$ is usually termed the **energy shell**. A Hamiltonian flow is called mixing if for all square integrable functions $\hat{f}(z), \hat{g}(z)$ on the energy shell, we have

$$\lim_{t \to \infty} \frac{1}{\Omega(E)} \int_E dz \hat{f}(z) \hat{g}(\mathcal{U}_t z) = \frac{1}{\Omega(E)^2} \int_E dz \hat{f}(z) \int_E dz \hat{g}(z) \tag{1.8.3}$$

where we use the shorthand notation

$$\int_E dz \cdots = \int dz \delta(\hat{H}(z) - E) \cdots \tag{1.8.4}$$

and

$$\Omega(E) \equiv \int dz \delta(\hat{H}(z) - E) \tag{1.8.5}$$

This function $\Omega(E)$ gives the overall measure of the energy shell.

By introducing the microcanonical average

$$\langle \cdots \rangle^E = \frac{1}{\Omega(E)} \int dz \delta(\hat{H}(z) - E) \cdots \tag{1.8.6}$$

the mixing property (1.8.3) is interpreted as a decorrelation property between arbitrary phase functions, this is

$$\lim_{t \to \infty} \langle \hat{f} \hat{g}_t \rangle^E = \langle \hat{f} \rangle^E \langle \hat{g} \rangle^E \tag{1.8.7}$$

Note that if we choose for the functions $\hat{f}(z), \hat{g}(z)$ the characteristic functions $\chi_A(z), \chi_B(z)$

of two measurable regions of phase space, then the mixing condition implies that the measure of the intersection of region A of the energy shell with the evolved region $\hat{\Phi}_t(B)$ becomes, as time goes to infinity the product of the original measures of the sets A, B.

In order to proof (1.8.2) whenever (1.8.7) holds, we first discuss what are the effects of the mixing property on the behaviour of the joint probability (1.7.9) of an arbitrary set of functions $\hat{A}(z)$. By including the dynamic invariants of the Hamiltonian flow as the variables, the joint probability is

$$P(a_0, E_0, 0; a_1, E_1, t) = \int dz_0 \rho(z_0, 0) \delta(\hat{A}(z_0) - a_0) \delta(\hat{H}(z_0) - E_0)$$
$$\times \delta(\hat{A}(\mathcal{U}_t z_0) - a_1) \delta(\hat{H}(\mathcal{U}_t z_0) - E_1) \qquad (1.8.8)$$

Being dynamic invariants, we have that $\hat{H}(\mathcal{U}_t z_0) = \hat{H}(z_0)$. Then

$$P(a_0, E_0, 0; a_1, E_1, t) = \delta(E_0 - E_1) \int_{E_0} dz \rho(z, 0) \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\mathcal{U}_t z) - a_1) \qquad (1.8.9)$$

where we have used the notation (1.8.4). The Dirac delta function $\delta(E_0 - E_1)$ reflects the fact that the probability to find the value of E_0 at the initial time and E_1 at time t is zero unless both values are exactly the same.

If we now use the mixing property (1.8.3) in (1.8.9) by choosing as the functions $\hat{f}(z) = \rho_0(z) \delta(\hat{A}(z) - a_0)$ and $\hat{g}(\mathcal{U}_t z) = \delta(\hat{A}(\mathcal{U}_t z) - a_1)$ we have that

$$\lim_{t \to \infty} P(a_0, E_0, 0; a_1, E_1, t) = \delta(E_0 - E_1) \frac{1}{\Omega(E_1)} \int_{E_0} dz \rho(z, 0) \delta(\hat{A}(z) - a_0) \Omega(a_1, E_1)$$
$$= \delta(E_0 - E_1) P(a_0, E_0, 0) \frac{\Omega(a_1, E_1)}{\Omega(E_1)} \qquad (1.8.10)$$

where $P(a_0, E_0, 0)$ is the probability of finding a_0, E_0 at the initial time,

$$P(a_0, E_0, 0) = \int dz \rho(z, 0) \delta(\hat{H} - E_0) \delta(\hat{A}(z) - a_0) \qquad (1.8.11)$$

and we have introduced

$$\Omega(a_1, E_1) = \int dz \delta(\hat{E}(z) - E_1) \delta(\hat{A}(z) - a_1) \qquad (1.8.12)$$

The interest of the result (1.8.10) is manifested when we consider the long time limit of the one time probability distribution that is inferred from (1.8.10). In fact, we have

$$P^{\text{eq}}(a_1, E_1) = \lim_{t \to \infty} P(a_1, E_1, t) = \lim_{t \to \infty} \int da_0 dE_0 P(a_0, E_0, 0; a_1, E_1, t) \qquad (1.8.13)$$

which on inserting (1.8.10) gives

$$P^{\text{eq}}(a_1, E_1) = \int da_0 dE_0 \delta(E_0 - E_1) P(a_0, E_0, 0) \frac{\Omega(a_1, E_1)}{\Omega(E_1)}$$

$$= P_0(E_1) \frac{\Omega(a_1, E_1)}{\Omega(E_1)} \tag{1.8.14}$$

where $P_0(E_1)$ is the probability of dynamic invariants at the initial time

$$P_0(E) = \int dz \rho_0(z) \delta(\hat{H}(z) - E) \tag{1.8.15}$$

By using (1.8.12) we have finally

$$P^{\text{eq}}(a_1, E_1) = \int dz \frac{P_0(\hat{E}(z))}{\Omega(\hat{E}(z))} \delta(\hat{E}(z) - E_1) \delta(\hat{A}(z) - a_1) \tag{1.8.16}$$

But, according to (1.7.2) this is just the way that a microscopic probability relates with the probability of functions in phase space. As this is for arbitrary phase functions $\hat{A}(z)$, we expect that in a weak sense, any initial ensemble $\rho_0(z)$ will evolve under a mixing Hamiltonian flow towards an equilibrium ensemble given by

$$\rho^{\text{eq}}(z) = \rho_N^0 \frac{P_0(\hat{H}(z))}{\Omega_0(\hat{H}(z))}. \tag{1.8.17}$$

where we have introduced ρ_N^0 with dimensions of $(\text{action})^{-N}$ in order to have

$$\Omega_0(E) = \rho_N^0 \int dz \delta(\hat{H}(z) - E) \tag{1.8.18}$$

with the same physical dimensions as $P_0(E)$.

This completes the proof that the mixing property (1.8.7) on microcanonical correlations implies that, in a weak sense, the system reaches an stationary distribution function $\rho^{\text{eq}}(z)$. The mixing property ensures that this stationary distribution is a function of the dynamic invariants of the system. In that sense, it is a stationary solution of the Liouville's equation.

    Eq. (1.8.17) is a fundamental result of equilibrium Statistical Mechanics that reflects the foundational statement (1.8.1). The intuitive meaning of (1.8.17) is very suggestive. $\Omega_0(E)$ is the measure of the submanifold of microstates corresponding to the energy E and we may think that it is proportional to "the number" of microstates that have a value E of the energy. Therefore, the probability density $\rho^{\text{eq}}(z)$ of a given microstate z is the "probability $P_0(\hat{H}(z))$ of being in the submanifold $\hat{H}(z) = E$ divided by the number of microstates of that submanifold". This is often referred to as the **the principle of equiprobability** that states that at equilibrium "all microstates with the same value of $\hat{H}(z)$ are equiprobable". Of course, this is more of a mnemotechnical rule for (1.8.17) than a rigorous statement because $\rho_{\text{eq}}(z)$ is not a probability but rather a probability density and, more important, being a continuum submanifold, the "number of microstates"

satisfying $\hat{H}(z) = E$ is infinite.

Usually, the proof that a given Hamiltonian system ergodic (see below) or mixing is very difficult but we will assume that the systems considered in this book are of the mixing type and have, therefore, a tendency to reach a well defined equilibrium state given by "equiprobability of microstates".

The function $\Omega_0(E)$ is usually termed the structure function [?] and contains all the macroscopic thermodynamic information about the system [?]. Equations of state, in particular, are derived from this function. In order to have results that agree with corresponding results obtained directly from Quantum Mechanics, the appropriate value for ρ_N^0 should be taken as

$$\rho_N^0 = \frac{1}{h^{3N} N!} \tag{1.8.19}$$

Planck's constant h gives the appropriate dimensions and allows for the correct classic limit of quantum results. The factor $N!$ is known as the **correct Boltzmann counting** and there are conflicting views on the physical origin of this factor, attributing it either to the quantum indistinguishability of the particles [?] or to the distinguishability of the classical particles [?]. It turns that the $N!$ factor can be obtained with entirely classic arguments [?] (see the macrocanonical ensemble later on in section **??**). In any case, we take the particular form (1.8.19) as given. As it will be seen for the ideal gas latter, it ensures the extensivity of the macroscopic entropy. Of course, if we have a mixture of particles that can be distinguished by some property (for example, they have different mass), then the factorial coefficient changes accordingly to account for the different equivalent ways of ordering the particles.

The *functional form* of the equilibrium ensemble $\rho^{\text{eq}}(z)$ depends on the initial preparation of the system through the probability density $P_0(E)$. For example, if at the initial time the system is prepared to have certain fixed energy E_0 and certain values a_0 of the phase functions $\hat{A}(z)$ then the initial probability density $P_1(a, E, 0)$ will be

$$P_1(a, E, 0) = \delta(a - a_0)\delta(E - E_0) \tag{1.8.20}$$

which corresponds to an initial ensemble density which is microcanonical in all the variables a, E

$$\rho(z, 0) = \rho_N^0 \frac{\delta(\hat{A}(z) - a_0)\delta(\hat{H}(z) - E)}{\Omega(a_0, E)} \tag{1.8.21}$$

where we introduce the prefactor ρ_N^0 also in

$$\Omega(a_0, E) = \int dz \rho_N^0 \delta(\hat{A}(z) - a_0)\delta(\hat{H}(z) - E) \tag{1.8.22}$$

The initial distribution of the energy, irrespective of the distribution of values of a will be

$$P_0(E) = \int da P(a, E, 0) = \delta(E - E_0) \tag{1.8.23}$$

Therefore, according to (1.8.21) the final equilibrium ensemble density will be

$$\rho_{eq}(z) = \rho_N^0 \frac{\delta(\hat{H}(z) - E_0)}{\Omega(E_0)} \tag{1.8.24}$$

which is microcanonical only in the energy. All the a's dependence in the initial ensemble (1.8.21) has been washed out by the mixing property. Other initial preparations different from (1.8.20) can be conceived. For example, an initial distribution canonical in the energy but microcanonical in the variables a which will lead to a canonical equilibrium distribution. The different *initial* preparations of the system at the initial time will be reflected in the *final* equilibrium ensemble density.

1.8.2 Ergodicity and the lack of it

Another important property of Hamiltonian dynamics is ergodicity. Ergodicity is related but not equivalent to mixing. A mixing flow is ergodic if the correlations not only decay as in (1.8.7) but they do it rapidly in such a way that their time integral is finite [?, ?]. The Hamiltonian dynamics is said to be ergodic, if the probability that the system starting from a microstate z_0 goes through an arbitrarily small neighborhood of *any* other microstate z is one, provided that we wait sufficiently long time. More precisely, given a microstate z and another microstate z', then the solution of Hamilton's equations with initial condition z will pass through an arbitrarily small vicinity of z' in finite time. This means, in particular, that any microstate z in phase space may be reached from any other and that there are not "disconnected" parts in the flow in phase space, a property of phase space known as metric transitivity. Birkoff[1] proved that for an ergodic system, the time average of an arbitrary phase function $\hat{A}(z)$, which is defined as

$$\overline{A} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \hat{A}(\mathcal{U}_t z_0) dt \tag{1.8.25}$$

is independent of z_0 and equals the microcanonical ensemble average

$$\overline{A} = \int dz \frac{1}{\Omega(E)} \delta(\hat{H}(z) - E) \hat{A}(z) \tag{1.8.26}$$

where $E = \hat{H}(z_0)$. By selecting as the function $\hat{A}(z)$ the characteristic function $\hat{\chi}_M(z)$ of a region M in phase space ($\hat{\chi}_M(z) = 1$ if $z \in M$ and zero otherwise), it easy to conclude that for an ergodic system, the fraction of time spent in that region, is proportional to the measure of the intersection of that region with the energy shell.

One of the most fascinating and intriging aspects of a Hamiltonian flow is that they may loose ergodicity as some of the parameters (like energy or volume[2]) in the system change. An example of loosing ergodicity as the volume is reduced is given in Fig. 1.2. For a large box, all configurations may be reached in principle from any given

---

[1]For a discussion see, Dorfman J R, 1999 An Introduction to Chaos in Nonequilibrium Statistical Mechanics (Cambridge: Cambridge University Press)

[2]Note that the volume of the container is a parameter of the Hamiltonian, because the container may be modelled with an external stiff confining potential.
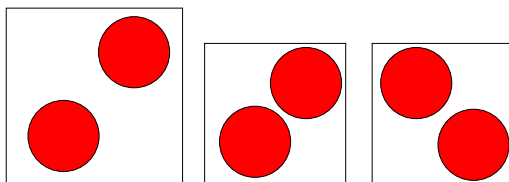
Figure 1.2: Two hard spheres in a 2D box [**?**, **?**].

configuration. If the size of the box is smaller than twice the diameter of the sphere, the number of available configurations is drastically reduced (in a discontinuous manner!). In addition, there are configurations that are not accessible from others by following the dynamic evolution of the system. In particular the configuration in the third box cannot be reached from the configuration on the second box. This is an example of an "entropic" non-ergodicity transition.

Another simple example of loosing ergodicity is shown in Fig. (1.3) where a particle in one dimension moves under the forces generated by a double well potential. At high energy E_1, all points with the same energy are accessible by the dynamics from any of them. At low energy E_2 when the particle is confined in one of the wells, it cannot reach configurations corresponding to the other well. This is an example of an "energetic" non-ergodicity transition. The Hamiltonian $\hat{H}(z) = \frac{p^2}{2m} + \frac{k}{2}(q^2 - 1)^2$ of the system is symmetric under the mirror symmetry $q \to -q$. This implies, for example, that the microcanonical average $\langle q \rangle = 0$. However, in a dynamic evolution at low energy, the system will always be trapped in one of the minima (say the right one) and the time average will be $\overline{q} \neq 0$. In a non-ergodic situation the time average does not coincide with the microcanonical ensemble average that involves all the energy shell.
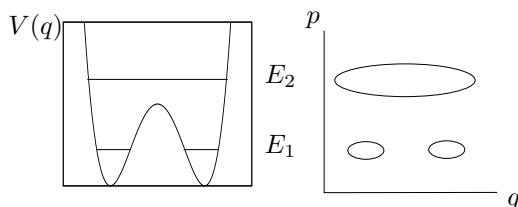
Figure 1.3: A double well potential for a particle in one dimension and a cartoon of the trajectory of a particle in the corresponding phase space q, p at two different energies E_1, E_2. The energy shell that for E_2 is one closed curve disconnects into two closed curves at the energy E_1. At energy E_1, it is not possible to go from a point in the first curve into another point of the second curve.

One example of loosing ergodicity is when a fluid system cristalizes into a solid as the available energy is lowered. There are many possibilities for the resulting cristalline solid, all of them differing by translations and rotations. However, it is extremely difficult that once in a given crystal configuration, the system may visit any other crystal configuration.

Conse-
quences of
loosing
ergodicity

The system would need to spontaneusly melt due to a fluctuation and recrystallize into the second cristalline configuration. As this is an extremelly rare event, the actual phase space is restricted to be those configurations that are around the crystalline space (and related to the vibrations of the atoms around this crystalline configurations). This loosing of ergodicity reflects also into a loose of symmetry, from the isotropic fluid into the less symmetric crystal configuration.

When a system looses ergodicity, the time average no longer coincides with the ensemble average because the system becomes trapped in a region M of phase space. One possibility to keep using the ensemble average is to restrict to M the phase space over which the ensemble averages are performed[3]. This is equivalent to include a confining potential in the Hamiltonian.

[Break of reading flow here] As Stillinger and Debenedetti [?] clearly put it: "Thus, the notion of a single system being at "equilibrium" only implies that its classical dynamics has visited a tiny but accurately representative sample of the available basin set that is relevant at the prevailing observational conditions. By contrast, individual liquid systems supercooled below their respective glass transition temperatures will also sample only minuscule fractions of the available basins; however, in that case the sampling is far from representative for the prevailing temperature."

1.8.3 Irreversibility

We have seen that the way to account for the fact that we cannot prepare a system with many degrees of freedom in a prescribed microstate leads to a probabilistic description in terms of ensembles. Also, we have seen that when the Hamiltonian has the property of mixing, an effective ensemble develops which at long time is stationary. This means that if we prepare our system with a prescribed (necesarily macroscopic) procedure many times, the average value $\langle \hat{A} \rangle_t$ *over the different realizations of the experiment* of any phase function $\hat{A}(z)$ will evolve in time towards a time independent value given by $\langle \hat{A} \rangle^{\text{eq}}$. However, experimentalist may perform the experiment just once and we have a single realization of the trajectory of $\hat{A}(\hat{\Phi}_t(z))$. Nonetheless, it is common everyday experience that in the evolution of this single experiment the values of $\hat{A}(\hat{\Phi}_t(z))$ tend to go towards a definite value. The classic example of a diffusing drop of ink in water shows that the concentration field evolves from a highly peaked value at the initial time towards an almost constant profile at long times. In a single realization of the experiment, the signal $\hat{A}(\hat{\Phi}_t(z))$ allows as to detect "past" from "present". This poses an interesting question because, as we have seen, the microscopic evolution $\hat{\Phi}_t(z)$ is time reversible and we cannot tell in Fig 1.1 what of the to panels is running forward or backward in time.

---

[3]See [?]"Statistical mechanics of time independent non-dissipative nonequilibrium states"

The green dog-flea model

To grasp what is going on, it may be helpful to consider a version of the urn model of Erhenfest []. Imagine that we have a large number of dogs whose colors are either white or green. Green dogs are very rare[4] and we only have a few of them in the collection. Green dogs like to be close to green dogs. Now consider a flea that jumps at random from dog to dog, with no respect for its color (as the red blood tastes the same for all dogs). It is clear that if the flea starts in a green dog it has a large chance to land on another close green dog, but after several jumps, it is highly probable that the dog it lands on is white, just because there are so many white dogs. Once in a white dog the probability of landing on a green dog is then much smaller. Therefore, if we pay attention just to the color of the dog the flea is in, going from green to white is much probable than going from white to green, even though going from dog A to dog B is as probable as going from dog B to dog A. Note that typically, *any single* realization of the process where the flea starts in a green dog will give very similar color recordings, particularly if the number of dogs increases.

We may translate now the basic idea above in phase space terms. In the above example, each particular dog is "the microstate", the color of the dogs are "the macrostate", rare green dogs are "non-equilibrium microstates" while copious white dogs are "equilibrium microstates". In a phase space description, to every value a of the observable $\hat{A}(z)$ we may associate the "number of microstates" that is compatible with that value a. Therefore, as the microstate travels through phase space it will encounter regions of different value of the observable $\hat{A}(z)$. However "blindly" it travels with no distinction of back and forth motion in time, it will typically spend more and more time in regions where the macrostate a has more microstates.

This notion of irreversibility was unveiled and fully understood already by Boltzmann more than 140 years ago. Nevertheless, the issue of irreversibility has been a source of debate for a long time, even recently, as can be appreciated in Refs. [?] and [?]. Boltzmann introduced the **Boltzmann entropy**, that we will refer also as **bare entropy**

$$S_B(a) = k_B \ln \Omega(a) \tag{1.8.27}$$

as a measure of the "number of microstates" associated to a macrostate described by a set of phase functions $\hat{A}(z)$. The function $\Omega(a)$ is defined as[5]

$$\Omega(a) = \int dz \delta(\hat{A}(z) - a) \tag{1.8.28}$$

and it is a measure of the region of phase space satisfying $\hat{A}(z) = a$. As the trajectories of a system in phase space will travel regions of larger measures, the entropy $S_B(\hat{A}(\mathcal{U}_t z))$ will be typically larger and larger as the system evolves. The genius of Boltzmann was to relate the microscopic object (1.8.27) with Clausius thermodynamic entropy, thus providing a microscopic view of the Second Law of Thermodynamics.

---

[4] According to a Spanish saying.
[5] In the original work, instead of using the Dirac delta function, shells of finite width Δa were introduced.

The Detailed Balance theorem

The intuitive picture described in the green dog-flea model can be cast into the form of a nice theorem known as Detailed Balance. This theorem is one among many that fall under the name of Fluctuation Theorems, that will be considered explicitly in Chapter 4. Detailed Balance reflects microscopic reversibility at the level of CG variables or phase functions. This theorem is described in, for example, de Groot and Mazur [**?**].

Consider the joint probability distribution $P(a_1, t_1, a_2, t_2)$ that the observable $\hat{A}(z)$ take the value a_1 at time t_1 *and* the value a_2 at time t_2, given in microscopic terms in (1.7.9). We will assume that the initial probability density $\rho_0(z)$ has a very special form, depending on the microstate only through the observables, this is

$$\rho_0(z) = \frac{P_0(\hat{A}(z))}{\Omega(\hat{A}(z))} \tag{1.8.29}$$

where $P_0(a)$ is the (constant) probability distribution of the observables in the preparation of the system. This assumption can be justified when we prepare the system initially to have prescribed values of the observable. Because the observable has fixed values, the dynamics in this preparation phase must be such that the observable is a dynamic invariant of the dynamics. If we wait long enough, the system will equilibrate until having an equilibrium ensemble which is a function of the dynamic invariants and, hence, a function of $\hat{A}(z)$ as described in (1.8.29). In Sec **??** we will justify such assumption from an inference point of view.

By using (1.8.29) into (1.7.9) we have

$$P(a_0, t_0, a_1, t_1) = \frac{P_0(a_0)}{\Omega(a_0)} \int dz \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\hat{\Phi}_{t_0}^{t_1}(z)) - a_1) \tag{1.8.30}$$

where we have used the flow map representation for the evolution operator. The conditional probability introduced in (1.7.10) of having a_1 at time t_1 given that the macrostate was a_0 at t_0 now becomes

$$P(a_0, t_0 | a_1 t_1) = \frac{P(a_0 t_0, a_1 t_1)}{P_0(a_0)} = \frac{1}{\Omega(a_0)} \int dz \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\hat{\Phi}_{t_0}^{t_1}(z)) - a_1) \tag{1.8.31}$$

Now consider the reversed conditional probability of finding a_0 at time t_1 given that at time t_0 the macrostate was a_1

$$P(a_1, t_0 | a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz \delta(\hat{A}(z) - a_1) \delta(A(\hat{\Phi}_{t_0}^{t_1}(z)) - a_0) \tag{1.8.32}$$

If we use now (1.5.20) in the form

$$\hat{\Phi}_{t_0}^{t_1}(z) = \epsilon \cdot \hat{\Phi}_{t_1}^{t_0}(\epsilon \cdot z) \tag{1.8.33}$$

we have

$$P(a_1, t_0 | a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz \delta(\hat{A}(z) - a_1) \delta(\hat{A}(\epsilon \cdot \hat{\Phi}_{t_1}^{t_0}(\epsilon \cdot z)) - a_0) \qquad (1.8.34)$$

We will assume that the observable function has a well defined parity under time reversal, meaning that

$$\hat{A}_\mu(\epsilon \cdot z) = \varepsilon_\mu \hat{A}_\mu(z) \qquad (1.8.35)$$

where the values of ε_μ are ± 1 and the sign determines the time reversal character of the relevant variable. Then

$$P(a_1, t_0 | a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz \delta(\hat{A}(z) - a_1) \delta(\varepsilon \cdot \hat{A}(\cdot \hat{\Phi}_{t_1}^{t_0}(\epsilon \cdot z)) - a_0)$$

$$= \frac{1}{\Omega(a_1)} \int dz \delta(\hat{A}(z) - a_1) \delta(\hat{A}(\hat{\Phi}_{t_1}^{t_0}(\epsilon \cdot z)) - \varepsilon \cdot a_0) \qquad (1.8.36)$$

Now we consider the change of variables $z' = \hat{\Phi}_{t_1}^{t_0}(\epsilon \cdot z)$ with unit Jacobian and arrive at

$$\frac{P(a_1, t_0 | a_0, t_1)}{P(\varepsilon \cdot a_0, t_0 | \varepsilon \cdot a_1, t_1)} = \frac{\Omega(a_0)}{\Omega(a_1)} \qquad (1.8.37)$$

that can be written also in terms of the bare entropies

$$\frac{P(a_0, t_0 | a_1, t_1)}{P(\varepsilon \cdot a_1, t_0 | \varepsilon \cdot a_0, t_1)} = e^{S_B(a_1) - S_B(a_0)} \qquad (1.8.38)$$

This property of the transition probabilities reflecting microscopic reversibility is also known as **detailed balance** [**?**]. For the case that the observables are invariant under time reversal $A(\epsilon \cdot z) = \hat{A}(z)$, i.e. $\varepsilon = 1$, (1.8.38) shows that the ratio of the transition probability of jumping from a_0 to a_1 over that of jumping from a_1 to a_0 is given by the ratio of the measures of $\Omega(a_0)$ and $\Omega(a_1)$. If it happens that $\Omega(a_1) > \Omega(a_0)$ or, equivalently $S_B(a_1) > S_B(a_0)$, then it is much probable that the system jumps from a_0 to a_1 rather than the other way around. Many observables are such that the bare entropy scales typically with the number N of degrees of freedom, $S_B(a) \sim N s(a)$, where $s(a)$ does not depend on N. When this happens, for a macroscopic system, the probability of going to states of smaller bare entropy is extremely small (exponentially small in the number of degrees of freedom).

The Detailed Balance Theorem (1.8.38) is a remarkable exact result for any Hamiltonian system that puts the heuristic ideas behind the green-dog-flea model into firm mathematical ground, and demonstrates that the observables behave irreversibly in spite of the fact that microstates do behave reversibly. Note that quite interestingly, the condition of microscopic reversibility is what ensures the validity of this theorem on irreversibility.

We will obtain in Chapter 4 some other exact results that, based on microscopic reversibility, display in one way or another the features of irreversibility. We will defer these results after the consideration of external forcing in the Hamiltonian, because many

of the results are valid even when the Hamiltonian is time-dependent.

1.9 The single, all encompassing problem of Non-Equilibrium Statistical Mechanics

The theory of coarse-graining that we present in this book assumes and is limited by the presupposition that all macroscopic processes can be ultimately understood in terms of the classical laws of motion of the constituent particles. So far, we have seen how an isolated classical system is governed at the microscopic level by Hamilton's equations with a time-independent Hamiltonian and how any initial distribution over the phase space will evolve, if the Hamiltonian is of the mixing type, towards the equilibrium ensemble. This means that we restrict ourselves, until Chapter 4 to the study of the evolution towards the equilibrium state of an initial distribution which is not the stationary solution of the Liouville equation corresponding to the Hamiltonian of the system. Our limitation on isolated systems that decay to equilibrium precludes, apparently, the possibility to study non-equilibrium stationary states that are maintained with "external couplings". These experimental situations, however, do fit into the framework of an "isolated decaying system" whenever we consider the system under study in contact with reservoirs, in such a way that the composed system of system+reservoirs is isolated. In this view, a stationary state is just an extremely long-lived decay towards the global equilibrium of the system+reservoir, where the time scale towards equilibrium is dictated by the size of the reservoir. We will consider the action of time-dependent external fields on the dynamics of the system in Chapter 4. While the unperturbed system still has an equilibrium state, the perturbed system does not need to reach and equilibrium state.

Therefore, the basic process that we study is how an arbitrary initial ensemble decays towards the equilibrium ensemble

$$\rho_0(z) \ \longrightarrow \ \rho^{\mathrm{eq}}(z) \tag{1.9.1}$$

A system left to evolve will reach the equilibrium state corresponding to the Hamiltonian of the system. Because no matter how an isolated system is prepared[6] it will go towards the same equilibrium state, the equilibration of a system is a particularly simple way to *prepare and control* the initial state of a system in the preparation phase of an experiment. Therefore, we assume that the initial ensemble $\rho_0(z)$ is the equilibrium state of certain Hamiltonian H_0. At $t = 0$ some parameter of the Hamiltonian changes and transforms the original Hamiltonian H_0 into another Hamiltonian H_1. The ensemble ρ_0 is no longer the equilibrium ensemble of H_1 and it will evolve according to the dynamics generated by H_1 until it reaches the equilibrium state of H_1.

Callen in his magnificent book "Thermodynamics" states in a crystalline sentence what is, perhaps, the essential tenet of the book: *"The single, all encompassing problem of thermodynamics is the determination of the equilibrium state that eventually results after the removal of internal constraints in a closed, composite system"*. If we think about this problem in microscopic terms, the removal of an internal constraint can always be described at a molecular level as a change of the Hamiltonian of the system and, therefore,

---

[6]Provided the distribution of dynamic invariants is the same in all preparations!

is a process of the form described in the previous paragraph. For this reason, we may state that the fundamental problem of Thermodynamics is, indeed the very same problem of Non-Equilibrium Statistical Mechanics, except that in the latter case not only the final equilibrium state is seeked for, but the richer question of *how this state is reached in time* is answered.

Example: Two systems equilibrating to each other

Think, for example, of two isolated systems with different energy content that are brought to contact. This problem will be considered at length in Chapter **??**. Energy will be transferred from one system to the other until the equilibrium state is reached. This experimental protocol can be understood as the following preparation. The Hamiltonian of the system is

$$\hat{H}(z) = \hat{H}_1(z_1) + \hat{H}_2(z_2) + \lambda \hat{H}_{\text{int}}(z_1, z_2) \tag{1.9.2}$$

Here $z = (z_1, z_2)$ where $z_1 = (\mathbf{q}_{1_1}, \mathbf{p}_{1_1}, \cdots, \mathbf{q}_{N_1}, \mathbf{p}_{N_1})$ and $z_2 = (\mathbf{q}_{1_2}, \mathbf{p}_{1_2}, \cdots, \mathbf{q}_{N_2}, \mathbf{p}_{N_2})$ are the coordinates and momenta of each subsystem. $\hat{H}_1(z_1), \hat{H}_2(z_2)$ are the Hamiltonian of systems 1 and 2, respectively, and $\hat{H}_{\text{int}}(z_1, z_2)$ stands for the potential energy of interaction between both systems. Initially $\lambda = 0$ and there is no interaction between the two systems. The equilibrium state of the composite system will be the product of the ensembles of each system because they are not interacting. If we know the exact amount of energy in each system, the distribution of each subsystem will be of microcanonical type, as in (27.0.15) and the equilibrium ensemble of the composite non-interacting system will be

$$\rho^{\text{eq}}_{\lambda=0}(z_1, z_2) = \frac{\delta(\hat{H}_1(z_1) - E_1)}{\Omega_1(E_1)} \frac{\delta(\hat{H}_2(z_2) - E_2)}{\Omega_2(E_2)} \tag{1.9.3}$$

This is the equilibrium ensemble corresponding to $\lambda = 0$. At a given moment we set $\lambda = 1$ by, for example, reducing the distance between the centers of mass of each system until the particles of different systems are so close to exert forces on the other system and then, interact with each other. When the interactions between the two systems are switched on, the new Hamiltonian is $\hat{H}_{\lambda=1}(z)$. This new Hamiltonian has an equilibrium ensemble which is different from (1.9.3) and which is given by

$$\rho^{\text{eq}}_{\lambda=1}(z) = \phi(\hat{H}_{\lambda=1}(z)) \tag{1.9.4}$$

where the function $\phi(E)$ can be obtained by following an identical argument as that leading to (27.0.12), this is,

$$\rho^{\text{eq}}_{\lambda=1}(z) = \frac{P(\hat{H}_{\lambda=1}(z))}{\Omega(\hat{H}_{\lambda=1}(z))} \tag{1.9.5}$$

Here, $P(E)$ is the initial distribution of the energy which is given by

$$P(E) = \int dz_1 dz_2 \rho^{\text{eq}}_{\lambda=0}(z_1, z_2)\delta(\hat{H}_1 + \hat{H}_2 + \hat{H}_{\text{int}} - E)$$

$$= \int dz_1 dz_2 \frac{\delta(\hat{H}_1(z_1) - E_1)}{\Omega_1(E_1)} \frac{\delta(\hat{H}_2(z_2) - E_2)}{\Omega_2(E_2)}\delta(\hat{H}_{\text{int}}(z_1, z_2) - E + E_1 + E_2) \quad (1.9.6)$$

and the measure is

$$\Omega(E) = \int dz_1 dz_2 \delta(\hat{H}_1 + \hat{H}_2 + \hat{H}_{\text{int}} - E) \tag{1.9.7}$$

If we assume that, in order to compute these integrals, the term $\hat{H}_{\text{int}}$ can be neglected because it is just a surface term, then we have

$$P(E) \approx \delta(E_1 + E_2 - E)$$

$$\Omega(E) \approx \int dz_1 dz_2 \delta(\hat{H}_1 + \hat{H}_2 - E) \int dE' \delta(H_1 - E')$$

$$= \int dE' \int dz_2 \delta(\hat{H}_2 + E' - E) \int dz_1 \delta(\hat{H}_1 - E') = \int dE' \Omega_2(E - E')\Omega_1(E')$$

$$\tag{1.9.8}$$

Therefore, we have for (1.9.9) the following expression

$$\rho^{\text{eq}}_{\lambda=1}(z) = \frac{\delta(E_1 + E_2 - \hat{H}_{\lambda=1}(z))}{\int dE' \Omega_1(E')\Omega_2(\hat{H}_{\lambda=1}(z) - E')} \tag{1.9.9}$$

The system, therefore, will evolve under the mixing hypothesis from (1.9.3) to (1.9.9). This transition from one ensemble to another is accompanied by an energy transfer from one system to another. We will discuss in Chapter **??** how is the dynamics of the energy transfer in this problem.

1.10   Summary

2

Levels of description: the CG variables, the relevant ensemble, and the entropy

2.1   What is CG

The theory of coarse-graining is a formalization of the process of representing a given system with less information than the one encoded in the actual microstate of the system. By coarse-graining, one eliminates the "uninteresting" fast variables and keeps the coarse-grained variables, with time scales much larger than typical molecular scales. In the simplification process involved in CG we may acquire a lot of understanding about the system by focusing on the essential details and not being distracted by an overwhelming number of irrelevant details.

If we want to simulate a system with a computer, constructing a CG model for it is a good strategy because we not only gain in terms of a reduction of computational units (particles, nodes, or sites), but also on the possibility of exploring much larger time scales than possible with molecular dynamics (MD) [?]. In MD, the time step of the method should be small enough to resolve the molecular collisions and vibrations of the particles. Typically, this time step is on the order of femtoseconds $= 10^{-15}$ s. The largest MD simulations at the time of writing allow some research teams to simulate billions of particles during hundreds of nanoseconds [?]. However impressive this may be, the resulting systems are still very small. A raindrop of water contains on the order of $\sim 10^{23}$ molecules! This already shows that simulating with MD a macroscopic system during seconds is absolutely unfeasible by any actual or foreseeable computer. On the other hand, in order to describe how a droplet oscillates or evaporates we do not want to follow the fate of every single water molecule inside the droplet. To describe these macroscopic phenomena we are entirely happy with a much smaller amount of information about the droplet, like its radius or its shape, conveniently parametrized with a small number of coefficients.

One and the same system may be described at different **levels of description**

depending on the amount of information which one retains macroscopically. The state of a system at a given level of description is described by a set of **CG variables**, which are functions of the microscopic state z of the system and, therefore, are phase functions $\hat{A}(z)$. The CG variables have received in the past a large number of different names: Macrostates, macroscopic variables, gross variables, CG variables, collective variables, reaction coordinates, order parameters, internal variables, structural variables, etc. With the symbol $\hat{A}(z)$ we denote a collection of phase functions each one labeled with a discrete index like in, for example, $\hat{A}(z) = \{\hat{A}_\mu(z), \mu = 1, \cdots, M\}$. We will often consider phase functions that are fields $\hat{A}(z) = \{\hat{A}_{\mathbf{r}}(z), \mathbf{r} \in \mathbb{R}^3\}$ and we understand that the CG variables are labeled with a continuum index $\mathbf{r}$. When we do not know exactly the total number of particles in the system and a macrocanonical description is required, then $\hat{A}_\mu(z)$ or $\hat{A}_{\mathbf{r}}(z)$ are understood as *a set* of phase functions, with each element of the set defined in each phase space of N particles.

The word "level" suggests a hierarchical structure and, in fact, the levels of description for a given system order themselves in terms of the amount of information captured by the CG variables and their time scale of evolution. Less detailed levels (coarser levels) have a smaller number of variables and capture less information. Typically, coarser levels evolve slower than finer levels. A coarse grained level is valid for describing phenomena that occur at time scales equal or larger than the typical time scale of the level, but it cannot reproduce the behaviour at shorter time scales.

In any system, among all the possible levels of description two of them are particularly important because they lay at the extremes of the hierarchy. These levels are the *micro-scopic* and *macroscopic* levels. The microscopic level has the position and momenta of all the atoms of the system as the set of variables characterizing the state of the system. The dynamic equations are Hamilton's equations, and the time scale is a typical collision or vibration time. This level has been studied in Chapter 1. The macroscopic level is the level of Thermodynamics where the CG variables are the dynamical invariants of the system (mass, momentum, energy). At the macroscopic level, there is no equation of motion, because the CG variables are constant in time and the time scale is infinite. The macroscopic level will be studied in the Chapter **??**. Any other level of description is in between these two levels and could be named as a *mesoscopic* level, although we will reserve this name and the related notion of mesoscale, for those situation in which thermal fluctuations are important. It is apparent that a given system may have many different levels of description, each one of different detail.

The evolution of the CG variables is governed by a **dynamic equation**. *The principal objective of the Theory of Coarse-Graining is to derive these dynamic equations.* The two intertwined crucial ideas of the Theory of Coarse-Graining that allow for the derivation of *simple and tractable* dynamic equations are **quasi-equilibrium** and **separation of time scales**, to be discussed in Sec. 2.3.

2.2   Example: CG a colloidal suspension

As an illustrative example of the notion of levels of description, let us consider a colloidal suspension, which is a system made of a collection of small solid particles of sub-micron size which are suspended in a liquid. Fig. 2.1 below displays different levels of description

(i.e. different resolutions of the same system) that are appropriate to resolve different time scales.
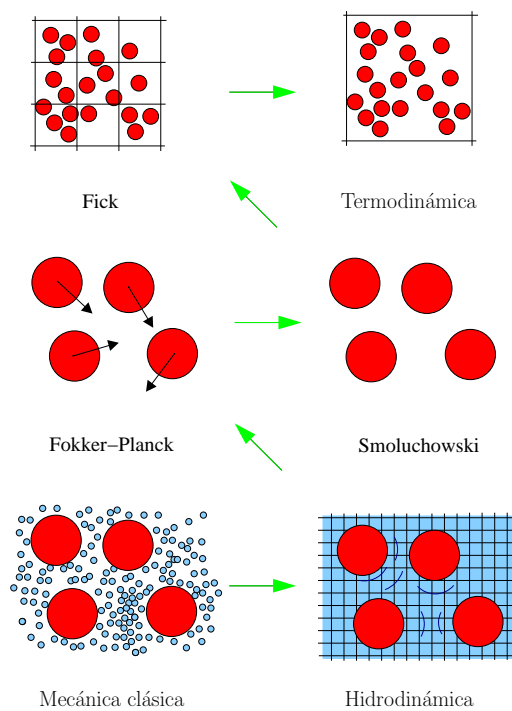
Figure 2.1: Different levels of description in a colloidal suspension. Arrows denote the direction of coarse graining, from the Classical Mechanics level down left to the top right Thermodynamics level.

Level 0: Chemistry detailed classical mechanics

At the microscopic level everything is made of atoms that, in this book, are modeled with point particles interacting classically with prescribed force fields. Therefore, the colloidal suspension is described at this microscopic level with the positions and momenta of all the solvent atoms, plus the corresponding positions and momenta of the atoms that constitute each solid suspended particle. To characterize microscopically the system, we should know what kind of atoms constitute the system. Assume that we have a colloidal suspension made of spherical silica particles in water. For describing water (H_2O), we need one oxygen atom and two hydrogen atoms for each water molecule. However, due to its small mass, hydrogen needs to be described with Quantum Mechanics, which requires to include an infinitely large number of degrees of freedom (given by the wave function of the problem). As this is too costly, the strategy is to model the water molecules with *sites* (two to five depending on the water model) that interact classically, either with springs or constraints for the bonds, and potential energies of the Coulomb and

Lennard-Jones type for the unbounded interactions. In a similar way, a silica particle is made of silicon dioxide SiO_2, and one needs the potential of interaction between the silicon and oxygen atoms. In addition, one requires to describe the interaction of the water molecules with the silicon dioxide in the silica particle. The construction of the classic potentials governing the interactions of the atoms in the system is a crucial step and requires a very laborious work involving fitting against *ab initio* quantum mechanic simulations and macroscopic thermodynamic parameters. Once we have constructed these approximate classical potentials, we may solve the resulting Hamilton's equations with molecular dynamics.

Level 1: Microscopic CGMD

With accurate classic force fields for all the atoms in the system we can look at very detailed information about how water and silica atoms interact. This information may be, perhaps too detailed for our interest, and we could be satisfied with much simpler models for both the water molecules and the whole silica particles. Note that the silica particles are solid and move more or less rigidly. For spherical silica particles we could focus on the center of mass of the particle and consider the water molecules as point particles. Of course, this would require to parametrize the potentials of interaction of these new "atoms", let us call them particles from now on as their similitude with real atoms is now very vague. Nevertheless, the particles still move according to Hamilton's equations of motion. For this reason, the simulation of these oversimplified models for the force fields between particles receives the name of **Coarse-Grained Molecular Dynamics** (CGMD). The state of the system in this level of description is given by $z = \{\mathbf{q}_i, \mathbf{p}_i, \mathbf{Q}_i, \mathbf{P}_i\}$, where $\mathbf{Q}_i$ is the position and $\mathbf{P}_i$ is the momentum of the center of mass of the silica sphere's, and $\mathbf{q}_i$ is the positions and $\mathbf{p}_i$ the momentum of the center of mass of the molecules that constitute the fluid. If instead of spherical particles we have irregular particles we would need also to consider orientation, angular velocities, etc. The evolution of the microstate is governed by Hamilton's equations.

The particular functional form of the potential of interaction between particles needs to be specified. The simplest models are sums of pair-wise interactions, with typical potential functions like Morse or Lennard-Jones potentials. Then the parameters of these potentials need to be adjusted, either by fitting against more detailed simulations of the kind described above, or by trying to reproduce some thermodynamic behaviour of water. The smallest typical time scale in this model is a collision time in the range of picoseconds. In order to resolve numerically these time scales, we need to use a time step for the numerical solution of Hamilton's equations that is much smaller than the collision time scale. However, if the mass of the colloidal particles is much larger than the mass of the solvent particles, as it occurs in reality, the evolution of the colloidal particles will be very slow in comparison with the evolution of the solvent molecules. If we are interested in the motion of the colloidal particles, then we would need an enormous number of time steps (and, therefore, of computer time) to observe an appreciable motion of the colloidal particles. To study these large time scales in a colloidal suspension, molecular dynamics is absolutely impracticable.

Level 2: Hydrodynamics

If we look at the motion of the molecules of the solvent, we will see that they collide with each other resulting in a rapid motion of the molecules. However, if we look "from a distant point" to the multitude of molecules, a collective motion will be appreciated in which molecules in a region of space move coherently (overwhelming the small erratic motions due to collision). It will be possible to appreciate slowly evolving waves, vortices and other sort of collective motion. The variables that capture this collective motions are the hydrodynamic variables. These variables are the mass density field $\hat{\rho}_\mathbf{r}(z)$, the momentum density field $\hat{\mathbf{g}}_\mathbf{r}(z)$, and the energy density field $e_\mathbf{r}(z)$, defined by

$$\hat{\rho}_\mathbf{r}(z) = \sum_i m\overline{\delta}(\mathbf{r} - \mathbf{q}_i),$$

$$\hat{\mathbf{g}}_\mathbf{r}(z) = \sum_i \mathbf{p}_i\overline{\delta}(\mathbf{r} - \mathbf{q}_i),$$

$$\hat{e}_\mathbf{r}(z) = \sum_i e_i\overline{\delta}(\mathbf{r} - \mathbf{q}_i), \tag{2.2.1}$$

where $\overline{\delta}(\mathbf{r} - \mathbf{q}_i)$ is a coarse-grained Dirac delta function (a function with a support over a finite small region and normalized to unity). The definition (2.2.1) is not unique as several different functional forms for $\overline{\delta}(\mathbf{r} - \mathbf{q}_i)$ can be used. In the above equations, e_i is the energy of particle i (the sum of its kinetic energy plus half the potential energy due to the interaction with its neighbours). It may appear as a contradiction to coarse-grain through a set of *field* variables (which have, in principle an infinite number of degrees of freedom). However, we should note that the above fields involving the coarse-grained delta functions are "smooth" fields (which have a *small* number of Fourier components with large wavelengths). Implicit in the definition of the hydrodynamic variables in (2.2.1) there is a partition of physical space in little cells that contain many solvent molecules. The hydrodynamic variables tell us about how many particles have a given average velocity and energy in a certain region of space, rather than providing the exact location and velocity of each solvent molecule. The reduction of information in passing from solvent atoms to smooth hydrodynamic fields is then apparent. We will discuss in great detail the definition of discrete hydrodynamics in Chapter **??** and, in particular a specific functional form of the coarse-grained delta functions.

The state of the system in this level of description of the colloidal suspension is described by the set of hydrodynamic fields *and* the position and momenta of the colloidal particles. The equations that govern the hydrodynamic variables are the well-known hydrodynamic equations [**?**]. In the hydrodynamic equations, the microscopic details of the solvent (what *particular* molecules constitute the solvent) are included through the pressure and temperature equations of state and through the particular values of the transport coefficients (shear and bulk viscosities, and thermal conductivity). The very fact of using equations of state for the local values of the pressure and temperature, already points out to the fact that quasi-equilibrium is in action in this level of description. In fact, we may regard each little cell as a small thermodynamic subsystem, where the atoms of the fluid are in a state of equilibrium corresponding to the values of the

hydrodynamic variables. This assumption is also referred as local equilibrium.

The time scales of the hydrodynamic variables depend on the length scale being considered. Typically, perturbations with large wavelengths evolve slowly. In a colloidal suspension these length scales are either the radius of the particles or the typical distance between colloidal particles. Roughly speaking there is a hydrodynamic time scale for mass evolution which is a sound traversal time, a time scale for momentum evolution given by a a viscous time, and a time scale for energy transport, given by a thermal diffusion time. The hydrodynamic equations for the solvent fields are coupled with the Newton's equation for the colloidal particles through boundary conditions, whereas the forces on the colloidal particles are given through the integral of the stress tensor of the fluid over the surface of the colloidal particle. We observe that the motion of a given colloidal particle depends on the motion of the rest through the coupling with the hydrodynamic fields of the solvent. This solvent mediated interactions between colloidal particles are known as *hydrodynamic interactions*.

As a result of the coarse-graining of the solvent, the hydrodynamic variables are subject to fluctuations. Physically, these fluctuations came from the fact that we have a finite (not infinite) number of solvent molecules on average in each of the cells that are used to defined the hydrodynamic variables. Molecules can enter and go out from the cells, causing its number and overall velocity to fluctuate. Mathematically, the fluctuations are described as stochastic contributions to the stress tensor and heat flux in the Navier-Stokes equations, as proposed by Landau and Lifshitz in their theory of fluctuating hydrodynamics [?]. A very remarkable property of these fluctuations is that their variance is proportional to the transport coefficients of the solvent, this is, to the shear and bulk viscosities and to the thermal conductivity. This is the fluctuation-dissipation theorem for this level of description. The hydrodynamic fluctuations are transmitted to the colloidal particles through the boundary conditions on the surface of the particles and are the ultimate responsible for the Brownian motion of the particles [?]. In the highly dilute limit (that is, for an isolated colloidal particle) and for neutrally buoyant particles (whit density equal to the solvent density) it is possible to compute explicitly the velocity autocorrelation of the colloidal particle [?], which exhibits the famous long-time tail, an algebraic decay as $t^{-3/2}$ [?]. The physical origin of this slow algebraic decay is that, due to a spontaneous fluctuation, the colloidal particle pushes the fluid in front of it. The push generates a vortex centered on the particle that, in turn, pushes the particle from its rear, thus maintaining the initial state of motion of the particle. These long-time tails have been observed experimentally for colloidal systems [?].

Level 3: Fokker-Planck

If the colloidal particles are massive their motion will be slow, and we could expect that the propagation of the hydrodynamic interactions is very fast, instantaneous, compared with the time scale of variation of the colloidal position and momenta. If this is the case, we can eliminate the solvent hydrodynamic variables from the description. The state of the system at this mesoscopic level of description is given by $x = \{\mathbf{Q}_i, \mathbf{P}_i\}$. Quasi-Equilibrium here means that the solvent is equilibrated conditional to the particular values of the positions and momenta of the colloidal particles.

If the colloidal particles are very far from each other, as it happens when the sus-

pension is dilute, we may expect that the mutual influence among colloidal particles is negligible and that the friction tensor is diagonal, this is $\boldsymbol{\zeta}_{ij} = \delta_{ij}\mathbf{1}\zeta$, where ζ is called the friction coefficient. In this case, the SDE equivalent to the FPE (**??**) decouple into a set of independent equations, called Langevin equations. The Langevin equations for a dilute suspension predict that the velocity autocorrelation function of a colloidal particle decays exponentially. As we have seen, this is at variance with experiments that show a clear long-time tail for neutrally buoyant particles. The reason for this discrepancy between theory and experiments for neutrally buoyant particles can only be attributed to the fact that the time scales of evolution of the colloidal variables $\mathbf{Q}_i, \mathbf{P}_i$ are not neatly separated from the time scales of hydrodynamic variables. We have here one example of the hazards of selecting a set of CG variables with no separation of time scales. Even though $\mathbf{Q}_i, \mathbf{P}_i$ look as a reasonable set of CG variables, they are not. If we are interested in the time scale in which the momentum of the particles evolves, we have to include, necessarily, the hydrodynamic modes as a part of the description. This time scale is typically $10^{-6}s$. It turns out, however, that the density ratio between the colloidal particle density and the solvent density governs the separation of time scales between colloidal momenta and hydrodynamic variables [**?**]. For colloidal suspensions of particles which are *much denser* than the solvent, these time scales are well-separated and the Fokker-Planck level of description is a good level of description. The method to solve numerically Eqs. (3.5.39) is known as **Langevin dynamics** [**?**],[**?**].

Level 4: Smoluchowski

Even though the CG variables $\mathbf{Q}_i, \mathbf{P}_i$ are not a good set of coarse variables for neutrally buoyant suspensions, it turns out that the $\mathbf{Q}_i$ alone do actually define a proper level of description. The reason is that the time scale of evolution of the positions of the colloidal particles is much longer than the time scale of evolution of the momentum *and* hydrodynamic modes. The typical time scale of evolution of the position variables for a colloidal suspension is $10^{-3}s$. Therefore, by using $\mathbf{Q}_i$ alone, we will be able to describe correctly the phenomena that occur in time scales *above* $10^{-3}s$.

Level 5: Fick

A further coarse graining can be performed if we are interested not in the actual positions $\mathbf{Q}_i$ of the colloidal particles but on the number of colloidal particles that are in a region of space located around $\mathbf{r}$. This amounts to the introduction of the number of colloidal particles per unit volume and to the use of a concentration field $c(\mathbf{r})$ as the variables x describing the state of the system. The concentration variable contains a much reduced amount of information because it tells us how many particles are in a region, but not which one is exactly where. We expect that the concentration field evolves in a time scale much larger than the time scale of evolution of the positions of the colloidal particles because in order to have an appreciable change in the number of particles in a given sufficiently large region, these particles must move (diffuse) for a long time in order to traverse this region.

Level 6: Macroscopic Thermodynamics

Finally, we might be interested in the very long time scale in which the system has arrived at equilibrium. In this case, the only CG variables are the dynamical invariants, that is, those coarse-grained variables that are independent of time due to particular symmetries of the microscopic Hamiltonian, like total energy, or those CG variables like mass and volume, that are constant parameters in the Hamiltonian. Note that the volume of the container of the colloidal suspension can be understood as a parameter of a confining potential in the Hamiltonian. Obviously, there is no equation of motion for this level of description because we are interested in the long time, equilibrium state of the system.

Other levels of description

The levels of description presented above for a colloidal suspension do not exhaust the possibilities and they have been selected as providing a pedagogical presentation of the main ideas involved in coarse-graining. Of course, other levels of intermediate complexity can, and need to, be formulated depending on the physical situation that one wants to describe. For example, one can define a level of description where the CG variables are the hydrodynamic solvent variables *plus* the concentration field for the colloidal particles. This purely hydrodynamic level of description is particularly useful for the description of transport of dilute pollutants by fluid flows. In this case, the equations are the hydrodynamic equations (that contain the effect of the dilute suspension through an osmotic pressure term) plus an advection-diffusion equation for the concentration field. If the colloidal suspension is more concentrated and anisotropic effects due to flow start to play a role, then it is necessary to enrich the level of description by including a vectorial or tensorial quantity representing the degree of local anisotropy in the system [?]. Another level of description is given by minimal blob models for colloids [?] where each particle is described with the position of the colloidal particle and the hydrodynamic fields. These models can be justified microscopically by considering discrete hydrodynamic variables defined in cells that are *much larger* than the colloidal particle [?], as suitable for nanocolloids.

In summary, in this section we have illustrated several levels of description that can be defined for the study of a colloidal suspension. The general theme is that every level captures less information than its predecessor and allows one to describe time scales *at and above* the typical time scale of evolution of the CG variables. In principle, one can construct a Theory of Coarse-Graining that allows to consider the derivation of any level of description from any other level down in the hierarchy [?, ?]. However, in the present book we will only consider the transition from the microscopic level governed by Hamilton's equation to the particular level of our interest. Note that one always require further modelling assumptions and approximations to make explicit the functional forms of the different terms of the dynamic equation for the CG variables. The hierarchy of approximatioins of going from one level, say level 1, to level 2, and then from level 2 to level 3 are usually not consistent with those that lead from level 1 to level 3.

2.3 Quasi-Equilibrium

In a consistent theory, the universe itself would be treated according to the same laws that govern any part of it. This poses an interesting situation because the universe itself is clearly in a non-equilibrium state – some of its parts (like galaxies) are moving in macroscopic terms! How is it that we may talk about equilibrium situations of subsystems of the whole non-equilibrium universe? Of course, separating the world in two, as discussed in Sec 1.2, helps. By isolating a system by reducing the interactions with the rest of the universe, and provided that the Hamiltonian dynamics is of the mixing type, the system has a chance of reaching an equilibrium state independently of the rest of the universe. More interestingly, a system may display different equilibrium states, depending on the time scale of observation. Indeed if, for some reason, the dynamics of our system is such that there are functions that evolve very slowly in the time scale of our observations, they may be considered as dynamic invariants that determine the equilibrium properties of the system –at the time scale of our observations–. One would say that the system reaches a state of **quasi-equilibrium**. All global equilibrium states of a system are, in fact, quasi-equilibrium states, because there is no system completely isolated. The non-equilibrium flow of the whole universe will eventually drag the system out of this equilibrium state. After pouring sugar and stirring our cup of hot coffee it finally calms down to a state of rest with all its parts at the same temperature and with the same concentration of sugar. We would say that the coffee cup has equilibrated. But if we wait long enough the coffee will disappear from our cup through evaporation, meaning that the equilibration of the system was not a real equilibrium state. Of course, we may close the lid of our cup, but this does not avoid the fact that all materials degrade in time and, eventually, the very cup and lid will deteriorate. Before that, anyway, any tiny bacteria feeding on sugar will chemically process the interior of the cup to produce, in a time scale much larger than the time the coffee settles down at rest, a rather different substance from our delicious original coffee.

The theory of non-equilibrium that we present in this book is based on this notion of having several "equilibrium time scales" and takes advantage of the partial equilibration of the system in the time scale of evolution of the selected CG variables. This crucial insight was introduced by Bogoliubov [**?**] in the context of kinetic theory, but it is extensible the Theory of Coarse-Graining: the evolution towards global equilibrium proceeds in a hierarchical way, each level of the hierarchy characterized by a set of CG variables with characteristic time scales.

If we look at the system during the time scales typical of a given level of description, we will observe that it reaches a state very similar to an equilibrium one in which the relevant variables were themselves dynamic invariants. At that particular time scale "all degrees of freedom except the CG variables" reach equilibrium, *conditional* to the slowly evolving CG variables. We refer to this state of affairs as quasi-equilibrium. The fast degrees of freedom equilibrate rapidly and in the time scales of these fast degrees of freedom, the system approximately behaves like an equilibrium system where the CG variables were "frozen", so to speak, at particular (non-equilibrium) values. Quasi-Equilibrium means that a system behaves approximately as if the CG variables were truly dynamic invariants of the system. When quasi-equilibrium is valid, the CG variables of a given level of description evolve much slowly than the CG variables of any other more detailed

The relevant ensemble

level of description down in the hierarchy of levels of descriptions. To these faster levels, the CG variables look as if they were quasi-dynamic invariants of the system. Therefore, the system is approximately described, in the appropriate time scale, with a generalized equilibrium ensemble, called **the relevant ensemble** or **quasi-equilibrium ensemble**, that takes into account the CG variables in its definition, as if they were dynamical invariants of the system. We need to emphasize that quasi-equilibrium does not mean that the theory presented is restricted to situations near global equilibrium. Arbitrarily far from equilibrium situations can be described with the Theory of Coarse-Graining. We have said that the single encompassing problem of Non-Equilibrium Statistical Mechanics is the study of *how* the system approaches the equilibrium state. This process is one in which a cloud of points initially concentrated in a region of the energy shell (or a cloud inhomogeneous in any other way) spread uniformly in that energy shell. Now, imagine that the system has an additional invariant (you may think, for example, about total linear momentum). This means that the flow in phase space will be "stratified" in layers, in which the points on every layer never leaves the layer (in order to conserve the dynamical invariant). Now, imagine that we do not have such additional dynamic invariant, but the flow in phase space is quasi-stratified, in a way that it becomes rapidly homogeneous in layers and the flow from one layer to another occurs slowly. If this is the dynamical scenario at the microscopic level, we expect that the phase function that implicitly defines the layers will be slow variables and good candidates to be CG variables. It is clear what is the role of the relevant ensemble now, it just describes the equilibrium ensemble (the one that gives equiprobability to all microscopic states of a given layer) of a system in which the CG variables where really invariants and not just slow variables. In this heuristic picture, it becomes apparent the central role in the Theory of Coarse-Graining of the relevant ensemble as an approximation of the real time dependent ensemble, and the close connection of the relevant ensemble with the notion of slow variables.

[The theory is not limited to near equilibrium situations, shocks and turbulence are quite far from equilibrium phenomena that are described however with the compressible Navier-Stokes equation which is based on the quasi-equilibrium assumption]

Associated to this picture of quasi-equilibrium we have the notion of separation of time scales. In fact, in order to have quasi-equilibrium we need that the CG variables evolve slowly compared with the rest of degrees of freedom in the system. If (and only if) the time scales of the different levels of description are well separated from each other, it is possible to obtain *simple* differential equations for the evolution of the CG variables. If the time scales are not well-separated, it is still possible to obtain formal dynamic equations, but they are not simple at all. In particular, they contain complicated memory terms and the equations become integro-differential. This situation is referred to as a **non-Markovian** dynamics, to distinguish it from **Markovian** descriptions governed by differential equations where the future state of the system is determined by the present, but not the past of the system. As we will discuss later, only if the selected level of description has scales well separated from the rest we expect to have a reasonable and tractable theory for that level of description.

2.4   How to identify the CG variables?

The identification of the CG variables $\hat{A}(z)$ is an art of the theory of coarse-graining and a crucial step, perhaps the most important one, in order to describe macroscopically a system with many degrees of freedom. As we have stressed in the previous chapter, we are concerned with the transition (1.9.1) from an initial ensemble towards the equilibrium ensemble. Usually, this happens in a way that it is possible to identify collective motions and patterns that emerge in the course of the relaxation towards equilibrium. When we stop stirring our cup of coffee, which at its most microscopic level is made of colliding atoms, vortices are clearly visible that suggest that the relaxation happens following "paths" in the phase space. These paths are characterized by phase functions whose values evolve in time much slowly than other phase functions. We will see that from a practical point of view only when there is a clear separation of time scales between the selected CG variables and the correlations of their time rate of change, it is possible to have simple dynamic equations for the relevant variables. When this happens we have that the CG variables "forget" their past rapidly and their future is essentially determined by their present values. In these cases, we say that the description is Markovian.

There are few guiding principles for the selection of CG variables. Whenever we have conserved or quasiconserved variables, we expect that they will need to be included in the description. Therefore, *we will always include in the set of relevant variables the dynamic invariants of the system. In particular, the number of particles $\hat{N}(z)$ and the Hamiltonian $\hat{H}(z)$ will be included in the description.*

Many but minuscule

A rule of thumb for a phase function to be a good candidate for a CG variable is that its evolution results from *many but minuscule* contributions. This happens in general when the phase function changes because of the cumulative effect of "collisions" or "particles entering regions". This micro events have a very short time scale while the overall appreciable change on the CG variable occurs in a much larger time scale. A sketch of the situation is given in Fig. 2.2. If this is the case, there is a chance that in the time scale in which the CG variables have hardly changed, a lot of independent things have already happened, in a way that can be described as a sort of equilibration conditioned to the actual values of the relevant variables. A somewhat more technical argument for this criteria about selecting CG variables is as follows. Time scales are usually identified with the help of correlation functions. By looking how fast a variable decorrelates we obtain a typical time scale for the evolution of this variable. As we will see, the Theory of Coarse-Graining gives exact integro-differential equations where the memory term is given by the correlation of the *time derivatives* of the CG variables. Therefore, to the extent that the evolution of the CG variables occur in a time scale much larger than the evolution of its time derivatives, we will be able to take the Markovian approximation and obtain a differential equation. This may seem a contradiction, because if a function evolves in a time scale, then its derivative evolves also at that time scale (just think of a sine function)! However, this is true for continuous and differentiable functions. The key point is that if the CG variables are functions of time that display two very different time scales, with a fast component (prone to be modeled with a random function) superimposed to a slow component, we may have a correlation function of the time derivatives that decay in a much shorter time scale than the correlation function of the relevant variables themselves. Therefore, a CG variable that changes slowly due to
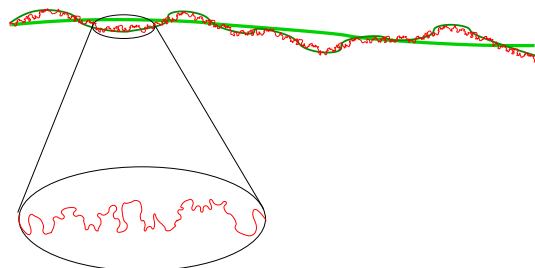
Figure 2.2: Cartoon of the time evolution of a CG variable, displaying three different time scales. For this type of evolution, the correlation of the time derivative will display two decay regions, one very fast, on the time scale of the zoomed part, and a longer trend, in the time scale of the green curves.

a large number of minute fast contributions is a good candidate for the construction of a Markovian level of description. Fig. 2.2 gives a cartoon of the situation in which a CG variable changes according to different time scales.

Another rule of thumb for selecting CG variables is the following. If you can "see", or "appreciate", or "identify" some feature in your system, most probably you should find a relevant variable that captures that feature, because in the time scales that you are living in (where the seeing-appreciating-identifying occurs), these CG variables are slow. If they were fast variables, you would not be able to "see" them, as they would appear to you averaged and blurred out. Therefore, if for some reason, we expect a characteristic feature (orientation, stretching, elongation, etc.) to play a role in the dynamics of the system, then we need to include in the collection of CG variables the phase functions that best capture such feature. Usually we choose the CG variables because there is "something salient" in the system that we notice, for example a systematic orientation of molecules, for which we need a vector to describe this salient feature. Salient means here usually a symmetry. Note, however, that the concept of symmetry is rather subtle. For example, a crystal is a rather symmetric object but, in fact, is *less* symmetric than its melted counterpart, which is symmetric under *arbitrary* translations, not only lattice translations. We say that in passing from a fluid to a crystalline solid, there has been a **symmetry breaking**. An important theorem due to Goldstone and Nambu says that whenever there is a symmetry breaking in the system then a slow mode appears. These modes are called Goldstone modes and they are variables that need to be incorporated into the description. The physical intuition is that when a symmetry is broken, long range correlations appear for certain variables. These variables, therefore, will change slowly because in order to change at one point, they need to change over large distances, involving large portions of material. Therefore, these long ranged variables will be slow variables.

Finally, we note that the selection of CG variables may depend also on the experimental set up. We may "excite" variables or not, depending on how we prepare the system. Unexcited variables in our experiment need not to be considered in the description. For example, if we consider flow of a fluid in a nanochanel made of parallel walls, it seems

Salient

Broken
symmetry

Excited

natural to take as relevant variable the parallel component of the flow field, but not any perpendicular component, that will not be (presumably) excited along such a flow field.

Quality
control

But the truth is that we never know in advance whether a set of CG variables will describe a appropriately the situation. Some times, we forget to include in the set of CG variables (or ignorantly ignore) some variable that turn out to evolve in a time scale comparable to the selected CG variables. When this occurs, the prediction of the theory will fail. The theory of coarse-graining allows for a rule of thumb "quality control" of the selected level of description. If the memory function (related to the correlation of the time derivatives of the CG variables) decay in a time scale comparable to the correlation of the relevant variables themselves, forget about this level of description because it will not be Markovian.

Non-
Markovian

In this case you should direct your attention to the possibility that there are "missing processes" that may be occurring in the time scales of interest and try to identify the corresponding CG variables. The general strategy when there is no such a separation of time scales is to look for additional variables that also evolve in comparable time scales as the ones that we believe are the slow variables. By enlarging the set of CG variables, we hope that the resulting description may be Markovian.

2.5   The Maximum Entropy Principle

The **macroscopic information** that we will retain in a statistical description of a system are either the averages of the CG variables

$$a = \langle \hat{A} \rangle = \int dz \rho(z) \hat{A}(z) \tag{2.5.1}$$

or their full probability distribution function

$$P(a) = \int dz \rho(z) \delta(\hat{A}(z) - a) \tag{2.5.2}$$

Imagine that at a given time we have access to this macroscopic information and this is all the information that we have about the state of the system. It is obvious that there are many possible ensembles $\rho(z)$ that would reproduce such a macroscopic information. We would like to find "the best" ensemble that reproduces the macroscopic information. By the best ensemble we mean an ensemble that captures in the least biased way the macroscopic information. In this way, we are faced with the problem of determining a probability density out of partial information about it. The problem of finding the least biased probability distribution compatible with some macroscopic information has been solved in the field of Information Theory and makes use of the Principle of Maximum Entropy, first introduced by Gibbs and elaborated by Jaynes in the context of Statistical Mechanics [?].

In order to find the least biased distribution function $\overline{\rho}(z)$ that is compatible with a

given macroscopic information one introduces the **Gibbs-Jaynes entropy functional**

$$\mathcal{S}[\rho] = - \int dz \rho(z) \ln \frac{\rho(z)}{\rho_N^0} \tag{2.5.3}$$

The dimensional factor ρ_N^0 is given by (1.8.19) and renders the argument of the logarithm dimensionless. Usually the Gibbs entropy is made dimensional by multiplying it with the Boltzmann constant k_B. With the definition (2.5.3) we assume that we measure the Gibbs entropy in units of k_B.

The basic idea of the maximum entropy principle [?] is that in order to find the least biased probability density $\rho(z)$ in phase space (i.e. the least biased ensemble) when only partial information is known about $\rho(z)$, one should maximize the entropy functional (2.5.3) subject to the constraints of this partial information. The entropy functional represents, from an information theory point of view, the "amount of uncertainty" of a distribution function [?] and by maximizing it subject to whatever is known, we are taking the distribution which is "maximally noncommittal with regard to missing information, and that it agrees with what is known, but expresses maximum uncertainty with respect to all other matters" in Jaynes words [?]. In order to find the least biased distribution $\overline{\rho}$ we should maximize the entropy functional (2.5.3) subject to the constraint that this ensemble should provide the information that we have available. The constrained maximization is done with the use of Lagrange multipliers and we will have many occasions to see how this works along this book. The least biased ensemble will be also referred as **the relevant ensemble** $\overline{\rho}(z)$.

A measure of the information content of each level of description is given by the **entropy of the level of description**. The entropy of a given level of description is defined by evaluating the Gibbs entropy functional (2.5.3) at its maximum, given by the relevant ensemble. Because the relevant ensemble is fixed by the macroscopic information corresponding to the given level of description, the entropy becomes a function, or functional, of this macroscopic information. We stress that there is no such a thing as *the* entropy of a system. Instead, to every level of description of a given system it corresponds a different entropy function. The entropy of the macroscopic Thermodynamics level of description corresponds to the "equilibrium" entropy that depends, for that particular level of description on the dynamic invariants of the system only. However, entropies of more detailed levels of description depend in general on the CG variables characterizing such a level. Many examples will be given along the book about the different explicit forms of the entropy function corresponding to the level of description selected.

2.6   Compatibility between mixing and the Maximum Entropy Principle

In this section, we show that there is a deep connection between the mixing property of Hamiltonian systems and the Principle of Maximum Entropy. This connection is revealed by deriving the mesocanonical ensemble (27.0.12) with the Principle of Maximum Entropy. Recall that the mesocanonical ensemble (27.0.12) is the effective equilibrium probability that emerges at long times, no matter what initial distribution of probability of microstates we may have, when the Hamiltonian system is mixing, This ensemble "remembers" from the initial ensemble the distribution of dynamic invariants but nothing else, and gives the same probability density to all microstates with the same values of the dynamic invariants. In this way, the mixing property is the randomization process at the basis of the "principle of equiprobability".

Note that there is a particular selection of a measure in phase space (1.8.4) implicit in the mixing condition stated in (1.8.3). This measure is proportional to $\delta(\hat{H}(z) - E_0)$. Other measures could be used, in principle. For example we could measure the extension of the energy shell with the Euclidian surface area, whose measure is

$$\int dz |\nabla \hat{H}(z)| \delta(\hat{H}(z) - E) \cdots \tag{2.6.1}$$

However, the use of a metric implicit in the definition of such an area is strange in the context of phase space. Phase space has no metric, but a symplectic structure that provides for a natural symplectic volume form [?]. The measure (1.8.4) used in the mixing property is "natural" in the sense that is the symplectic volume of an "infinitely thin" shell. Another argument in favour of the choosing the measure (1.8.4) for the mixing property stems from the ergodic theorem (1.8.26), where the measure (1.8.4) appears in a natural way. But it should be clear that the selection of the measure is an aditional physical ingredient in the theory.

On the other hand, the Maximum Entropy Principle is a method for obtaining the least biased distribution function given partial information. If the only known information is that the microstate is in a certain energy shell, and nothing else, the most judicious choice for assigning a probability is to say that all these microstates are equally probable. While in a discrete setting this assignations is somewhat trivial– we usually assign $1/6$ to the probability of outcome of a dice– in a continuous setting equiprobability requires a *measure* (i.e volume or area) associated to regions in order to say that regions of equal measure are equally probable. The measure induced by the mixing property is proportional to $\delta(\hat{H}(z) - E)$. We show now that the same measure is obtained with the Maximum Entropy Principle.

By definition, the probability distribution $P(E)$ that the system has an energy E is given by the microscopic probability $\rho(z)$ through

$$P(E) = \int dz \rho(z) \delta(\hat{H}(z) - E) \tag{2.6.2}$$

There are many different microscopic probability densities $\rho(z)$ that give the same $P(E)$, i.e. that satisfy (2.6.2). According to the Maximum Entropy Principle, the least

biased distribution function $\rho(z)$ which is compatible with the macroscopic information (2.6.2) is the one that maximizes (2.5.3) with the constraint (2.6.2). Eqs. (2.6.2) are a set of conditions on $\rho(z)$, one for each value of E. By using the method of Lagrange multipliers, we will need a Lagrange multiplier $\nu(E)$ for each value of E. Then we need to maximize without constraints the Lagrangian functional

$$
\begin{aligned}
\mathcal{L}[\nu, \rho] =& \mathcal{S}[\rho] - \int dE \, \nu(E) \left[P(E) - \int dz \rho(z) \delta(\hat{H}(z) - E) \right] \\
=& - \int dz \rho(z) \ln \frac{\rho(z)}{\rho_N^0} \\
& - \int dE \, \nu(E) \left[P(E) - \int dz \rho(z) \delta(\hat{H}(z) - E) \right]
\end{aligned}
\tag{2.6.3}
$$

By setting to zero the functional derivatives of $\mathcal{L}[\nu, \rho]$ with respect to ν we recover the restriction (2.6.2). The functional derivative with respect to $\rho(z)$ set to zero gives the maximum $\rho^{\mathrm{eq}}(z)$ at

$$
\rho^{\mathrm{eq}}(z) = \rho_N^0 \exp\{\nu(\hat{H}(z))\}
\tag{2.6.4}
$$

Substitution into the constraint (2.6.2) allows to eliminate the Lagrange multipliers $\nu(E)$ in terms of $P(E)$ and leads to the final result

$$
\rho^{\mathrm{eq}}(z) = \rho_N^0 \frac{P(\hat{H}(z))}{\Omega(\hat{H}(z))}
\tag{2.6.5}
$$

This is identical to the equilibrium ensemble (27.0.12). In Sec. 27 we have obtained the general form (27.0.12) of the equilibrium ensemble under the mixing condition of the Hamiltonian flow, that ensures that the equilibrium ensemble depends on the microscopic state only through the value of its energy $\hat{H}(z)$. The mixing property of the dynamic flow ensures that all microscopic states with the same energy have the same probability density. We have just seen that the "equiprobability" of all microscopic states produced by the mixing property can also be obtained from the Maximum Entropy Principle as the least biased ensemble compatible with the macroscopic information $P(E)$.

We may say that the Maximum Entropy Principle favours a particular selection of a measure in phase space, precisely the one that it is used in the mixing property. Eventually, the ultimate reason for the particular selection of a measure on which to draw conclusions about statistical events should be based on the effectiveness in reproducing observed regularities. For a further discussion of this topic see Ref [**?**].

The relevance of the Principle of Maximum Entropy in the Theory of Coarse-Graining, beyond finding the limiting equilibrium ensemble, stems from the following observation. Imagine that the CG variables were really, really slow, almost dynamic invariants of the system. In this case, we would expect that the mixing Hamiltonian would produce a "quasi equilibrium" ensemble compatible with the dynamic invariants *and* the quasi-invariants due to the CG variables. In the short time scales in which the relevant variables act as dynamic invariants, the system equilibrates due to the mixing property of the Hamiltonian, and adopts an ensemble in which all microstates compatible with

the same values of the Hamiltonian *and* the CG variables are equally probable. This quasi-equilibrium ensemble is just the relevant ensemble obtained by maximization of the Gibbs-Jaynes entropy functional. The relevant ensemble should be very similar to the actual ensemble satisfying the Liouville's equation, but not identical, because the CG variables do actually move at long time scales. The keystone of the Theory of Coarse-Graining is then the approximation of the real ensemble by the relevant ensemble, plus "small" corrections. As we will see, these corrections, even if thought to be small, are extremely important as they give rise to dissipative effects. In fact, when the time-dependent statistics of the system is fully approximated with the relevant ensemble, the system evolves with a **reversible dynamics**, that do not produce any increase of the entropy. However, this approximation is not very good in general and one needs to include dissipative terms in the evolution equation. The deviations of the real dynamics with respect to the one described by the relevant ensemble is the responsible for the **irreversible part of the dynamics**, that causes entropy to increase. All the theories obtained within the Theory of Coarse-Graining respect the Second Law.

Inference approach

Non-Equilibrium Statistical Mechanics aims at describing how the macroscopic information (averages or probabilities) about the system evolves in time. This macroscopic information is, by assumption, the only information that is accessible by the experimental equipment at hand. In an experiment, the system needs to be initially prepared, then perturbed in some way, and then let it evolve as we record the macroscopic information. Even if we prepare the system always in the same way, the macroscopic constraints in the initial preparation of the state of the system do not allow us to specify what is the precise configuration of the atoms in the system. Our "clumsy macroscopic experimental hands" do not allow us to fix the initial microstate and, therefore, the system may go to different places in each realization of the experiment. We need statistical methods to describe this uncertainty about the evolution of the system. As, the laws of classical mechanics are deterministic, all the predictive problem lies in the initial conditions. After insufficient reason, in words of Bernouilli, it seems natural to assign "equal probabilities" to all the initial microscopic states compatible with the initial macroscopic information. Then, it seems that the objective of Non-Equilibrium Statistical Mechanics is to make the most honest predictions with the available initial macroscopic information. This purely statistical inference approach has been advocated by Jaynes [**?**] and has been elaborated afterwards in an explosive number of papers and books [**?**]. The Gibbs-Jaynes **information entropy** plays a crucial role in assigning equal probabilities to initial microstates compatible with the macroscopic information.

We believe that this inference approach is not all the story about Non-Equilibrium Statistical Mechanics, because it obviates the fact that the dynamics of ordinary systems is such that leads to quasi-equilibrium ensembles and, eventually, to global equilibrium ensembles. What makes information entropy useful is the fact that *naturally* the dynamics of the system leads to a statistical distribution of microstates approximately given by a quasi-equilibrium ensemble, with almost equiprobability of microstates compatible with the macroscopic state. How this occurs is one of the mysteries of Hamiltonian dynamics, as clearly stated by Zwanzig in the last paragraph of his book [**?**]: "So the final issue is to

understand why the initial states that we can construct experimentally are characterized by such simple ensembles. We can always assert that an ensemble is just a model of reality that can be confirmed by experiment. This brings us to the essential mystery of statistical mechanics, whether equilibrium or nonequilibrium – why do such models work in the first place?" The Theory of Coarse-Graining does not answer this question at all, but makes plausible that under some properties of the Hamiltonian dynamics the evolution of the system is such that leads to quasi-equilibrium. In particular, we believe that the mixing property of a Hamiltonian system, which is the responsible for ensuring an evolution towards effective *global* equilibrium ensembles is also the responsible for fast equilibration under quasi invariant CG variables leading to *quasi*-equilibrium ensembles.

2.7 The flavours of a level of description

A level of description is prescribed once we give the set of CG variables $\hat{A}(z)$. However, in order to obtain the relevant ensemble of this level of description, we need to maximize the Gibbs-Jaynes entropy subject to the known macroscopic information. This macroscopic information can be either the *averages* of the CG variables or their full *probability distribution function*. Depending on whether we keep track of the averages or of the distribution functions we will speak about different **flavours** of a level of description. In general terms, given a level of description defined by a set of CG variables $\hat{A}(z)$ (that include the dynamical invariants), the **averages flavour** of the level of description is characterized by the ensemble averages $a = \langle \hat{A} \rangle$ of the CG variables as macroscopic information. The **probabilistic flavour** is characterized by the distribution functions $P(a)$ of the CG variables. Of course, we may have some variables that are described in average terms and other variables that are treated at the level of distribution functions[1]. The probabilistic flavour involves much more information about the system because all the moments (covariances, etc.) of the distribution and not only the first moment (averages) are known. Therefore, the level of description given by the probability of CG variables $P(a)$ is more detailed than the level of description given by the averages $\langle \hat{A} \rangle$ of the CG variables. It gives information about not only averages but also about fluctuations.

The next three sections are devoted to the derivation of the relevant ensemble and the entropy of the corresponding three flavours: averages flavour, the probabilistic flavour, and the averages flavour suited for MD simulations. These are essential ingredients in the construction of the dynamics of the CG variables in later chapters.

2.7.1 The averages flavour

In this section, we obtain the relevant ensemble and the entropy of a level of description in the averages flavour, this is, when the macroscopic information is given by the averages of the CG variables.

---

[1]We will consider this situation when describing dynamic invariants with probability distributions, while the rest of evolving CG variables are described with their averages

The relevant ensemble

In order to obtain the relevant ensemble in the averages flavour we have to maximize the Gibbs-Jaynes entropy function (2.5.3) subject to the following restrictions

$$\text{Tr}[\rho] = \int dz \rho(z) = 1$$

$$\text{Tr}[\hat{A}\rho] = \int dz \hat{A}(z)\rho(z) = a \qquad (2.7.1)$$

The first condition ensures that the ensemble is normalized properly. By following the method of Lagrange multipliers, one first introduce the Lagrangian $\mathcal{L}[\rho, \lambda, \mu]$, which is a functional of $\rho(z)$ and a function of a set of $M + 1$ variables λ, μ, called the Lagrange multipliers

$$\mathcal{L}[\rho, \lambda, \mu] = \ \mathcal{S}[\rho] - \lambda \left(\text{Tr}[\hat{A}\rho] - a \right) - \mu \left(\text{Tr}[\rho] - 1 \right) \qquad (2.7.2)$$

The Lagrange multipliers λ are also known as **conjugate variables**. Here $\lambda \hat{A}(z)$ implies a sum or an integral, depending on whether the CG variables are discrete or continuum. For example, if $\hat{A}(z) = \{\hat{A}_\mu(z), \mu = 1, \ldots, M\}$ we have

$$\lambda \cdot \hat{A}(z) = \sum_{\mu=1}^{M} \lambda_\mu \hat{A}_\mu(z) \qquad (2.7.3)$$

Usually we will also make use of Einstein's summation convention over repeated indices, i.e. $\lambda \cdot \hat{A}(z) = \lambda_\nu \hat{A}_\nu(z)$. On the other hand, for a continuum set of CG variables as, for example, $\hat{A}(z) = \{\hat{A}_{\mathbf{r}}(z), \mathbf{r} \in \mathbb{R}^3\}$ we understand

$$\lambda \cdot \hat{A}(z) = \int d\mathbf{r} \lambda_{\mathbf{r}} \hat{A}_{\mathbf{r}}(z) \qquad (2.7.4)$$

The explicit form of (2.7.2) is

$$\mathcal{L}[\rho, \lambda, \mu] = \ -\int dz \rho(z) \ln \frac{\rho(z)}{\rho_N^0} - \lambda \left(\int dz \hat{A}(z)\rho(z) - a \right) - \mu \left(\int dz \rho(z) - 1 \right) \quad (2.7.5)$$

The method of Lagrange multipliers ensures that the maximum of the Lagrangian with respect to *all* its arguments provides the maximum of the entropy functional (2.5.3) subject to the constraints (2.7.1). Indeed, by setting to zero the derivatives with respect to λ, μ we recover the constraints (2.7.1).

By setting equal to zero the functional derivative of (2.7.5) with respect to $\rho(z)$ we obtain the relevant ensemble

$$\overline{\rho}(z) = \rho_N^0 \exp\{-1 - \mu - \lambda \cdot \hat{A}(z))\} \qquad (2.7.6)$$

Substitution of (2.7.6) into the first constraint allows one to eliminate the multiplier μ

and we obtain the normalized distribution function

$$\overline{\rho}(z) = \rho_N^0 \frac{\exp\{-\lambda \cdot \hat{A}(z))\}}{Z(\lambda)} \tag{2.7.7}$$

where the normalization is given by the **dimensionless partition function** defined as

$$Z(\lambda) \equiv \int dz \rho_N^0 \exp\{-\lambda \cdot \hat{A}(z))\} \tag{2.7.8}$$

Of course, we are assuming that this partition function (2.7.8) exists. This is in general ensured if, among the CG variables we include the dynamic invariants. Should this not be the case, we are in serious problems that reflect the fact that the system itself does not have a well defined equilibrium state. A well-known case where a system does not have an equilibrium state is for a system of self-gravitating particles.

The rest of constraints (2.7.1) can be expressed as

$$\frac{\partial}{\partial \lambda} \Phi(\lambda) = a \tag{2.7.9}$$

where we have introduced the **dimensionless thermodynamic potential** as

$$\Phi(\lambda) = -\ln Z(\lambda) \tag{2.7.10}$$

Eqs. (2.7.9) are a system of M equations that allow to obtain the M Lagrange multipliers λ as a function of the M averages a. Note that the solution is unique, this is, for every a there is one and only one set of Lagrange multipliers also known as conjugate variables λ. The proof of this reduces to show that the dimensionless thermodynamic potential (2.7.10) is a concave function of its arguments. Indeed, the matrix of second derivatives of the thermodynamic potential is (minus) the matrix of the covariances, as it is easily shown, this is

$$\frac{\partial^2 \Phi}{\partial \lambda \partial \lambda} = -\langle \delta \hat{A} \delta \hat{A} \rangle^\lambda \tag{2.7.11}$$

where $\delta \hat{A} = \hat{A}(z) - a$ and the averages $\langle \cdots \rangle^\lambda$ are taken with the relevant ensemble (2.7.7). A matrix of covariances is positive semi-definite. If we left and right multiply with an arbitrary vector the matrix of covariances we obtain

$$v^T \cdot \frac{\partial^2 \Phi}{\partial \lambda \partial \lambda} \cdot v = -\langle (v \cdot \delta \hat{A})^2 \rangle^\lambda \leq 0 \tag{2.7.12}$$

This quantity is always negative or zero for all vector v. The only possibility for the above quadratic form to be zero is whenever there exists a vector v for which the scalar product with $\delta \hat{A}$ vanishes[2]. Let us assume for a moment that such a vector does not exists and, therefore, the matrix of covariances is strictly positive definite. This is another way to express that the thermodynamic potential is a strictly concave functional of its

---

[2]The average of a function that is always positive or zero vanishes if and only if the function itself vanishes

arguments. According to (2.7.9), the matrix of second derivatives is, in fact, the Jacobian matrix $\partial a/\partial \lambda$ of the change of variables from $a \leftrightarrow \lambda$. Therefore, because the Jacobian matrix is definite negative, it is invertible. According with the theorem of the inverse function, the relationship between the average values and the conjugate variables is one to one [**?**, **?**, **?**].

The entropy

Substitution of the relevant ensemble (2.7.7) into the Gibbs entropy functional (2.5.3) gives the entropy of the level of description of the CG variables $\hat{A}(z)$. The result is

$$S(a) = -\Phi(\lambda) + \lambda a \qquad (2.7.13)$$

where here, λ is understood as being a function of a, i.e. $\lambda(a)$. Therefore, the entropy at the level of description determined by the averages of the CG variables $\hat{A}$ is (minus) the Legendre transform of the dimensionless thermodynamic potential. The derivatives of the entropy are given by

$$\frac{\partial S}{\partial a}(a) = -\frac{\partial \Phi}{\partial \lambda}(\lambda(a))\frac{\partial \lambda}{\partial a}(\lambda) + \frac{\partial \lambda}{\partial a}(a)a + \lambda(a) = \lambda(a) \qquad (2.7.14)$$

where we have used (2.7.9). Because the dimensionless thermodynamic potential is the Legendre transform of the entropy, one should probably call the dimensionless thermodynamic potential as the **dimensionless Massieu functions**, to comply with Callen's nomenclature when the CG variables only include dynamic invariants [**?**]. The thermodynamic potential is termed the Massieu-Planck functions in Zubarev's book [**?**]. We, nevertheless, stick to the equilibrium Statistical Mechanics nomenclature in which the logarithm of a partition function is a thermodynamic potential [**?**]. Also, we will refer to the entropy (2.7.13) in the averages flavour as the **dressed entropy**, to distinguish it from the bare entropy (1.8.27). As we will see, when the CG variables are extensive and fluctuations can be neglected both entropies are very similar, but in general they are different functions.

Second derivatives

We have seen in (2.7.11) that the second derivatives of the thermodynamic potential give the negative of the matrix of covariances of the CG variables. In addition, we have the one to one connection between the conjugate variables and the CG variables, i.e. we have the functions $\lambda(a)$ and also the inverse $a(\lambda)$. The composite function $a(\lambda(a))$ is just the identity, this is

$$a = a(\lambda(a)) \qquad (2.7.15)$$

If we take the partial derivatives of this function with respect to a and use the chain rule, we have

$$1 = \frac{\partial a}{\partial \lambda}(\lambda(a))\frac{\partial \lambda}{\partial a}(a) \qquad (2.7.16)$$

If we now use (2.7.9) and (2.7.14), we arrive at

$$1 = \frac{\partial^2 \Phi}{\partial \lambda \partial \lambda}(\lambda(a)) \frac{\partial^2 S}{\partial a \partial a}(a) \tag{2.7.17}$$

which shows that the matrix of second derivatives of the thermodynamic potential and of the entropy function are inverse of each other. This shows, in particular, that the matrix of second derivatives of the entropy is the negative of the inverse of the covariance matrix, i.e.

$$\frac{\partial^2 S}{\partial a \partial a}(a) = - \left[\langle \delta \hat{A} \delta \hat{A} \rangle^{\lambda(a)} \right]^{-1} \tag{2.7.18}$$

2.7.2   The probabilistic flavour

In this section, we consider the relevant ensemble and the entropy of a level of description described with the probabilistic flavour. The probabilistic flavour of a level of description is one in which the statistical information that we have about the CG variables is the probability $P(a)$ that the system has a particular realization N, a of the value of the CG variables $\hat{N}, \hat{A}(z)$. This probability is given in microscopic terms as

$$\int dz \rho(z) \delta(\hat{A}(z) - a) = P(a) \tag{2.7.19}$$

The Lagrangian functional to be maximized without constraints is

$$\mathcal{L}[\rho] = \mathcal{S}[\rho] - \int da \, \nu(a) \int dz \rho(z) \delta(\hat{A}(z) - a) \tag{2.7.20}$$

where $\nu(a)$ are the Lagrange multipliers enforcing the constraint (2.7.19). By setting to zero the functional derivative with respect to $\rho(z)$ gives

$$- \ln \frac{\overline{\rho}(z)}{\rho_N^0} - 1 - \nu \left(\hat{A}(z) \right) = 0 \tag{2.7.21}$$

this is

$$\overline{\rho}(z) = \rho_N^0 \exp\{-1 - \nu \left(\hat{A}(z) \right)\} \tag{2.7.22}$$

By inserting this relevant ensemble in the constraint (2.7.19) we obtain

$$\exp\{-1 - \nu(a)\} = \frac{P(a)}{\Omega(a)} \tag{2.7.23}$$

where we have introduced the measure

$$\Omega(a) = \int dz \rho_N^0 \delta(\hat{A}(z) - a) \tag{2.7.24}$$

In such a way that the relevant ensemble is

$$\overline{\rho}(z) = \rho_N^0 \frac{P(\hat{A}(z))}{\Omega(\hat{A}(z))} \tag{2.7.25}$$

In this level of description the relevant ensemble can be expressed *explicitly* in terms of the known macroscopic information $P(a)$.

The entropy in the probabilistic flavour is obtained by evaluating the Gibbs-Jaynes entropy (2.5.3) at the relevant ensemble (2.7.25) with the result

$$S[\overline{\rho}] = -\int dz \frac{P(\hat{A}(z))}{\Omega(\hat{A}(z))} \ln\left(\frac{P(\hat{A}(z))}{\Omega(\hat{A}(z))}\right) \tag{2.7.26}$$

By inserting the identity

$$\int da\delta(\hat{A}(z) - a) = 1 \tag{2.7.27}$$

inside the integral in (2.7.26) leads to

$$S[\overline{\rho}] = -\int daP(a) \ln\left(\frac{P(a)}{\Omega(a)}\right) \tag{2.7.28}$$

This is the explicit form of the entropy functional at the probabilistic flavour.

2.7.3   Mixed flavour for MD

In many occasions we will study the system through MD simulations for which the equilibrium ensemble is given by the microcanonical ensemble. In this case, we may need a mixed flavour in which the dynamic invariants are known with precision, in a microcanonical sense, but we are interested in average behaviour of the CG variables that evolve in time. In this section we study the resulting flavour.

Assume that the macroscopic information is given by $P(E)$, the probability distribution that the dynamic invariants $\hat{N}(z), \hat{H}(z)$ take values N, E, and the averages $A = \text{Tr}[\hat{A}\overline{\rho}]$ of the evolving CG variables. The macroscopic information is

$$1 = \int dz\rho(z)$$

$$P(E) = \int dz\rho(z)\delta(\hat{H}(z) - E)$$

$$A = \int dz\hat{A}(z)\rho(z) \tag{2.7.29}$$

The relevant ensemble is obtained by maximizing unrestrictedly the following La-

grangian

$$\mathcal{L}[\rho, \lambda, \nu, \mu] = -\int dz \rho(z) \ln \frac{\rho(z)}{\rho_N^0} - \int dz \rho(z) \nu(\hat{\mathcal{H}}(z))$$
$$- \lambda \left(\int dz \hat{A}(z) \rho(z) - a \right) - \mu \left(\int dz \rho(z) - 1 \right) \tag{2.7.30}$$

By setting to zero the functional derivative with respect to $\rho(z)$ gives the relevant ensemble

$$-\ln \frac{\overline{\rho}(z)}{\rho_N^0} - 1 - \nu \left(\hat{H}(z) \right) - \lambda \cdot \hat{A}(z) - \mu = 0 \tag{2.7.31}$$

this is

$$\overline{\rho}(z) = \rho_N^0 \exp \left\{ -1 - \mu - \nu \left(\hat{H}(z) \right) - \lambda \cdot \hat{A}(z) \right\} \tag{2.7.32}$$

Substitution into the restriction $P(E)$ gives

$$P(E) = \exp \left\{ -1 - \mu - \nu(E) \right\} \int dz \rho_N^0 \delta(\hat{H}(z) - E) \exp \left\{ -\lambda \cdot \hat{A}(z) \right\} \tag{2.7.33}$$

Substituting this into (2.7.32)

$$\overline{\rho}(z) = \rho_N^0 \frac{P(\hat{H}(z))}{\Omega(\hat{H}(z)) \left\langle \exp \left\{ -\lambda \cdot \hat{A} \right\} \right\rangle^{N,\hat{H}}} \exp \left\{ -\lambda \cdot \hat{A}(z) \right\} \tag{2.7.34}$$

which is the same as

$$\overline{\rho}(z) = \int dE P(E) \frac{\exp \left\{ -\lambda \cdot \hat{A}(z) \right\}}{\left\langle \exp \left\{ -\lambda \cdot \hat{A} \right\} \right\rangle^{NE}} \rho_{NE}^{\mathrm{mic}}(z) \tag{2.7.35}$$

where we have introduced the equilibrium microcanonical ensemble (27.0.15).

The entropy is obtained by evaluating the Gibbs-Jaynes entropy at the relevant ensemble. The result is

$$\mathcal{S}[\overline{\rho}] = -\int dz \overline{\rho}(z) \ln \frac{\overline{\rho}(z)}{\rho_N^0}$$
$$= -\int dz \left[\int dE P(E) \rho_{NE}^{\mathrm{mic}}(z) \frac{\exp \left\{ -\lambda \cdot \hat{A}(z) \right\}}{\left\langle \exp \left\{ -\lambda \cdot \hat{A} \right\} \right\rangle^{NE}} \right]$$
$$\times \left[\ln \left[\frac{P(\hat{H}(z))}{\Omega(\hat{H}(z))} \right] - \ln \left\langle \exp \left\{ -\lambda \cdot \hat{A} \right\} \right\rangle^{N,\hat{H}} - \lambda \cdot \hat{A}(z) \right] \tag{2.7.36}$$

This is, the entropy in the mixed flavour in which we know the probability $P(E)$ of the dynamic invariants and the averages of the CG variables A is given by

$$S[P, A] = -\int dE P(E) \ln \frac{P(E)}{\Omega(E)} + \int dE P(E) \ln \left\langle \exp\left\{-\lambda \cdot \hat{A}\right\}\right\rangle^{NE} + \lambda A \qquad (2.7.37)$$

This is a functional of $P(E)$ and a function of A. The conjugate variables are a function of A and a functional of $P(E)$ because

$$\frac{\partial S}{\partial A}[P, A] = \lambda \qquad (2.7.38)$$

Because $P(E)$ is given and known at the initial time, this is actually a one to one functional relationship between A and λ.

Consider a MD simulation with N_0 particles at the energy E_0. In this case, the probability distribution is given by

$$P(E) = \delta_{NN_0}\delta(E - E_0) \qquad (2.7.39)$$

and the entropy (2.7.37) in this case becomes

$$S[N_0, E_0, A] = \ln \Omega(N_0, E_0) + \ln \left\langle \exp\left\{-\lambda \cdot \hat{A}\right\}\right\rangle^{N_0 E_0} + \lambda A$$
$$= \ln \int dz_{N_0} \delta(\hat{H}(z) - E_0) \exp\left\{-\lambda \cdot \hat{A}\right\} + \lambda A \qquad (2.7.40)$$

The entropy of this mixed flavour in which the energy is known exactly but the coarse-grained variables are known on average is to be compared with the entropy (2.7.13) where the energy is known on average only, which is

$$S[\overline{N}, \overline{E}, , A] = \ln \int dz_{N_0} \exp\left\{-\alpha N - \beta \hat{H}\right\} \exp\left\{-\lambda \cdot \hat{A}\right\} + \alpha \overline{N} + \beta \overline{E} + \lambda A \qquad (2.7.41)$$

[I think that this result should allow to discuss that β does not depend on a, only on E in the usual averages flavour]

2.8   Summary

We have considered in this chapter the relevant ensembles that arise depending on the macroscopic information that we have available. When the information that we have are the averages of the dynamic invariants and the averages of the CG variables the relevant ensemble is given by (2.7.7) while when we know the averages of the dynamic invariants and the distribution of the CG variables we have the relevant ensemble given by (2.7.25). In both cases, we have the relevant ensemble as a modification of the macrocanonical equilibrium ensemble. When we consider the dynamic problem in the following chapters, we will always assume that the relevant ensemble is of this type. In the appendix we explore other possibilities for the relevant ensemble which are not given directly in terms of the equilibrium ensemble.

3

The stochastic dynamics of CG variables

3.1 The objective of coarse-graining

The main objective of the Theory of Coarse-Graining is to obtain a dynamical description of a system with much less information than the one given in the microscopic level. This is achieved by constructing dynamic equations for the time dependent averages of the CG variables. The CG variables change in time because the microscopic state of the system changes in time as shown in (1.4.32). Of course, because we do not know with certainty the initial microstate of the system, we do not know the actual value of the relevant variable at a given time with certainty. Therefore, it makes sense to take the average of the CG variables over initial conditions

$$a(t) \equiv \langle A \rangle_t = \int dz \, \hat{A}(\mathcal{U}_t z) \rho_0(z) \tag{3.1.1}$$

where $\rho_0(z)$ is the probability distribution over initial conditions. For simplicity, we understand $\int dz$ as a sum over microstates in a macrocanonical setting. Equivalently, we may write the average as

$$a(t) = \int dz \, \hat{A}(z) \rho_0(\mathcal{U}_{-t} z) = \int dz \, \hat{A}(z) \rho_t(z) \tag{3.1.2}$$

where we have used a change of variables $z \to \mathcal{U}_{-t} z$ with unit Jacobian, and (1.6.5). Therefore, the time dependent average over initial conditions is equal to the average of the CG variables with the real ensemble that is solution of the Liouville equation. By analogy with Quantum Mechanics Eqs. (3.1.1) may be called the Heisenberg representation for the average where the phase function is evolving, whereas (3.1.2) may be called the Schrodinger representation in which the ensemble is evolving in time.

The time derivative of the average of the CG variables is given by

$$\frac{d}{dt}a(t) = \int dz \hat{A}(z)\partial_t \rho_t(z) = -\int dz \hat{A}(z)i\mathcal{L}\rho_t(z) = \int dz \rho_t(z)i\mathcal{L}\hat{A}(z) \qquad (3.1.3)$$

where we have used the Liouville's equation (1.6.7). Of course, (3.1.3) is not a proper dynamic equation for $a(t)$ because the right hand side is not a function of $a(t)$ and this equation is not a *closed* differential equation for the averages. How to close this equation? Imagine for a moment that the solution of the Liouville's equation $\rho_t(z)$ could be very well approximated by a time-dependent relevant ensemble $\overline{\rho}_t(z)$ of the form (2.7.7) corresponding to the value of the average $a(t)$

$$\overline{\rho}_t(z) = \rho_N^0 \frac{\exp\{-\lambda(t)\hat{A}(z)\}}{Z(\lambda(t))} \qquad (3.1.4)$$

The conjugate variables $\lambda(t)$ are adjusted in such a way that the relevant ensemble does produce the correct macroscopic averages

$$a(t) = \int dz \hat{A}(z)\overline{\rho}_t(z) \qquad (3.1.5)$$

Because the connection between the averages and the conjugate variables is one to one, the relevant ensemble $\overline{\rho}_t(z)$ is fully determined by the average values $a(t)$. If we are allowed to approximate the real ensemble $\rho_t(z)$ by the relevant ensemble (3.1.4), this is

$$\rho_t(z) \approx \overline{\rho}_t(z) \qquad (3.1.6)$$

we obtain after introducing (3.1.6) into (3.1.3)

$$\frac{d}{dt}a(t) = \int dz \overline{\rho}_t(z)i\mathcal{L}\hat{A}(z) \qquad (3.1.7)$$

The right hand side of this equation is now a function of the averages through the relevant ensemble and, therefore, (3.1.7) is a closed equation for the averages.

Unfortunately, the real ensemble is not well approximated by the relevant ensemble in general. At later times, equations of the form (3.1.7) only capture the *reversible* part of the dynamics and all the interesting dissipative processes are eliminated in this approximation. What do we mean by this? Consider the time-dependent entropy function which is the result of evaluating the entropy (2.7.13) of the given level of description at the value of the average $a(t)$. If we take the time derivative we obtain

$$\frac{d}{dt}S(t) = \frac{\partial S}{\partial a}[a(t)]\frac{d}{dt}a(t) = -\lambda(t)\int dz \hat{A}(z)i\mathcal{L}\overline{\rho}_t(z) = \int dz i\mathcal{L}\overline{\rho}_t(z) = -\frac{d}{dt}\int dz \overline{\rho}_t(z) = 0$$
$$(3.1.8)$$

where we have used the chain rule, (2.7.14), (3.1.7), the form of the relevant ensemble (2.7.7) and, in the last identity, that the relevant ensemble is normalized. Therefore, the dynamic equation (3.1.7) does not produce any increase of the entropy function and can

be termed a reversible equation.

Despite of its shortcomings, the approximation (3.1.6) suggests the procedure in order to obtain closed equations for the averages. We may write the real ensemble as the sum of two contributions

$$\rho_t(z) = \overline{\rho}_t(z) + \delta\rho_t(z) \tag{3.1.9}$$

The term $\delta\rho_t(z)$ is known as the non-relevant part. If we could express the non-relevant part $\delta\rho_t(z)$ in terms of the relevant part $\overline{\rho}_t(z)$ we would solve the problem because then, the real ensemble $\rho_t(z)$ would be expressed in terms of the averages $a(t)$ on which the relevant ensemble depends. The main tenet of the Theory of Coarse-Graining is that one can fully express the real ensemble $\rho_t(z)$ in terms of the relevant ensemble at *present and past times*. That this can be achieved at all is shown by the following simple argument.

Assume that at the initial time the initial ensemble (to be used as initial condition of the Liouville's equation) is of the relevant type (2.7.7), this is

$$\rho_0(z) = \overline{\rho}_0(z) \tag{3.1.10}$$

The solution (1.6.5) of the Liouville's equation takes then the form

$$\rho_t(z) = \exp\{-i\mathcal{L}t\}\overline{\rho}_0(z) \tag{3.1.11}$$

We can then make use of the following trivial mathematical identity

$$\rho_t(z) = \overline{\rho}_t(z) - \int_0^t dt' \frac{d}{dt'} \left[\exp\{-i\mathcal{L}(t-t')\}\overline{\rho}_{t'}(z) \right] \tag{3.1.12}$$

This identity can be easily proved by performing explicitly the time integral of the full time derivative. Equivalently

$$\rho_t(z) = \overline{\rho}_0(\mathcal{U}_{-t}z) = \overline{\rho}_t(z) - \int_{t_0}^t dt' \frac{d}{dt'} \overline{\rho}_{t'}(\mathcal{U}_{t-t'}z) \tag{3.1.13}$$

Note that in this simple expression we have already written the solution $\rho_t(z)$ of the Liouville's equation with initial condition $\overline{\rho}_0(z)$ in terms of the relevant ensemble at previous times. Because the relevant ensemble depends on the averages $a(t)$ of the CG variables, *(3.1.12) gives the solution of the Liouville's equation in terms of the path history of the CG variables*. We could use now this equation to obtain a closed equation for the averages. We do not follow this route now but instead, we will follow the method of projection operators based on the Kawasaki-Gunton projector. It should be mentioned that (3.1.13) gives, in the Markov approximation, the same dynamic equations as the projection operator method. Projection operators are not essential to the Theory of Coarse-Graining, only a convenient tool. The essential concept is that of relevant ensemble.

3.2   Introduction

In many physical situations, it is important to keep track of not only the averages but also the *fluctuations* of the CG variables. Think, for example, on the case of a colloidal particle of small size suspended in a liquid solvent. This particle is subject to the bombardment of the surrounding molecules of the solvent and will cause the erratic motion of the center of mass of the colloidal particle. This erratic motion is known as *Brownian motion*. If we were to record the average value of the position of the center of mass of the colloidal particle, we would obtain a deceptively simple result. If the fluid is at equilibrium, the average position of the center of mass of the particle does not change. However, experimentally we have access to a much richer information about the colloidal particle. This information is encoded in the probability distribution of the position of the center of mass of the colloidal particle. This probability distribution does indeed change in time and describes how a colloidal particle diffuses as a consequence of the Brownian motion. While the average position computed with the time-dependent probability distribution is constant, the width of the probability describing the size of the fluctuations increases in time.

In the present chapter, we consider the probabilistic flavour of the Theory of Coarse-Graining. The CG variables are stochastic variables because the microstate is itself a stochastic variable, subject to the uncertainty on initial conditions. For example, assume that we know that the initial value of the CG variables is a_0. Many different microstates are compatible with this value (those of the submanifold $\hat{A}(z) = a_0$ of phase space). Two different microstates z_0, z_0' that are both compatible with a_0, this is, $\hat{A}(z_0) = \hat{A}(z_0') = a_0$ will evolve in such a way that the functions $A(\mathcal{U}_t z_0), A(\mathcal{U}_t z_0')$ will take different numerical values in general, i.e. $A(\mathcal{U}_t z_0) \neq A(\mathcal{U}_t z_0')$. Therefore, even if we know with certainty the actual initial state a_0 different outcomes are possible for the value of the CG variables at later times, just because there is no control over the actual microscopic state that produces the given macroscopic initial state a_0. The CG variables are inherently stochastic.

In this chapter we will obtain the dynamic equation for the probability distribution $P(a, t)$ of the CG variables. Eq. (1.7.4) shows that the probability distribution $P(a, t)$ is just the average of the following phase functions

$$\hat{\Psi}_a(z) = \delta(\hat{A}(z) - a) \tag{3.2.1}$$

We will refer to these phase functions constructed from the set of CG variables $\hat{A}(z)$ as **probabilistic CG variables**.

In many ways, the probabilistic flavour is to be preferred to the averages flavour when deriving evolution equations for CG variables. Obviously, probabilities give more information than averages and, as we will see, the cost for obtaining the equations is the same. Another argument is that, as we are interested in Markovian descriptions, we need to ensure that the set of CG variables includes *all* the slow variables of the system. In words of Green, they need to form a *complete set of gross variables*. If we forget some slow variable, then its effect will appear in the memory kernel that will not decay in a short time scale. In this case, the level of description will not be Markovian. But, if $\hat{A}(z)$ is a slow variable, we expect that any function of $\hat{A}(z)$ will be also a slow variable and

we should include in the description all functions of $\hat{A}(z)$ as CG variables also. The use of the variable $\hat{\Psi}_a(z)$ is a way to include all the functions of $\hat{A}(z)$ at once because

$$g(\hat{A}(z)) = \int da \hat{\Psi}_a(z) g(a) \tag{3.2.2}$$

and any function of $\hat{A}(z)$ is just a linear combination of the probabilistic CG variables. Therefore, as a rule of thumb we expect that a probabilistic description has a larger chance to be Markovian than an averages flavour, as *all* functions of the CG variables are included in the description.

However, there is one important class of levels of description that simply do not admit a probabilistic flavour, those where the CG variables $\hat{A}(z)$ are distributions themselves. The paradigmatic example is the microscopic density field

$$\hat{n}_{\mathbf{r}}(z) = \sum_i^N \delta(\mathbf{r} - \mathbf{q}_i) \tag{3.2.3}$$

These "spiky" CG variables are typical of field theories encountered in Vol 2. Note that $\hat{\Psi}_a(z) = \delta(\hat{A}(z) - a)$ is a distribution, not a function, and final results always have it inside expectations. Unfortunately, the composition of distributions does not exist[1] and, therefore, $\hat{A}(z)$ canot itself be a distribution like (3.2.3) but it must be a function. In this chapter, we will always assume that the CG variables $\hat{A}(z)$ are such that an equation like $\hat{A}(z) = a$ defines a proper (smooth and measurable) submanifold of phase space. This excludes a probabilistic description of CG variables of the form (3.2.3) whose evolution makes sense only in the averages flavour.

In this chapter, we obtain the evolution equation for the probability $P(a,t)$ that turns out to be a Fokker-Planck Equation (FPE). We show two different and complementary routes for the derivation of the FPE. The first route is due to Green [**?**] and is presented in Sec 3.4. He assumed that the *stochastic process* of CG variables was a Markov stochastic diffusion process. From this assumption, that requires the notion of separation of time scales, and the continuity of sample paths, and by looking at how the transition probability is linked to the microscopic Hamiltonian dynamics, he was able to produce microscopic expressions for the drift and diffusion terms in the FPE.

The second route to obtain the FPE is due to Zwanzig [**?**] and presented in Sec 3.5. His strategy was to obtain an *exact* generalized Fokker-Planck equation for $P(a,t)$ through a clever rewritting of Liouville's equation through projection operators. This exact equation contains memory (is an integro-differential equation). Again, under the assumption of clear separation of time scales, the exact equation can be approximated with a usual FPE that turns out to be identical to the one obtained by Green. Because the Markov approximation is not rigorous, approaching this issue from the complementary views of Green and Zwanzig offers further understanding of what is going on in the approximation process. Green and Zwanzig papers are arguably a cornerstone in the theory of Non-equilibrium Statistical Mechanics.

Neither Green's or Zwanzig's approaches payed much attention to the role of microscopic dynamic invariants in the CG dynamics. With a careful attention to this role, a

---

[1]What is the mathematical meaning of a Dirac delta function evaluated inside a Dirac delta function?

powerful structure is revealed from the FPE originally obtained by Green and Zwanzig. This structure, known as the GENERIC structure, was introduced phenomenologically by Öttinger and Grmela in 1997 [**?, ?**] and its statistical mechanics undepinning was given by Öttinger shortly after [**?**]. The GENERIC approach is fully discussed in the textbook [**?**].

3.3   Diffusion Markov processes

In this section we review the definition and some properties of the basic stochastic ingredients that are needed in the Theory of Coarse-Graining. We aim at an intuitive physicists perspective leaving aside mathematical rigour. More details, within this perspective, may be found in Gardiner [**?**] and Risken [**?**] textbooks.

A stochastic process is fully characterized [**?**] with the hierarchy of **joint probability distributions** $P(a_1, t_1, \cdots, a_n, t_n)$, *for all n*, of obtaining the value a_1 of the stochastic variable at time t_1 *and* the value a_2 at time t_2 etc. where $t_1 < t_2 < \cdots < t_n$. The one time probability $P(a_1, t_1)$ and the two time joint probability $P(a_1, t_1, a_2, t_2)$ allow one to introduce the **conditional probability**, also referred as **transition probability**, as

$$P(a_1, t_1 | a_2, t_2) = \frac{P(a_1, t_1, a_2, t_2)}{P(a_1, t_1)} \tag{3.3.1}$$

The one time probability is obtained by marginalization of the two time joint probability in the form

$$P(a_2, t_2) = \int da_1 P(a_1, t_1, a_2, t_2) \tag{3.3.2}$$

because both sides of this equation are the probability of finding a_2 at time t_2 irrespective of the value of a_1 at time t_1. This equation gives the following integral equation relating the one time probability and the transition probability

$$P(a_2, t_2) = \int da_1 P(a_1, t_1) P(a_1, t_1 | a_2, t_2) \tag{3.3.3}$$

This integral equation allows to "evolve" the one time probability.

A Markov process is characterized by the fact that the n-time joint probability can be fully expressed in terms of the one-time probability and the transition probability, in the form

$$P(a_1, t_1, \ldots, a_n, t_n) = P(a_1, t_1) P(a_1, t_1 | a_2, t_2) \cdots P(a_{n-1}, t_{n-1} | a_n, t_n) \tag{3.3.4}$$

Therefore, if we know $P(a_1, t_1)$ and $P(a_1, t_1 | a_2, t_2)$ we know the stochastic process in full. If we consider the conditional probability $P_2(a_1, t_1 | a_2, t_2, a_3, t_3)$ of having a_2 at t_2 and a_3 at t_3 provided that we had a_1 at t_1, the Markov property states

$$P(a_1, t_1 | a_2, t_2, a_3, t_3) = P(a_1, t_1 | a_2, t_2) P(a_2, t_2 | a_3, t_3) \tag{3.3.5}$$

If we integrate (3.3.5) over a_2 we need to have the following consistency condition

$$P_2(a_1, t_1 | a_3, t_3) = \int da_2 P(a_1, t_1 | a_2, t_2) P(a_2, t_2 | a_3, t_3) \qquad (3.3.6)$$

This consistency condition is known as the **Chapman-Kolmogorov equation** for the transition probabilities. The intuitive idea associated with this equation is that the probability of a transition from a_1 to a_3 is given as the sum of all the transition probabilities over an intermediate state a_2.

A Markov process with continuous paths is termed a diffusion process and, as mentioned, is fully characterized by the one time probability distribution and the transition probability. There are several equivalent ways to describe the diffusion process. We will typically consider three different ways. The first way is in terms of the Fokker-Planck equation governing both the one time probability and the conditional probability. The second, mathematically equivalent way is through the stochastic differential equation governing the generation of a particular realization of the stochastic process. And the third way is through the Onsager-Machlup functional path integral, that gives the probability of obtaining a particular realization of a given path of the stochastic variable.

The Fokker-Planck equation

The Chapman-Kolmogorov is an integral equation that links all the transition probabilities of a Markov process. There exists an equivalent differential form for the Chapman-Kolmogorov that, under the assumption that stochastic process is continuous becomes the **Fokker-Planck equation**. The derivation of the Fokker-Planck equation from the Chapman-Kolmogorov equation is presented in [**?**] and we only quote the final result. The Fokker-Planck equation governs the one time probability distribution

$$\frac{\partial}{\partial t} P(a, t) = -\frac{\partial}{\partial a_\mu} D_\mu^{(1)}(a) P(a, t) + \frac{1}{2} \frac{\partial^2}{\partial a_\mu \partial a_\nu} D_{\mu\nu}^{(2)}(a) P(a, t) \qquad (3.3.7)$$

Here the vector $D^{(1)}(a)$ is called **the drift term** while the matrix $D^{(2)}(a)$ is called **the diffusion matrix**.

Forward and backward differential Chapman-Enskog equations

Note that, from (3.3.3) if we take at $t_1 = t_0$ as initial condition $P(a_1, t_1) = \delta(a_1 - a_0)$ then $P(a, t) = P(a_0, t_0 | a, t)$, this is, the transition probability is identical to the one-time probability with a Dirac delta initial condition. As a consequence, the transition probability also satisfies the Fokker-Planck equation (3.3.7)

$$\frac{\partial}{\partial t} P(a_0, t_0 | a, t) = -\frac{\partial}{\partial a_\mu} D_\mu^{(1)}(a) P(a_0, t_0 | a, t) + \frac{1}{2} \frac{\partial^2}{\partial a_\mu \partial a_\nu} D_{\mu\nu}^{(2)}(a) P(a_0, t_0 | a, t) \quad (3.3.8)$$

where, by definition, we have that the initial condition for this equation is

$$P(a_0, t_0 | a, t_0) = \delta(a - a_0) \qquad (3.3.9)$$

Eq (3.3.8) is referred to as the **forward differential Chapman-Enskog equation** [**?**]. The **backward differential Chapman-Enskog equation** is a mathematically equivalent equation that is obtained by taking the conditions as independent variables. Its explicit form is given by [**?**]

$$\frac{\partial}{\partial t_0} P(a_0, t_0 | a, t) = -D_\mu^{(1)}(a_0) \frac{\partial}{\partial a_{0\mu}} P(a_0, t_0 | a, t) - \frac{1}{2} D_{\mu\nu}^{(2)}(a_0) \frac{\partial^2}{\partial a_{0\mu} \partial a_{0\nu}} P(a_0, t_0 | a, t)$$

(3.3.10)

Note that, as compared with the forward equation, the drift and diffusion precede the derivatives and the sign of the diffusion term is reversed.

The conditional probability at short times

The solution of the Fokker-Planck equation (3.3.8) is difficult to obtain in general due to the fact that the objects $D^{(1)}(a), D^{(2)}(a)$ may depend in principle in a non-linear way on the state a. Nevertheless, it is possible to obtain an explicit solution *for short times*. Because the initial condition (3.3.9) of the transition probability is a Dirac delta, we expect that for sufficiently short times $t = t_0 + \Delta t$ with $\Delta t > 0$ and $\Delta t \approx 0$, the transition probability will remain highly peaked. In this case, we may approximate in (3.3.8) $D^{(1)}(a) \approx D^{(1)}(a_0)$ and $D^{(2)}(a) \approx D^{(2)}(a_0)$, because the derivatives of these functions are much smaller than the derivatives of the highly peaked probability. This results in a Fokker-Planck equation with constant coefficients

$$\frac{\partial}{\partial t} P(a_0, t_0 | a, t) \simeq -D_\mu^{(1)}(a_0) \frac{\partial}{\partial a_\mu} P(a_0, t_0 | a, t) + D_{\mu\nu}^{(2)}(a_0) \frac{1}{2} \frac{\partial^2}{\partial a_\mu \partial a_\nu} P(a_0, t_0 | a, t)$$

(3.3.11)

This difussion equation with constant coefficients is now easy to solve. One way is by considering the Fourier transform

$$P(a_0, t_0 | k, t) = \int da \exp\{-ik{\cdot}a\} P(a_0, t_0 | a, t)$$

$$P(a_0, t_0 | a, t) = \frac{1}{(2\pi)^M} \int dk \exp\{ik{\cdot}a\} P(a_0, t_0 | k, t)$$

(3.3.12)

The Fourier components evolve according to

$$\partial_t P(a_0, t_0 | k, t) = -iD^{(1)}(a_0){\cdot}k P(a_0, t_0 | k, t) - \frac{1}{2} k^T {\cdot} D^{(2)}(a_0){\cdot}k P(a_0, t_0 | k, t)$$

(3.3.13)

that has as solution

$$P(a_0, t_0 | k, t) = \exp\left\{ \left[-iD^{(1)}(a_0){\cdot}k - \frac{1}{2} k^T {\cdot} D^{(2)}(a_0){\cdot}k \right] (t - t_0) \right\} P(a_0, t_0 | k, t_0) \quad (3.3.14)$$

Going back to real space

$$P(a_0, t_0|a, t) = \frac{1}{(2\pi)^M} \int dk \, \exp\{ik{\cdot}a\} \exp(t - t_0) \left\{ -iD^{(1)}(a_0){\cdot}k - \frac{1}{2}k^T{\cdot}D^{(2)}(a_0){\cdot}k \right\} P(a_0, t_0|k, t_0)$$

$$(3.3.15)$$

The initial condition $P(a, t_0|a_0, t_0) = \delta(a - a_0)$ in Fourier space is $P(k, t_0 0|a_0, t_0) = \exp\{-ik{\cdot}a_0\}$ and, therefore,

$$P(a_0, t_0|a, t) = \frac{1}{(2\pi)^M} \int dk \, \exp\{ik{\cdot}(a - a_0)\} \exp(t - t_0) \left\{ -iD^{(1)}(a_0){\cdot}k - \frac{1}{2}k^T{\cdot}D^{(2)}(a_0){\cdot}k \right\}$$

$$= \frac{1}{(2\pi)^M} \int dk \, \exp\{ik{\cdot}(a - a_0 - (t - t_0)D^{(1)}(a_0))\} \exp \left\{ -\frac{t}{2}k^T{\cdot}D^{(2)}(a_0){\cdot}k \right\}$$

$$(3.3.16)$$

when the inverse of $D^{(2)}$ exists, we have the explicit Gaussian solution

$$P(a_0, t_0|a_1, t_0 + \Delta t) = \exp \left\{ -\frac{1}{2\Delta t} (a_1 - a_0(\Delta t)) \, D_{(2)}^{-1}(a_0) \, (a_1 - a_0(\Delta t)) \right\}$$

$$\times \frac{1}{(2\pi\Delta t)^{M/2} \det(D^{(2)}(a_0))^{1/2}} \qquad (3.3.17)$$

where we have introduced $a_0(\Delta t)$ as the initial value evolved with the drift

$$a_0(\Delta t) = a_0 + \Delta t D^{(1)}(a_0) \qquad (3.3.18)$$

The transition probability is a Gaussian centered around $a_0(\Delta t)$ and with a width given by the inverse of the diffusion matrix $D^{(2)}(a_0)$, provided this inverse exists. In systems with inertia the inverse does not exist but it is nevertheless still possible to write down the transition probability, that will include some aditional delta functions [**?**]. The time scale in which the transition probability has this Gaussian form is for times Δt in which the stochastic variables have hardly evolved.

Kramers-Moyal coefficients

The drift vector $D^{(1)}(a)$ and the diffusion matrix $D^{(2)}(a)$ appearing in (3.3.7) are given in terms of moments of the transition probability at short times. Its particular form is, actually, specified as assumptions in the derivation of the Fokker-Planck equation from the Chapman-Kolmogorov equation [**?**]. We may also obtain the specific form of $D^{(1)}(a), D^{(2)}(a)$ from the following argument that will shed light on the physical aspects involved.

The transition probability (3.3.17) has the following moments

$$\int da_1 (a_1 - a_0) P(a_0, t_0|a_1, t_0 + \Delta t) = D^{(1)}(a_0)\Delta t$$

$$\int da_1 [a_1 - a_0(\Delta t)][a_1 - a_0(\Delta t)]^T P(a_0, t_0|a_1, t_0 + \Delta t) = D^{(2)}(a_0)\Delta t \qquad (3.3.19)$$

Therefore, the drift and diffusion terms are given in terms of the transition probability as

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int da_1 (a_1 - a_0) P(a_0, t_0 | a_1, t_0 + \Delta t)$$

$$D^{(2)}(a_0) = \frac{1}{\Delta t} \int da_1 [a_1 - a_0(\Delta t)][a_1 - a_0(\Delta t)]^T P(a_0, t_0 | a_1, t_0 + \Delta t)$$

$$= \frac{1}{\Delta t} \int da_1 [a_1 - a_0][a_1 - a_0]^T P(a_0, t_0 | a_1, t_0 + \Delta t) + D^{(1)}[D^{(1)}]^T \Delta t \quad (3.3.20)$$

We will neglect the last term in the diffusion matrix which is small for small Δt. Equations (3.3.20) are known as the **Kramers-Moyal** expression for the drift and diffusion terms. They give the objects of the Fokker-Planck equation in terms of the conditional probability and will play a crucial role in the development of microscopic expressions for these terms.

The Stochastic Differential Equation

We now introduce the notion of stochastic differential equation. We do it in a very pedestrian way by starting with the following updating difference equation

$$a'_\mu = a_\mu + D_\mu^{(1)}(a)\Delta t + B_{\mu\nu}(a)\cdot\zeta_\nu \Delta t^{1/2} \quad (3.3.21)$$

where $a'_\mu = a_\mu(t + \Delta t)$ and $a_\mu = a_\mu(t)$ are the future and past values of the stochastic variable a, and ζ is a M-dimensional vector of independent Gaussian random numbers distributed according to the following probability density

$$G(\zeta) = \frac{1}{(2\pi)^{M/2}} \exp\left\{-\frac{1}{2}\zeta^T\cdot\zeta\right\} \quad (3.3.22)$$

Therefore, the variables ζ have zero mean and unit variance. The matrix $B_{\mu\nu}(a)$ is selected in such a way that it is the square root, in matrix sense, of the diffusion matrix, this is

$$B_{\mu\sigma}(a)B_{\nu\sigma}(a) = D_{\mu\nu}^{(2)}(a) \quad (3.3.23)$$

The difference equation (3.3.21) can be understood as a way to generate states a' in the future, from the state a in the present.

Because of the random term ζ the resulting a' will be also a random variable. The objective now is to find out what is the probability distribution of a', given that we know the probability distribution of ζ. We denote this probability as $P(a|a', \Delta t)$ which is a conditional probability of finding a' at a later time Δt given that at present we know that the variable has the value a. Note that when we say that ζ is "independent" what we mean is that it is statistically independent of both a' and a, this is ζ is uncorrelated with the present and future values of the stochastic variable a.

In order to find out $P(a|a', \Delta t)$ we may understand a' in (3.3.21) as a function (in this case a linear function) of ζ. Therefore we have an induced probability on a' due to the probability density (3.3.22) of ζ. The probability density that a function of a stochastic

variable takes particular values is given by (1.7.2). Particularizing in the present case we have that

$$P(a|a', \Delta t) = \int d\zeta G(\zeta) \delta \left(a' - a - D^{(1)}(a)\Delta t - B(a)\cdot\zeta\Delta t^{1/2} \right) \tag{3.3.24}$$

This integral over a Gaussian is of the form (23.0.5) and has, therefore, the following explicit form, when the inverse of $D^{(2)}(a)$ defined in (3.3.23) exists,

$$P(a|a', \Delta t) = \exp \left\{ -\frac{1}{2\Delta t} \left(a' - a - D^{(1)}(a)\Delta t \right) D^{-1}_{(2)}(a) \left(a' - a - D^{(1)}(a)\Delta t \right) \right\}$$
$$\times \frac{1}{(2\pi\Delta t)^{M/2} \det(D^{(2)}(a))^{1/2}} \tag{3.3.25}$$

which is of the same form as (3.3.17).

Therefore, the stochastic variable a' is distributed with the same conditional probability that governs the diffusion process. This means that the difference equation (3.3.21) may be used to generate realizations of the stochastic process that is described, in the limit of small Δt by the Fokker-Planck equation.

Note that we can also write (3.3.21) in the form

$$a_\mu(t_{n+1}) = a_\mu(t_n) + D^{(1)}_\mu(a(t_n))\Delta t + B_{\mu\nu}(a(t_n))\cdot\zeta_\nu(t_n)\Delta t^{1/2} \tag{3.3.26}$$

that generates the sequence of random numbers $a(t_n)$ for $t_n = n\Delta t$, $n = 0, 1, 2, \cdots$, which is a realization of the stochastic process, under the realization $\zeta(t_n)$ of the noise. A Stochastic Differential Equation can be thought of the limit $\Delta t \to 0$ of the difference equation (3.3.26) and is usually written in the form

$$da(t) = D^{(1)}(a(t))dt + B(a(t))\cdot dW_t \tag{3.3.27}$$

where dW_t are independent increments of the Wiener process. We refer to [?] for a clear exposition of SDEs. The SDE (3.3.27) is a mathematical representation of the Markov Diffusion process which is equivalent to the FPE (3.3.7). The stochastic interpretation of the SDE is the Ito interpretation, that states that the Wiener increments dW_t are uncorrelated with the present and future values of the stochastic variable $a(t)$. For the purposes of the present book, whenever we encounter a SDE of the form (3.3.27) we may always go back to its discrete representation (3.3.21). Eq (3.3.21) is actually Euler's method for the solution of the SDE (3.3.27)[2]. We note in passing that we choose the notation Δt for the upper limit of integration in the Green-Kubo formula because it is the minimum physical value of the time step that can be used in the solution of the SDE.

Ito formula

The fact that the independent increments of the Wiener process scale as $dt^{1/2}$ has subtle consequences when we investigate how a function of the stochastic variables changes. If

---

[2]In practice, Euler method requires to use very small time steps. One should use more sophisticate methods that allow for larger time steps [?] .

we go back to the difference equation (3.3.21) and consider the variation $\Delta F(a)$ of a differentiable function $F(a)\mathbb{R}^M \to \mathbb{R}$ when we update the state we have

$$\Delta F(a) = F(a') - F(a) = F\left(a + D^{(1)}(a)\Delta t + B(a)\cdot\zeta\Delta t^{1/2}\right) - F(a) \qquad (3.3.28)$$

By Taylor expanding to second order (we will see why to second order) we obtain

$$\begin{aligned}
\Delta F(a) &= \frac{\partial F(a)}{\partial a_\mu}\left(D^{(1)}_\mu(a)\Delta t + B_{\mu\nu}(a)\zeta_\nu\Delta t^{1/2}\right) \\
&\quad + \frac{1}{2}\frac{\partial^2 F(a)}{\partial a_\mu \partial a_\nu}\left(D^{(1)}_\mu(a)\Delta t + B_{\mu\mu'}(a)\zeta_{\mu'}\Delta t^{1/2}\right)\left(D^{(1)}_\nu(a)\Delta t + B_{\nu\nu'}(a)\zeta_{\nu'}\Delta t^{1/2}\right) \\
&= \frac{\partial F(a)}{\partial a_\mu}B_{\mu\nu}(a)\zeta_\nu\Delta t^{1/2} \\
&\quad + \left(\frac{\partial F(a)}{\partial a_\mu}D^{(1)}_\mu(a) + \frac{1}{2}\frac{\partial^2 F(a)}{\partial a_\mu \partial a_\nu}B_{\mu\mu'}(a)\zeta_{\mu'}B_{\nu\nu'}(a)\zeta_{\nu'}\right)\Delta t + \mathscr{O}(\Delta t^{3/2}) \quad (3.3.29)
\end{aligned}$$

where we have grouped terms in increasing powers of Δt. We observe that because of the dependence on $\Delta t^{1/2}$ of the random terms, an additional term involving second derivatives appear in the linear order term.

Is it clear that because ζ is a random vector, $\Delta F(a)$ will also be a random variable. The conditional expectation (conditional on the present state a) is given by

$$\langle\Delta F(a)\rangle = \left(\frac{\partial F(a)}{\partial a_\mu}D^{(1)}_\mu(a)\Delta t + \frac{1}{2}\frac{\partial^2 F(a)}{\partial a_\mu \partial a_\nu}D^{(2)}_{\mu\nu}(a)\right)\Delta t + \mathscr{O}(\Delta t^{3/2}) \qquad (3.3.30)$$

This identity suggest that whenever we need to compute the infinitesimal increment of a function of a stochastic variable governed by the Ito SDE (3.3.27) we have to expand the function to second order, this is

$$dF(a) = \left(\frac{\partial F(a)}{\partial a_\mu}D^{(1)}_\mu(a) + \frac{1}{2}\frac{\partial^2 F(a)}{\partial a_\mu \partial a_\nu}D^{(2)}_{\mu\nu}(a)\right)dt + \frac{\partial F(a)}{\partial a_\mu}B_{\mu\nu}(a)dW_\nu(t) \qquad (3.3.31)$$

This is the **Ito formula** for computing the differential of a function of stochastic variable governed by the SDE (3.3.27).

Onsager-Machlup path integral
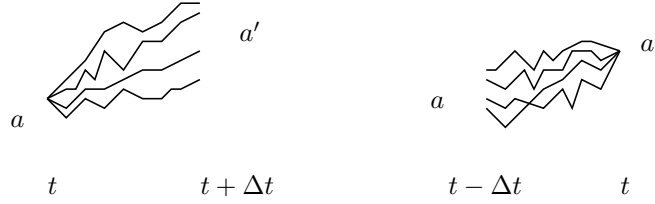
Figure 3.1: The forward (left) and backward derivatives (right) are computed from trajectories that start at a or end at a' at the present time t, respectively.

Forward and backward mean derivatives

The concept of forward and backward mean derivatives of a stochastic process was introduced by Nelson [?, ?]. The forward mean derivative is defined as

$$Da(t) \equiv \lim_{\Delta t \to 0^+} \int da' \frac{(a' - a)}{\Delta t} P(a, t | a', t + \Delta t) \tag{3.3.32}$$

The forward mean derivative gives an idea of how much the state drifts, on average, when we start at a and look for very small times ahead in the future. The backward mean derivative is defined according to

$$D_* a(t) \equiv \lim_{\Delta t \to 0^+} \int da \frac{(a' - a)}{\Delta t} P(a', t | a, t - \Delta) \tag{3.3.33}$$

Note that in this case, we are considering an average of the "velocity" over all the processes that start at some a in the past $t - \Delta t$ and end up at the same a' in the present t. Both forward and backward mean derivatives depend on the conditional state a or a', respectively, and are, therefore, stochastic variables themselves.

The objective is to express the forward and backward derivatives in terms of the drift vector $D^{(1)}(a)$ and diffusion matrix $D^{(2)}(a)$. As for the forward derivative, we recognize simply from the Kramers-Moyal coefficient (3.3.20) that it is given by the drift term

$$Da(t) = D^{(1)}(a) \tag{3.3.34}$$

The backward derivative is far more tricky because, to start with, we do not know what is the *backward* conditional probability of finding the value a of the stochastic variable in the past, conditional on the value a' in the future. Note that, up to now, we have always assumed that the conditional probabilities were defined as "the future conditional on the past". One way to get the meaning of the backward conditional probability is through a similar argument to the one that we have used for the calculation of the forward conditional probability (3.3.25) from the difference equation (3.3.21).

$$a'_\mu = a_\mu + D^{(1)}_\mu(a)\Delta t + B_{\mu\nu}(a) \cdot \zeta_\nu \Delta t^{1/2} \tag{3.3.35}$$

The question that we pose now is, according to the finite difference equation (3.3.21), what is the conditional probability of finding the present value a given that we know for sure the value a' in the future? The difference equation (3.3.21 allow us to understand a as a function of ζ, this is $a = F(\zeta)$, where the function $F(\zeta)$ is defined from its inverse, which is

$$F^{-1}(a) = \frac{1}{\Delta t^{1/2}} B^{-1}(a)(a' - a - D^{(1)}(a)\Delta t) \tag{3.3.36}$$

The backward conditional probability will be simply

$$P(a'|a, \Delta t) = \int d\zeta\, G(\zeta)\delta\left(a - F(\zeta)\right) \tag{3.3.37}$$

By considering the change of variable $y = F(\zeta)$ with $dy = \left|\frac{\partial F(\zeta)}{\partial \zeta}\right| d\zeta$

$$\begin{aligned} P(a'|a, \Delta t) &= \int dy \left|\frac{\partial F^{-1}(y)}{\partial y}\right| G\left(F^{-1}(y)\right)\delta\left(a - y\right) \\ &= \left|\frac{\partial F^{-1}(a)}{\partial a}\right| G\left(F^{-1}(a)\right) \end{aligned} \tag{3.3.38}$$

Note that this conditional probability is normalized when we integrate with respect to a, as it should. We can, if we please, compute explicitly the Jacobian matrix. However, it is not needed in what follows.

With this expression for the backward conditional probability we can now compute explicitly the expression of the backward mean derivative. By substitution of (3.3.38) into (**??**).

$$D_*a(t) \equiv \lim_{\Delta t \to 0^+} \int da \frac{(a' - a)}{\Delta t} \left|\frac{\partial F^{-1}(a)}{\partial a}\right| G\left(F^{-1}(a)\right) \tag{3.3.39}$$

Perfoming the change of variable $F^{-1}(a) = y$ we obtain the following expression for the backward mean derivative

$$D_*a(t) \equiv \lim_{\Delta t \to 0^+} \int dy \frac{(a' - F(y))}{\Delta t} G(y) \tag{3.3.40}$$

This integral is not inmediate because $F(y)$ is known explicitly only through its inverse, defined in (3.3.36). For this reason, we will compute the function $F(y)$ for small values of Δt, which is all what we need in order to compute the backward mean derivative. We assume that the function $F(y)$ has the following expansion in terms of Δt

$$a = F(y) = F_0(y) + F_1(y)\Delta t^{1/2} + F_2(y)\Delta t + \mathscr{O}(\Delta t^{3/2}) \tag{3.3.41}$$

The functions $F_0(y), F_1(y), F_2(y)$ are obtained by requiring that, order by order, the Eq (3.3.36) is fullfilled. By inserting a in (3.3.41) into (3.3.36) we have

$$
\begin{aligned}
y &= F^{-1}(a) \\
&= F^{-1}\left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right) \\
&= \frac{1}{\Delta t^{1/2}} B^{-1}\left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right) \\
&\quad \times \left(a' - \left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right) - D^{(1)}\left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right)\Delta t\right)
\end{aligned}
\tag{3.3.42}
$$

Which can be expressed as

$$
\begin{aligned}
\Delta t^{1/2}B\left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right)y = a' &- \left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right) \\
&- D^{(1)}\left(F_0 + F_1\Delta t^{1/2} + F_2\Delta t\right)\Delta t
\end{aligned}
\tag{3.3.43}
$$

The limit $\Delta t \to 0$ gives the zero-th order term, which is

$$
F_0(y) = a'
\tag{3.3.44}
$$

Then

$$
\Delta t^{1/2}B\left(a' + F_1\Delta t^{1/2} + F_2\Delta t\right)y = -F_1\Delta t^{1/2} - F_2\Delta t - D^{(1)}\left(a' + F_1\Delta t^{1/2} + F_2\Delta t\right)\Delta t
\tag{3.3.45}
$$

By collecting the terms of order $\Delta t^{1/2}$ we have

$$
F_1(y) = -B\left(a'\right)y
\tag{3.3.46}
$$

The order Δt terms are

$$
\Delta t^{1/2}B\left(a' + F_1\Delta t^{1/2} + F_2\Delta t\right)y = -F_1\Delta t^{1/2} - F_2\Delta t - D^{(1)}(a')\Delta t
\tag{3.3.47}
$$

$$
\frac{\partial B_{\mu\nu}}{\partial a'_\sigma}(a')F_{1\sigma}y_\nu = -F_{2\mu} - D_\mu^{(1)}(a')
\tag{3.3.48}
$$

which on account of (3.3.46) gives

$$
\frac{\partial B_{\mu\nu}}{\partial a'_\sigma}(a')B_{\sigma\sigma'}\left(a'\right)y_{\sigma'}y_\nu = F_{2\mu} + D_\mu^{(1)}(a')
\tag{3.3.49}
$$

Back to (3.3.41) we have

$$F_\mu(y) = a'_\mu - B_{\mu\nu}(a') y_\nu \Delta t^{1/2} + \left(-D_\mu^{(1)}(a') + \frac{\partial B_{\mu\nu}}{\partial a'_\sigma}(a') B_{\sigma\sigma'}(a') y_{\sigma'} y_\nu \right) \Delta t + \mathscr{O}(\Delta t^{3/2})$$

$$(3.3.50)$$

By inserting now this expansion into the form (3.3.40) for the backward mean derivative we have

$$D_* a_\mu(t) \equiv \lim_{\Delta t \to 0^+} \int dy \frac{-B_{\mu\nu}(a') y_\nu \Delta t^{1/2} + \left(-D_\mu^{(1)}(a') + \frac{\partial B_{\mu\nu}}{\partial a'_\sigma}(a') B_{\sigma\sigma'}(a') y_{\sigma'} y_\nu \right) \Delta t + \mathscr{O}(\Delta t^{3/2})}{\Delta t} G(y)$$

$$(3.3.51)$$

Note that the potentially dangerous term of order $\Delta t^{-1/2}$ vanishes because y has zero mean. By using that the covariance $\langle y_\sigma y_\nu \rangle = \delta_{\sigma\nu}$, the final result for the backward mean derivative

$$D_* a_\mu(t) = -D_\mu^{(1)}(a') + \frac{\partial B_{\mu\nu}}{\partial a'_\sigma}(a') B_{\sigma\nu}(a') \qquad (3.3.52)$$

[There must be an error, because I would like to express the forward derivative fully in terms of $D_{\mu\nu}^{(2)}$ and not to depend on the way the diffusion matrix is squared in terms of the B matrix]

[I have not understood yet the concept of forward and backward derivatives. In particular, I do not obtain the Nelson result for Ornstein-Uhlenbeck process]

Alternative derivation, not entirely succesful: In order to find the backward conditional probabity we return to the original definition of conditional probability in terms of the joint probability $P(a, t, a', t')$ of finding a at t and a' at t'. Note that in the joint probability there is no ordering of the times implied. From the joint probability we may define two conditional probabilities

$$P(a, t, a', t') = P(a, t)P(a, t|a', t') = P(a', t')P(a', t'|a, t) \qquad (3.3.53)$$

which implies

$$P(a', t'|a, t) = \frac{P(a, t)}{P(a', t')} P(a, t|a', t') \qquad (3.3.54)$$

If we understand primed variables as "future" variables this equation allows to define the backward conditional probability $P(a', t'|a, t)$ of finding the value a in the past t given the value a' in the future t'.

We seek now an equation of motion for the backward conditional probability. It proves convenient to write the Fokker-Planck equations for the one time probability (3.3.7), for the forward conditional probability (3.3.8), and the backward differential

Chapman-Kolmogorov equation (3.3.10) in the following form

$$\partial_t P(a,t) = \mathcal{L}P(a,t)$$
$$\partial_{t'} P(a,t|a',t') = \mathcal{L}'P(a,t|a',t')$$
$$\partial_t P(a,t|a',t') = \mathcal{L}^B P(a,t|a',t') \tag{3.3.55}$$

where the operator $\mathcal{L}, \mathcal{L}', \mathcal{L}^B$ are

$$\mathcal{L} \equiv -\frac{\partial}{\partial a_\mu} D^{(1)}_\mu(a) + \frac{1}{2}\frac{\partial^2}{\partial a_\mu \partial a_\nu} D^{(2)}_{\mu\nu}(a)$$

$$\mathcal{L}' \equiv -\frac{\partial}{\partial a'_\mu} D^{(1)}_\mu(a') + \frac{1}{2}\frac{\partial^2}{\partial a'_\mu \partial a'_\nu} D^{(2)}_{\mu\nu}(a')$$

$$\mathcal{L}^B \equiv -D^{(1)}_\mu(a)\frac{\partial}{\partial a_\mu} - \frac{1}{2}D^{(2)}_{\mu\nu}(a)\frac{\partial^2}{\partial a_\mu \partial a_\nu} \tag{3.3.56}$$

We now take the derivative with respect to t (which is a "past" time) of Eq (3.3.54)

$$\partial_t P(a',t'|a,t) = \frac{1}{P(a',t')}P(a,t|a',t')\mathcal{L}P(a,t) + \frac{P(a,t)}{P(a',t')}\partial_t P(a,t|a',t')$$

$$= \frac{P(a',t'|a,t)}{P(a,t)}\mathcal{L}P(a,t) + \frac{P(a,t)}{P(a',t')}\mathcal{L}^B P(a,t|a',t')$$

$$= \frac{P(a',t'|a,t)}{P(a,t)}\mathcal{L}P(a,t) + \frac{P(a,t)}{P(a',t')}\mathcal{L}^B \frac{P(a',t')}{P(a,t)}P(a',t'|a,t)$$

$$= \frac{P(a',t'|a,t)}{P(a,t)}\mathcal{L}P(a,t) + P(a,t)\mathcal{L}^B \frac{1}{P(a,t)}P(a',t'|a,t) \tag{3.3.57}$$

3.4   Coarse-graining à la Green

Green in his 1952 remarkable paper presented the essentials of the theory of coarse-graining as we know it. The basic assumption taken by Green and on which the whole construction of coarse-graining is based is that *the stochastic process of the CG variables is a continuous diffusion Markov process.* As we will see in this section, this single hypothesis is sufficient to obtain a dynamical equations for the CG variables, the Fokker-Planck equation, with all the objects appearing in the dynamics defined in terms of microscopic expressions. The crucial insight of Green was to note that, when the CG variables are assumed to evolve following a diffusion process, then the transition probability has also a microscopic expression that when substituted in (3.3.20) leads to microscopic expressions for the drift and diffusion terms.

Microscopic expression for the transition probability

The microscopic expression for the transition probability (3.3.1) can be obtained from the microscopic expressions for the one time and two time probabilities given in (1.7.5) and (1.7.9), this is

$$P(a_0, 0|a_1, t_1) = \frac{P(a_0, 0; a_1, t_1)}{P(a_0, 0)} = \frac{\int dz_0 \rho(z_0, 0)\delta(\hat{A}(z_0) - a_0)\delta(\hat{A}(U_{t_1} z_0) - a_1)}{\int dz_0 \rho(z_0, 0)\delta(\hat{A}(z_0) - a_0)} \quad (3.4.1)$$

It is possible to go one step further if the initial ensemble is specified. Let us assume that after the preparation of the system at the initial time the values a of the CG variables appear distributed according to a *known* probability density $P(a, 0)$ that is all the information available about the microstate at the initial time. The Maximum Entropy Principle provides the following initial ensemble density $\rho(z, 0)$ that describes in the best way this situation

$$\rho(z, 0) = \rho_N^0 \frac{P(\hat{A}(z), 0)}{\Omega(\hat{A}(z))} \quad (3.4.2)$$

where the measure $\Omega(a)$ is defined by

$$\Omega(a) \equiv \int dz \rho_N^0 \delta(\hat{A}(z) - a) \quad (3.4.3)$$

By using the form of the initial ensemble (3.4.2) into (3.4.1) we have

$$P(a_0, 0|a_1, t_1) = \frac{1}{\Omega(a_0)} \int dz \rho_N^0 \delta(\hat{A}(z) - a_0)\delta(\hat{A}(\mathcal{U}_{t_1} z) - a_1) \quad (3.4.4)$$

This expression for the transition probability has a nice heuristic interpretation as follows. The numerator $\Omega(a_0)$ is "the number of microstates compatible with a_0" while the denominator in (3.4.4) is "the number of microstates compatible with a_0 that after a time are at a_1". Therefore, the transition probability is just the fraction of microstates compatible with a_0 that after a given time are at a_1.

Microscopic expression for drift and diffusion

We can now obtain the explicit microscopic expressions for $D^{(1)}(a)$ and $D^{(2)}(a)$. Consider the first term $D^{(1)}(a_0)$ and use (3.4.4) in (3.3.20)

$$D^{(1)}(a) = \frac{1}{\Delta t} \int da_1 (a_1 - a) \frac{1}{\Omega(a_1)} \int \delta(\hat{A}(z) - a)\delta(\hat{A}(\mathcal{U}_{\Delta t}z) - a_1)dz \qquad (3.4.5)$$

We perform the integral over the variable a_1 of the Dirac delta function and obtain the desired microscopic expression for the drift $D^{(1)}(a)$.

$$D^{(1)}_\mu(a) = \frac{1}{\Delta t} \int dz \frac{\delta(\hat{A}(z) - a)}{\Omega(a)}(\hat{A}_\mu(\mathcal{U}_{\Delta t}z) - \hat{A}_\mu(z)) = \left\langle \frac{\hat{A}_\mu(\Delta t) - \hat{A}_\mu}{\Delta t} \right\rangle^a \qquad (3.4.6)$$

where we have denoted with $\hat{A}(\Delta t) = \hat{A}(\mathcal{U}_{\Delta t}z)$, and we have introduced the **conditional expectation**

$$\langle \cdots \rangle^a \equiv \int dz \rho^{\text{mic}}_a(z) \cdots \qquad (3.4.7)$$

where the **generalized microcanonical ensemble** is

$$\rho^{\text{mic}}_a(z) = \rho^0_N \frac{\delta(\hat{A}(z) - a)}{\Omega(a)} \qquad (3.4.8)$$

In a similar way, by substitution of (3.4.4) into (3.3.20) we obtain a microscopic expression for the diffusion tensor

$$D^{(2)}_{\mu\nu}(a) = \frac{1}{\Delta t} \left\langle [\hat{A}_\mu(\Delta t) - \hat{A}_\mu][\hat{A}_\nu(\Delta t) - \hat{A}_\nu] \right\rangle^a \qquad (3.4.9)$$

This expression is known as the **Einstein-Helfand formula** for the diffusion coefficient. It basically says that the mean square displacement of the CG variables grows linear in time. As it is obvious from this microscopic expression, the diffusion matrix is a symmetric positive definite matrix, a property that is inherited from its very definition in (3.3.20).

In principle, (3.4.6) and (3.4.9) could be used to compute the drift and diffusion term (at a particular value a) by running MD simulation with initial conditions compatible with $\hat{A}(z) = a$.

The issue of time scales

In summary, *if* the CG variables describe a Markov process *and* the time lapse Δt is short compared to the time-scale of the CG variables (in such a way that the transition probability remains peaked and (3.3.17) holds), then we have explicit microscopic expressions (3.4.6) and (3.4.9) for the drift and diffusion tensor. Note that the left hand side of these expressions are, presumably, independent of Δt, while the right hand sides do depend on Δt. Obviously, for these expressions to make sense, it must happen that (3.4.6) and (3.4.9) should be very weakly dependent on Δt. The microscopic expressions

should inform us that $D^{(1)}(a), D^{(2)}(a)$ are independent on Δt.

The first naive thing to do is to take the limit $\Delta t \to 0$, because Δt is "small", and use a Taylor expansion

$$\hat{A}_\mu(\mathcal{U}_{\Delta t} z) = \hat{A}_\mu(z) + i\mathcal{L}\hat{A}_\mu(z)\Delta t + \mathcal{O}(\Delta t)^2 \qquad (3.4.10)$$

in (3.4.9). However, this is problematic. Indeed

$$D^{(2)}_{\mu\nu}(a) = \Delta t \left\langle i\mathcal{L}\hat{A}_\mu i\mathcal{L}\hat{A}_\nu \right\rangle^a \qquad (3.4.11)$$

Eq. (3.4.11) shows that the diffusion matrix vanish in the strict limit $\Delta t \to 0$! Clearly, Δt has to be "sufficiently large", larger than a typical time τ_{mic}, in order for the mean square displacement to become a linear function of Δt . However, at this "large" time scales the Taylor expansion (3.4.10) is no longer valid! At the same time Δt has to be sufficiently short, shorter than a typical time τ_{mac}, for the transition probability to remain highly peaked. Therefore, we need to have the two conditions

$$\tau_{\mathrm{mic}} \ll \Delta t \ll \tau_{\mathrm{mac}} \qquad (3.4.12)$$

Only if these two conditions are met simultaneously we may have a reasonable Markovian description for the selected relevant variables. Obviously, we need to have a clear separation of time scales $\tau_{\mathrm{mic}} \ll \tau_{\mathrm{mac}}$ in order to have a Markovian description. We also expect that if this separation of time scales exists, then the actual value of Δt, subject to (3.4.12), does not play a significant role. In this case, the drift and diffusion terms are independent of Δt.

Because the Taylor expansion (3.4.10) cannot be used at the microscopically large time scale Δt, we will use instead the following exact expression

$$\hat{A}_\mu(\mathcal{U}_{\Delta t} z) - \hat{A}_\mu(z) = \int_0^{\Delta t} dt \frac{d}{dt} \hat{A}_\mu(\mathcal{U}_t z) = \int_0^{\Delta t} dt\, i\mathcal{L}\hat{A}_\mu(\mathcal{U}_t z) \qquad (3.4.13)$$

As we are assuming that $\Delta t \ll \tau_{\mathrm{mac}}$ and, therefore, the CG variables hardly change in the interval Δt, the quantity (3.4.13) is small. It will make sense, therefore, to make expansions in this quantity, and neglect higher order terms. This will allow us to obtain interesting expressions for the drift (3.4.6) and diffusion (3.4.9) that are explicitly independent of Δt.

By using the identity (3.4.13) in (3.4.6) and (3.4.9), we obtain the equivalent expressions

$$D^{(1)}_\mu(a) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \left\langle i\mathcal{L}\hat{A}_\mu(t) \right\rangle^a$$

$$D^{(2)}_{\mu\nu}(a) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_0^{\Delta t} dt' \left\langle i\mathcal{L}\hat{A}_\mu(t) i\mathcal{L}\hat{A}_\nu(t') \right\rangle^a \qquad (3.4.14)$$

In what follows we will use these expressions to find alternative forms for the drift and diffusion matrix that are particularly convenient.

The Green-Kubo form of the diffusion matrix

First note that the diffusion matrix involves a correlation matrix. As we expect that the variables $i\mathcal{L}\hat{A}$ will decorrelate at long times, we write (3.4.14) in such a way that a correlation that decays to zero appears. This is

$$
\begin{aligned}
D_{\mu\nu}^{(2)}(a) = {} & \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\left\langle\left(i\mathcal{L}\hat{A}_\mu(t)-\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\right)\left(i\mathcal{L}\hat{A}_\nu(t')-\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a\right)\right\rangle^a \\
& + \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\left\langle i\mathcal{L}\hat{A}_\mu(t)\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\right\rangle^a \\
& + \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\left\langle\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a i\mathcal{L}\hat{A}_\nu(t')\right\rangle^a \\
& - \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\left\langle\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a\right\rangle^a
\end{aligned}
\tag{3.4.15}
$$

By using the definition of the drift in (3.4.14) and performing trivial time integrals this becomes

$$
\begin{aligned}
D_{\mu\nu}^{(2)}(a) = {} & \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\left\langle\left(i\mathcal{L}\hat{A}_\mu(t)-\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\right)\left(i\mathcal{L}\hat{A}_\nu(t')-\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a\right)\right\rangle^a \\
& + \Delta t D_\mu^{(1)}(a)\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a + \Delta t\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a D_\nu^{(1)}(a) - \Delta t\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a
\end{aligned}
\tag{3.4.16}
$$

We will consistently neglect the terms in the second line as they are proportional to Δt (note that terms of this order were already neglected in (3.3.20)). Therefore, we have the following form for the diffusion matrix

$$
D_{\mu\nu}^{(2)}(a) = \frac{1}{\Delta t}\int_0^{\Delta t}dt\int_0^{\Delta t}dt'\,K_{\mu\nu}(t,t',a)
\tag{3.4.17}
$$

were we have introduced the correlation

$$
K_{\mu\nu}(t,t',a) \equiv \left\langle\left(i\mathcal{L}\hat{A}_\mu(t)-\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a\right)\left(i\mathcal{L}\hat{A}_\nu(t')-\left\langle i\mathcal{L}\hat{A}_\nu\right\rangle^a\right)\right\rangle^a
\tag{3.4.18}
$$

Now, we divide the region of integration over t, t' in the diffusion matrix in (3.4.17) in two regions with $t' > t$ and $t' < t$. Therefore, we have

$$
D_{\mu\nu}^{(2)}(a) = \frac{1}{\Delta t}\int_0^{\Delta t}dt'\int_0^{t'}dt\,K_{\mu\nu}(t,t',a) + \frac{1}{\Delta t}\int_0^{\Delta t}dt'\int_{t'}^{\Delta t}dt\,K_{\mu\nu}(t,t',a)
\tag{3.4.19}
$$

By interchanging the order of the integrations in the second integral, it can be expressed as

$$\frac{1}{\Delta t}\int_0^{\Delta t} dt' \int_{t'}^{\Delta t} dt K_{\mu\nu}(t,t',a) = \frac{1}{\Delta t}\int_0^{\Delta t} dt \int_0^t dt' K_{\mu\nu}(t,t',a)$$

$$= \frac{1}{\Delta t}\int_0^{\Delta t} dt' \int_0^{t'} dt K_{\mu\nu}(t',t,a) \qquad (3.4.20)$$

where in the last equality we have interchanged the variables t and t'. Therefore, we conclude that the diffusion matrix can be written as

$$D_{\mu\nu}^{(2)}(a) = \frac{1}{\Delta t}\int_0^{\Delta t} dt' \int_0^{t'} dt [K_{\mu\nu}(t,t',a) + K_{\mu\nu}(t',t,a)] \qquad (3.4.21)$$

The correlation matrix (3.4.18) is not stationary and, therefore, it does not satisfy

$$K_{\mu\nu}(t,t',a) = K_{\mu\nu}(t-t',a) \qquad (3.4.22)$$

because the generalized microcanonical ensemble $\rho_a^{\text{mic}}(z)$ introduced in (3.4.8) and appearing in its definition is not invariant under the dynamics. However, within the time scale Δt the CG variables are assumed to be almost constant and, therefore, the generalized microcanonical ensemble is approximately invariant. Perhaps in more precise terms, we assume that averages computed with $\rho_a^{\text{mic}}(z)$ and with $\rho_a^{\text{mic}}(\mathcal{U}_{-t}z)$ are essentially the same. Therefore, we will assume that (3.4.22) is very well satisfied. The stationarity condition implies that

$$K_{\mu\nu}(t,a) = K_{\nu\mu}(-t,a) \qquad (3.4.23)$$

Under the stationarity approximation (3.4.22) the diffusion matrix (3.4.21)

$$D_{\mu\nu}^{(2)}(a) = \frac{1}{\Delta t}\int_0^{\Delta t} dt' \int_0^{t'} dt [K_{\mu\nu}(t-t',a) + K_{\nu\mu}(t-t',a)] \qquad (3.4.24)$$

All the problem now is to eliminate one of the time integrals by exploiting the fact that the integrand depends only on the time difference. Perform the change of variable

$t \to \tau = t - t'$

$$D^{(2)}_{\mu\nu}(a) = \frac{1}{\Delta t} \int_0^{\Delta t} dt' \int_{-t'}^0 d\tau [K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)]$$

$$= \frac{1}{\Delta t} \int_0^{\Delta t} dt' \int_0^{t'} d\tau [K_{\mu\nu}(-\tau, a) + K_{\nu\mu}(-\tau, a)]$$

$$= \frac{1}{\Delta t} \int_0^{\Delta t} d\tau \int_\tau^{\Delta t} dt' [K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)]$$

$$= \frac{1}{\Delta t} \int_0^{\Delta t} d\tau (\Delta t - \tau)[K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)]$$

$$= \int_0^{\Delta t} d\tau [K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)] - \frac{1}{\Delta t} \int_0^{\Delta t} d\tau \tau [K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)]$$

$$(3.4.25)$$

Note that the second integral is related to the correlation time of the correlation of $K(\tau, a)$. This correlation time is of the order of τ_{mic}. Therefore, the second integral is of the order $\tau_{\text{mic}}/\Delta t$ which, on account of (3.4.12) will be neglected. The diffusion matrix then becomes

$$D^{(2)}_{\mu\nu}(a) = \int_0^{\Delta t} d\tau [K_{\mu\nu}(\tau, a) + K_{\nu\mu}(\tau, a)] \qquad (3.4.26)$$

It proves convenient to introduce the **dissipative matrix** as

$$D_{\mu\nu}(a) = \int_0^{\Delta t} d\tau \left\langle \left[i\mathcal{L}\hat{A}_\mu(t) - \left\langle i\mathcal{L}\hat{A}_\mu \right\rangle^a \right] \left[i\mathcal{L}\hat{A}_\nu - \left\langle i\mathcal{L}\hat{A}_\nu \right\rangle^a \right]^T \right\rangle^a \qquad (3.4.27)$$

The diffusion matrix (3.4.26) is just twice the symmetric part of the dissipative matrix

$$D^{(2)}_{\mu\nu}(a) = D_{\mu\nu}(a) + D_{\nu\mu}(a) \qquad (3.4.28)$$

While the diffusion matrix (3.4.28) is symmetric by construction, the dissipative matrix (3.4.27) needs not to be symmetric. Eq. (3.4.27) is the celebrated **Green-Kubo** expression of the dissipative matrix.

The drift term

We now consider an expression for the drift term (3.4.14) that reveals an interesting physical structure

$$D^{(1)}_\mu(a) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \left\langle i\mathcal{L}\hat{A}_\mu(t) \right\rangle^a \qquad (3.4.29)$$

We know that the naive Taylor expansion (3.4.10) would lead to

$$D^{(1)}_\mu(a) = \left\langle i\mathcal{L}\hat{A}_\mu \right\rangle^a \qquad (3.4.30)$$

but this cannot be correct because we cannot use the Taylor expansion for the times in the range $(0, \Delta t)$. We aim at obtaining the corrections to (3.4.30). To this end, consider (3.4.29) in the form

$$D_\mu^{(1)}(a) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \int dz \rho_N^0 \frac{\delta(\hat{A}(\mathcal{U}_{-t}z) - a)}{\Omega(a)} i\mathcal{L}\hat{A}_\mu(z) \qquad (3.4.31)$$

where we have transferred the time dependence to the microcanonical ensemble through the usual change of variables $z \to \mathcal{U}_{-t}z$. It is obvious that now we have to take the evolution of the microcanonical ensemble into account, otherwise we are back to the naive (3.4.30). To this end, consider the identity, similar to (3.4.13)

$$\hat{A}_\mu(\mathcal{U}_{-t}z) = \hat{A}_\mu(z) - \int_0^t dt' i\mathcal{L}\hat{A}_\mu(\mathcal{U}_{-t'}z) \qquad (3.4.32)$$

This integral term is assumed to be small for times t which are smaller than Δt because the CG variables hardly change in this range. Therefore, we can now formally expand the Dirac delta function in (3.4.31) around $\hat{A}(z) - a$,

$$\delta(\hat{A}(\mathcal{U}_{-t}z) - a) = \delta\left(\hat{A}(z) - a + \int_0^t dt' i\mathcal{L}\hat{A}(\mathcal{U}_{-t'}z)\right)$$

$$= \delta\left(\hat{A}(z) - a\right) + \frac{\partial}{\partial a_\nu}\delta\left(\hat{A}(z) - a\right) \int_0^t dt' i\mathcal{L}\hat{A}_\nu(\mathcal{U}_{-t'}z) + \cdots \quad (3.4.33)$$

By using (3.4.33) into (3.4.31) we obtain

$$D_\mu^{(1)}(a) = \left\langle i\mathcal{L}\hat{A}_\mu \right\rangle^a + \frac{1}{\Omega(a)} \frac{\partial}{\partial a_\nu} \Omega(a) \frac{1}{\Delta t} \int_0^{\Delta t} dt' \int_0^t dt \left\langle i\mathcal{L}\hat{A}_\mu i\mathcal{L}\hat{A}_\nu(-t) \right\rangle^a + \cdots$$

$$(3.4.34)$$

The first term is clearly independent of Δt. We have to asses how is the scaling with Δt of the rest of the terms. For example, the integral in the second term has the same structure of the diffusion matrix. We follow exactly the same steps (substracting a negligible term of order Δt as in (3.4.16) in order to have a properly decaying correlation, assume stationarity, reorderings of integrals, and the definition of the dissipative matrix (3.4.27)) to arrive at the result

$$D_\mu^{(1)}(a) = \left\langle i\mathcal{L}\hat{A}_\mu \right\rangle^a + \frac{1}{\Omega(a)} \frac{\partial}{\partial a_\nu} \Omega(a) D_{\mu\nu}(a) + \cdots \qquad (3.4.35)$$

The analysis of the third and higher order terms in (3.4.35) becomes readily very complicated. We will assume, though, that these terms are of higher order in Δt and that can be neglected. This approximation is not rigorous and one has to judge the resulting expression a posteriori. Note that, as compared with the naive result (3.4.30), we have an additional term in the drift.

3.4.1 The Fokker Planck Equation

By collecting the results (3.4.35) for the drift and (3.4.28) for the diffusion, the FPE (3.3.7) becomes

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[\left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a + \frac{1}{\Omega(a)}\frac{\partial}{\partial a_\nu}\Omega(a)D_{\mu\nu}(a)\right]P(a,t) + \frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a_\nu}D_{\mu\nu}^S(a)P(a,t)$$

(3.4.36)

The FPE is usually written in the equivalent form

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[V_\mu(a) + D_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu}(a)\right]P(a,t) + \frac{\partial}{\partial a_\mu}D_{\mu\nu}^S(a)\frac{\partial}{\partial a_\nu}P(a,t) \quad (3.4.37)$$

where we have introduced the **reversible drift**

$$V_\mu(a) = \left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a \tag{3.4.38}$$

and the **generalized Boltzmann entropy** or **bare entropy** according to

$$S_B(a) = \ln\frac{\Omega(a)}{\Omega_0} \tag{3.4.39}$$

This is a generalization of Boltzmann entropy

$$S_B(E) \equiv \ln\Omega(E) = \ln\int dz\rho_N^0\delta(\hat{H}(z) - E) \tag{3.4.40}$$

when we have, in addition to the energy, other CG variables. Recall that in this book we measure entropies in units of the Boltzmann constant k_B and, therefore, they are dimensionless quantities. The factor Ω_0 has the same physical dimensions as $\Omega(a)$ and renders the argument of the logarithm dimensionless as it should. The actual value of Ω_0 is irrelevant, because what enters the FPE (3.4.37) are the derivatives of the generalized Boltzmann entropy.

3.4.2 The Stochastic Differential Equation

The SDE governs the instantaneous value of the stochastic variable $a(t)$ that represents the stochastic values of the CG variables. The SDE that is mathematically equivalent to the FPE (3.4.37) is given by

$$da_\mu(t) = V_\mu(a(t))dt + D_{\mu\nu}(a(t))\frac{\partial S_B}{\partial a_\nu}(a(t))dt + \frac{\partial D_{\mu\nu}}{\partial a_\nu}(a(t))dt + d\tilde{a}_\mu(t) \tag{3.4.41}$$

and should be interpreted in Ito sense [**?**]. Here, the random term, or noise, $d\tilde{a}_\mu(t)$ is a linear combination of Wiener processes $dW_i(t)$, this is

$$d\tilde{a}_\mu(t) = B_{\mu i}(a(t))dW_i(t) \tag{3.4.42}$$

where the matrix B (not necessarily a square matrix) satisfies the **Fluctuation-Dissipation theorem**

$$\sum_i B_{\mu i}(a) B_{\nu i}(a) = 2D^S_{\mu\nu}(a) \tag{3.4.43}$$

The term involving the divergence of the diffusion matrix in (3.4.41) is a consequence of the Ito stochastic interpretation and the fact that the diffusion matrix appears *in between* the two derivatives in the FPE (3.4.37). By remembering that the independent increments of the Wiener process have a variance given by dt, the Fluctuation-Dissipation theorem is written in a mnemotechnically simpler way

$$d\tilde{a}_\mu d\tilde{a}_\nu = 2D^S_{\mu\nu}(a)dt \tag{3.4.44}$$

Note that there is no profound "physics" in the Fluctuation-Dissipation theorem, which is nothing else than the rule for obtaining the SDE from the FPE and establishes that the amplitude of the noise is given by the dissipative matrix. The profound physics is buried in the fact that the stochastic process of the CG variables is approximately Markovian.

3.5   Coarse-graining à la Zwanzig

In the previous section, we have presented the theory of coarse-graining as was given by Green. In principle, and quite often in practice, this is all what we would need to know about the theory of coarse-graining. Of course, the whole construction is based on the assumption that the stochastic process of CG variables has the Markov property. We have seen that this assumption is equivalent to *postulate* that the one-time and transition probabilities obey the Fokker-Planck equation. If this is true, then we have explicit molecular expressions for the dynamics of the CG variables. Of course, in the process of obtaining these molecular expressions we have followed a number of non-rigorous approximations that may leave us with a somewhat uncomfortable feeling. In the present section, in order to get some more insight into the problem we will not postulate the Fokker-Planck equation but rather will derive it directly from the microscopic Hamilton's equations by following the route of projection operators. Of course, we will need to do approximations in the process that, essentially convey the same physics as those in the previous section, but their mathematical implementation is somewhat different. We hope that in this way, we will have a complementary view to the problem.

3.5.1   The projector

As we take the probabilistic CG variables (4.3.2) as the central objects in the probabilistic flavour, it makes sense to construct a projection operator that transforms any phase function into a linear combination of the probabilistic CG variables. A natural one is Zwanzig's projector

$$\mathcal{P}^\dagger \hat{F}(z) = \int da \left\langle \hat{F} \right\rangle^a \hat{\Psi}_a(z) \tag{3.5.1}$$

where the conditional expectation is defined in (3.4.7). If we perform in (3.5.1) the integral over the Dirac delta function $\hat{\Psi}_a(z)$ we have

$$\mathcal{P}^\dagger \hat{F}(z) = \langle F \rangle^{\hat{A}(z)} \tag{3.5.2}$$

The action of Zwanzig's projector is to take the average of the phase function $\hat{F}(z')$ on the submanifold $\hat{A}(z') = a$ of constant CG variables a, and then evaluate this average at $a = \hat{A}(z)$. As we will see in Sec **??**, Zwanzig's projector is just a particular case of the Kawasaki-Gunton projector when the CG variables are (4.3.2). Therefore, the projector (3.5.2) satisfies all the properties described in Sec. **??**.

3.5.2   The exact generalized FPE and Langevin equations

We consider the time evolution $\Psi_a(\mathcal{U}_t z)$ of the probabilistic CG variables, where $\mathcal{U}_t z$ is the solution of Hamilton's equations with initial condition z. By taking the time derivative we have

$$\partial_t \hat{\Psi}_a(\mathcal{U}_t z) = \exp\{i\mathcal{L}t\} i\mathcal{L} \hat{\Psi}_a(z) \tag{3.5.3}$$

Now, consider the following decomposition of the evolution operator

$$\exp\{i\mathcal{L}t\} = \exp\{i\mathcal{L}t\}\mathcal{P}^\dagger + \int_0^t dt' \exp\{i\mathcal{L}t'\}\mathcal{P}^\dagger i\mathcal{L} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger + \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger \tag{3.5.4}$$

where $\mathcal{Q}^\dagger = 1 - \mathcal{P}^\dagger$ is the complementary operator. This identity can be easily proved by differentiation with respect to t. If two operators that depend on time have the same derivative and they coincide at the initial time, then they are the same operator.

By substitution of (3.5.4) into (3.5.3) and using the form (3.5.2) for Zwanzig projector $\mathcal{P}^\dagger$, we arrive at

$$\partial_t \hat{\Psi}_a(z) = \langle i\mathcal{L}\psi_a \rangle^{\hat{A}(\mathcal{U}_t(z))} + \int_0^t dt' \left\langle i\mathcal{L} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L} \hat{\Psi}_a \right\rangle^{\hat{A}(\mathcal{U}_u(z))}$$
$$+ \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger i\mathcal{L} \hat{\Psi}_a(z) \tag{3.5.5}$$

The term inside the time integral can be written as

$$\left\langle i\mathcal{L} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L} \hat{\Psi}_a \right\rangle^{\hat{A}(u)} = \int da' \Psi_{a'}(\mathcal{U}_u \hat{A}(z)) \left\langle i\mathcal{L} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L} \hat{\Psi}_a \right\rangle^{a'} \tag{3.5.6}$$

The conditional expectation inside this integral (3.5.6) can be written as

$$\left\langle i\mathcal{L}\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a\right\rangle^{a'} = \frac{1}{\Omega(a')}\int dz\rho_N^0\Psi_{a'}(z)i\mathcal{L}\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a(z)$$

$$= -\frac{1}{\Omega(a')}\int dz\rho_N^0(i\mathcal{L}\Psi_{a'}(z))\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a(z)$$

$$(3.5.7)$$

where in the last equality we have used the hermitian nature of the Liouville operator. Next, we note that the chain rule implies

$$i\mathcal{L}\Psi_a(z) = -\frac{\partial}{\partial a_\mu}\Psi_a i\mathcal{L}\hat{A}_\mu(z) \qquad (3.5.8)$$

therefore, Eq. (3.5.7) becomes

$$\left\langle i\mathcal{L}\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a\right\rangle^{a'}$$

$$= -\frac{1}{\Omega(a')}\frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a_\nu'}\int dz\rho_N^0\Psi_{a'}(z)i\mathcal{L}\hat{A}_\nu(z)\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger\Psi_a i\mathcal{L}\hat{A}_\mu(z)$$

$$= -\frac{1}{\Omega(a')}\frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a_\nu'}\int dz\rho_N^0\Psi_{a'}(z)i\mathcal{L}\hat{A}_\nu(z)\mathcal{Q}^\dagger\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger\Psi_a i\mathcal{L}\hat{A}_\mu(z)$$

$$= -\frac{1}{\Omega(a')}\frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a_\nu'}\int dz\rho_N^0\Psi_{a'}(z)(\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\nu(z))\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger\Psi_a i\mathcal{L}\hat{A}_\mu(z)$$

$$= -\frac{1}{\Omega(a')}\frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a_\nu'}\Omega(a')\left\langle(\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\nu)\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger\Psi_a i\mathcal{L}\hat{A}_\mu\right\rangle^{a'} \qquad (3.5.9)$$

By using (3.5.6)-(3.5.9) into (3.5.5) we obtain (after a partial integration with respect to a')

$$\partial_t\hat{\Psi}_a(\mathcal{U}_t z) = -\frac{\partial}{\partial a_\mu}V_\mu(a)\hat{\Psi}_a(\mathcal{U}_t z)$$

$$+ \int_0^t du\frac{\partial}{\partial a_\mu}\int da'\Omega(a')D_{\mu\nu}(a,a';t-u)\frac{\partial}{\partial a_\nu'}\frac{\Psi_{a'}(\mathcal{U}_u z)}{\Omega(a')} + F_a^\dagger(t) \qquad (3.5.10)$$

In this equation we have introduced the **reversible drift** (3.4.38). The **memory kernel** of probabilistic CG variables is defined as

$$D_{\mu\nu}(a,a';t-t') \equiv \left\langle(i\mathcal{L}\hat{A}_\nu - V_\nu(a'))\exp\left\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\right\}\Psi_a(i\mathcal{L}\hat{A}_\mu - V_\mu(a))\right\rangle^{a'} \qquad (3.5.11)$$

Finally, the **projected current** is defined as

$$F_a^\dagger(t)(z) \equiv \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a(z) \qquad (3.5.12)$$

Eq (3.5.10) is the basis for obtaining the generalized FPE for the probability $P(a,t)$

and the transition probability $P(a_0, 0|a, t)$, and the generalized Langevin equation for the CG variables.

The generalized Fokker-Planck Equation

The probability distribution of the CG variables is given by (1.7.5) that we quote again

$$P(a, t) = \int dz \rho_0(z) \hat{\Psi}_a(\mathcal{U}_t z) \tag{3.5.13}$$

If we take the time derivative and insert (3.5.10) into (3.5.13) we obtain the following exact **generalized Fokker-Planck equation** for the probability distribution

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} V_\mu(a) P(a, t) + \frac{\partial}{\partial a_\mu} \int da' \int_0^t dt' D_{\mu\nu}(a, a'; t - t') \Omega(a') \frac{\partial}{\partial a'_\nu} \frac{P(a', t')}{\Omega(a')}$$
$$\tag{3.5.14}$$

Because the memory kernel (3.5.11) does not depend on the probability $P(a, t)$, this is a *closed* and *linear* partial integro-differential equation for $P(a, t)$.

Note that the last term in (3.5.10) does not contribute upon averaging, because

$$\text{Tr}[\overline{\rho}_0 F_a^\dagger(t)] = \text{Tr}[\overline{\rho}_0 \mathcal{Q}^\dagger \exp\{i\mathcal{L}\mathcal{Q}^\dagger t\} i\mathcal{L}\hat{\Psi}_a] = 0 \tag{3.5.15}$$

because of the fact that $\mathcal{Q}\overline{\rho}_0 = 0$ for relevant ensembles.

Evolution of the transition probability

The transition probability is defined in (3.4.4) that we quote again

$$P(a_0, 0|a_1, t_1) = \frac{1}{\Omega(a_0)} \int dz \rho_N^0 \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\mathcal{U}_{t_1} z) - a_1) \tag{3.5.16}$$

If we take the time derivative of (3.5.16) and use (3.5.10) we obtain

$$\partial_t p(a_0, 0|a, t) = -\frac{\partial}{\partial a_\mu} V_\mu(a) p(a_0, 0|a, t)$$
$$+ \int_s^t du \frac{\partial}{\partial a_\mu} \int da' \Omega(a') D_{\mu\nu}(a, a'; t - u) \frac{\partial}{\partial a'_\nu} \frac{p(a_0, 0|a', u)}{\Omega(a')} \tag{3.5.17}$$

where we have used

$$\text{Tr}[\overline{\rho}_0 \Psi_{a_0} F_a^\dagger(t, 0)] = \text{Tr}[\overline{\rho}_0 \Psi_{a_0} \mathcal{Q}^\dagger \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\} i\mathcal{L}\hat{\Psi}_a] = 0 \tag{3.5.18}$$

Therefore, the transition probability obeys the same generalized Fokker-Planck equation as the one-time probability.

The Generalized Langevin Equation

We may also obtain from (3.5.10) a dynamic equation for the CG variables themselves. Indeed, if we multiply (3.5.10) with a_μ and then integrate over a, we obtain **the generalized Langevin equation (GLE)**

$$\partial_t \hat{A}_\mu(\mathcal{U}_t z) = V_\mu(\hat{A}(\mathcal{U}_t z))$$
$$+ \int_0^t du \left[D_{\mu\nu}(\hat{A}(\mathcal{U}_u z); t - u) \frac{\partial S}{\partial a_\nu}(\hat{A}(\mathcal{U}_u z)) + \frac{\partial}{\partial a_\nu} D_{\mu\nu}(\hat{A}(\mathcal{U}_u z); t - u) \right] + F_\mu^\dagger(t)$$
$$(3.5.19)$$

where we have defined **the memory kernel** of CG variables as

$$D_{\mu\nu}(a; t) = \int da' D_{\mu\nu}(a, a', t) = \left\langle F_\nu^\dagger(0) F_\mu^\dagger(t) \right\rangle^a$$
$$F_\mu^\dagger(t) = \exp\{i\mathcal{L}\mathcal{Q}^\dagger t\} \mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu \qquad (3.5.20)$$

The phase function $F_\mu^\dagger(t)$ is usually termed the **random force**, although it is evidently a time dependent phase function without any stochastic component. Observe that the random force is a total time derivative

$$F_\mu^\dagger(t) = \frac{d}{dt} \exp\{i\mathcal{L}\mathcal{Q}^\dagger t\} \hat{A}_\mu \qquad (3.5.21)$$

3.5.3 The Fokker Planck Equation

The exact generalized FPE (3.5.14) is not very useful because it is an integro-differential equation but, most blatantly, the memory kernel is a formal object due to the presence of the projected dynamics. It is necessary to take some approximations to render this equation into a usable form. The idea is to use again the separation of time scales and assume that the CG variables change in a time scale τ_{mac} which is much larger than the time scale τ_{mic} in which the memory kernel decays. Note that the memory kernel (3.5.11) involves, essentially, the correlation of the time derivatives $i\mathcal{L}\hat{A}$ of the CG variables, we are assuming that the CG variables evolve in two distinct time scales. This allows to take the approximation

$$\int_0^t dt' D_{\mu\nu}(a, a'; t - t') \frac{\partial}{\partial a'_\nu} \frac{P(a', t')}{\Omega(a')} \approx \left[\int_0^{\Delta t} D_{\mu\nu}(a, a'; \tau) d\tau \right] \frac{\partial}{\partial a'_\nu} \frac{P(a', t)}{\Omega(a')} \qquad (3.5.22)$$

Here, Δt is any time sufficiently larger than the decay time τ_{mic} of the memory kernel. As we will see, due to the plateau problem, we cannot extend Δt to infinity, but should remain smaller that the typical evolution time of the CG variables τ_{mac}. Note that this approximation is equivalent to approximate

$$D_{\mu\nu}(a, a'; t - t') = \delta^+(t - t') \int_0^{\Delta t} D_{\mu\nu}(a, a'; \tau) d\tau \qquad (3.5.23)$$

where $\delta^+(t)$ is a Dirac delta function satisfying

$$\int_0^\infty dt \delta^+(t) = 1 \qquad\qquad (3.5.24)$$

instead of the usual definition

$$\int_{-\infty}^\infty dt \delta(t) = 1 \qquad\qquad (3.5.25)$$

Under the assumption of separation of time scales, we may approximate the memory kernel (3.5.11), that we quote again

$$D_{\mu\nu}(a, a'; t - t') = \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a'))) \exp\left\{ \mathcal{Q}^\dagger i\mathcal{L}(t - t') \right\} \Psi_a (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^{a'}$$
$$\approx \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a'))) \Psi_a \exp\left\{ \mathcal{Q}^\dagger i\mathcal{L}(t - t') \right\} (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^{a'}$$
$$(3.5.26)$$

where we have transferred $\hat{\Psi}_a$ to the left of the evolution operator under the assumption that in the time scale of decay of the kernel the CG variables hardly evolve. By noting that the conditional expectation involves $\Psi_{a'}(z)$ and that $\hat{\Psi}_a(z)\Psi_{a'}(z) = \delta(a - a')\hat{\Psi}_a(z)$ we have that in this approximation the memory kernel becomes diagonal in the space of CG variables

$$D_{\mu\nu}(a, a'; t) \approx \delta(a - a') \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a))) \exp\left\{ \mathcal{Q}^\dagger i\mathcal{L}t \right\} (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^a$$
$$= \delta(a - a') D_{\mu\nu}(a, t) \qquad\qquad (3.5.27)$$

where we have used the definition of the memory kernel (3.5.20).

By inserting this approximate expression into the generalized Fokker-Planck equation (3.5.14) we obtain the following FPE for the probability distribution of CG variables

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} V_\mu(a) P(a, t) + \frac{\partial}{\partial a_\mu} D_{\mu\nu}(a) \Omega(a) \frac{\partial}{\partial a_\nu} \frac{P(a, t)}{\Omega(a)} \qquad (3.5.28)$$

where the **dissipative matrix** is given by

$$D_{\mu\nu}(a) = \int_0^{\Delta t} d\tau \left\langle \hat{F}_\nu^\dagger \hat{F}_\mu^\dagger(t) \right\rangle^a$$
$$= \int_0^{\Delta t} d\tau \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a))) \exp\left\{ \mathcal{Q}^\dagger i\mathcal{L}\tau \right\} (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^a \qquad (3.5.29)$$

We encounter again the plateau problem in this expression. Because the random force $\hat{F}_\mu^\dagger(t)$ is a total time derivative as seen in (3.5.21), we cannot take the limit $\Delta t \to \infty$. Therefore, the value of Δt is subject to the restrictions (3.4.12).

By introducing the bare entropy (3.4.39), it is easily seen that the FPE (3.5.28) is identical to the FPE (3.4.36) obtained by Green. In this way, the FPE postulated

by Green is here derived from the microscopic Hamiltonian dynamics. Of course, the assumptions needed to reduce the exact generalized FPE into an ordinary FPE are essentially the ones that already ensure the validity of Green's FPE.

The projected dynamics is similar to the real dynamics for $t < \Delta t$

The attentive reader will note that there is a slight difference between the dissipative matrix (3.5.29) and the one obtained in Green's approach (3.4.27). The former involves the projected dynamics $\exp\left\{\mathcal{Q}^{\dagger}i\mathcal{L}\tau\right\}$ while the latter involves the real dynamics $\exp\left\{i\mathcal{L}\tau\right\}$. We show now that within the approximations considered both expressions are equivalent.

The GLE (3.5.19) can be cast in the following form

$$F_{\mu}(t) = F_{\mu}^{\dagger}(t) + \int_{0}^{t} du \left[D_{\mu\nu}(\hat{A}(\mathcal{U}_{u}z); t-u)\frac{\partial S}{\partial a_{\nu}}(\hat{A}(\mathcal{U}_{u}z)) + \frac{\partial}{\partial a_{\nu}}D_{\mu\nu}(\hat{A}(\mathcal{U}_{u}z); t-u) \right] \tag{3.5.30}$$

where we have introduced the fluctuation of the real force as

$$\begin{aligned}
F_{\mu}(t) &= \frac{d}{dt}\hat{A}_{\mu}(\mathcal{U}_{t}z) - V_{\mu}(\hat{A}(\mathcal{U}_{t}z)) \\
&= \exp\{i\mathcal{L}t\}i\mathcal{L}\hat{A}_{\mu}(z) - \mathcal{P}^{\dagger}[\exp\{i\mathcal{L}t\}i\mathcal{L}\hat{A}_{\mu}](z) \\
&= \exp\{i\mathcal{L}t\}\mathcal{Q}^{\dagger}i\mathcal{L}\hat{A}_{\mu}(z)
\end{aligned} \tag{3.5.31}$$

To be compared with the random force (3.5.20) that we quote again

$$F_{\mu}^{\dagger}(t) = \exp\{\mathcal{Q}^{\dagger}i\mathcal{L}t\}\mathcal{Q}^{\dagger}i\mathcal{L}\hat{A}_{\mu}(z) \tag{3.5.32}$$

We see that the random force $F_{\mu}^{\dagger}(t)$ evolves with $\exp\{\mathcal{Q}^{\dagger}i\mathcal{L}t\}$ while the real force $F_{\mu}(t)$ evolves with $\exp\{i\mathcal{L}t\}$. The difference between these two dynamics is in the integral term in (3.5.30). This is an integral of a function of the CG variables that hardly vary during the time scale Δt. Therefore, it is expected that the above integral is, in macroscopic terms, a very small quantity of order Δt.

3.5.4   The Stochastic Differential Equation

Note that the Markov property (3.5.23) implies the following approximation for the memory kernel (3.5.20) of the GLE (3.5.19)

$$D_{\mu\nu}(a, t-t') = \delta^{+}(t-t')\int_{0}^{\Delta t} D_{\mu\nu}(a, \tau)d\tau = \delta^{+}(t-t')D_{\mu\nu}(a) \tag{3.5.33}$$

By inserting the expression of the Markov approximation (3.5.33) into the GLE (3.5.19) we obtain

$$\frac{d}{dt}\hat{A}_{\mu}(\mathcal{U}_{t}z) = V_{\mu}(\hat{A}(\mathcal{U}_{t}z)) + D_{\mu\nu}(\hat{A}(\mathcal{U}_{t}z))\frac{\partial S}{\partial a_{\nu}}(\hat{A}(\mathcal{U}_{t}z)) + \frac{\partial}{\partial a_{\nu}}D_{\mu\nu}(\hat{A}(\mathcal{U}_{t}z)) + F_{\mu}^{\dagger}(t) \tag{3.5.34}$$

This is an approximate equation for the evolution of the CG variables. As all terms, except the random force $F_\mu^\dagger(t)$ are functions of $\hat{A}(\mathcal{U}_t z)$ this can be interpreted as a closed equation for the values $a(t)$ of the CG variables, this is

$$\frac{d}{dt}\tilde{a}(t) = V_\mu(\tilde{a}(t)) + D_{\mu\nu}(\tilde{a}(t))\frac{\partial S}{\partial a_\nu}(\tilde{a}(t)) + \frac{\partial}{\partial a_\nu}D_{\mu\nu}(\tilde{a}(t)) + F_\mu^\dagger(t) \qquad (3.5.35)$$

This is a **Langevin equation** which is an ordinary differential equation with a random term. However, we should be very cautious in interpreting this equation as an ordinary differential equation. Due to the approximations made, we cannot take it seriously as a differential equation because for sure such an equation cannot capture the features of the CG variables that happen below the time scale Δt. Therefore, the very concept of "differentiability" is already put in question by the Markovian assumption made.

Note that (3.5.33) can be understood, from (3.5.20) in the form

$$\left\langle F_\nu^\dagger(0)F_\mu^\dagger(t)\right\rangle^a = \delta^+(t)D_{\mu\nu}(a) \qquad (3.5.36)$$

where this is valid for $t > 0$. For $t = -|t| < 0$, we have, from the approximate stationarity of the generalized microcanonical ensemble

$$\left\langle F_\nu^\dagger(0)F_\mu^\dagger(-|t|)\right\rangle^a = \left\langle F_\nu^\dagger(|t|)F_\mu^\dagger(0)\right\rangle^a = \delta^+(|t|)D_{\nu\mu}(a) \qquad (3.5.37)$$

We may combine (3.5.36) and (3.5.37) in a form valid for all times as

$$\left\langle F_\nu^\dagger(0)F_\mu^\dagger(t)\right\rangle^a = \delta(t)[D_{\mu\nu}(a) + D_{\nu\mu}(a)] \qquad (3.5.38)$$

The interest of this "mechanistically" derived Langevin equation arises when we compare it with the SDE (3.4.41) that is mathematically equivalent to the FPE (3.4.37). By dividing (3.4.41) with dt we obtain an equation resembling an ordinary differential equation

$$\frac{da_\mu}{dt}(t) = V_\mu(a(t)) + D_{\mu\nu}(a(t))\frac{\partial S_B}{\partial a_\nu}(a(t)) + \frac{\partial D_{\mu\nu}}{\partial a_\nu}(a(t)) + \tilde{F}_\mu(t) \qquad (3.5.39)$$

where we have introduced the white noise forces

$$\tilde{F}_\mu(t) \equiv \frac{d\tilde{a}_\mu}{dt}(t) \qquad (3.5.40)$$

The white noise is a mathematically ill-defined object because the Wiener process is not differentiable. Nevertheless, intuitively is the prototype of an "extremely irregular function of time". The random forces have zero mean and have a variance given, from (3.4.44), by

$$\langle\tilde{F}_\mu(t)\tilde{F}_\nu(t')\rangle = \delta(t - t')[D_{\mu\nu}(a_t) + D_{\nu\mu}(a_t)] \qquad (3.5.41)$$

where the average here is over realizations of the stochastic process. Note the formal similitude of this result for the stochastic process $\tilde{F}_\mu(t)$ and the behaviour of the random force $\hat{F}_\mu^\dagger(t)$. We see therefore, that what the Markov approximation does in practice is to model the mechanistic $\hat{F}_\mu^\dagger(t)$ with a white noise.

In principle, the theory gives us is a microscopic expression for the dissipative matrix given by (3.5.29). However, in order to solve numerically the SDE (3.4.41), we have to obtain not $D_{\mu\nu}$ but rather "its square root" $B_{\mu i}$ satisfying (3.4.43). Note that there may be many different random terms $d\tilde{a}_\mu$ that are equivalent, because there may be many different matrices $B_{\mu\nu}$ whose "square" in the sense of (3.4.43) gives the same dissipative matrix $D_{\mu\nu}$. We see now the interest of the "mechanistic" Langevin equation in that it suggests that the stochastic forces $\tilde{F}_\mu(t)$ are related to the projected currents, this is

$$\tilde{F}_\mu(t) \sim \exp\{i\mathcal{L}t\}(i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \tag{3.5.42}$$

where we identify a phase function with a stochastic white noise. This identification is certainly not rigorous, but shows that the Fluctuation-Dissipation theorem and the Green-Kubo expression are closely related. In fact, the Markovian approximation *models* the projected currents as *white noise*, this is functions that are delta-correlated in time, which is the same as saying that the correlation of the projected currents decays in a very short time scale as shown in (3.5.36).

The heuristic identification (3.5.42) of the random forces with the projected currents is very helpful when one has to construct the matrix $B_{\mu i}$ needed in the formulation of the SDE for the CG variables. Indeed, it is always very instructive to look at the time derivative $i\mathcal{L}\hat{A}_\mu$ of the CG variables because its *structure* will also determine the structure of the random noises suggesting the form of the matrix $B_{\mu i}$.

3.6 The GENERIC structure

Up to now, when discussing the evolution of the probability of the CG variables, we have always assumed that the Hamiltonian, as well as any other dynamic invariant are included in the list of the CG variables. In many situations, however, the Hamiltonian may be expressed in terms of the CG variables themselves

$$\hat{H}(z) = E(\hat{A}(z)) \tag{3.6.1}$$

where $E(a)$ is a function referred to as the **CG energy function**. When (3.6.1) holds we should exclude the Hamiltonian from the list of CG variables, as it is redundant information. In fact, as the CG variables are of the form (4.3.2), if we include the Hamiltonian in the list of variables we would have from (3.6.1)

$$\Psi_{Ea}(z) = \delta(\hat{H}(z) - E)\delta(\hat{A}(z) - a) = \delta(E(a) - E)\delta(\hat{A}(z) - a) \tag{3.6.2}$$

which has a Dirac delta function $\delta(E(a) - E)$ that does not depend on the microstate z anymore and which is just a trivial overall prefactor. In order to understand the role of energy conservation (or of any other dynamic invariant in the system), it is useful to exploit the consequences of the form (3.6.1) in the form of the FPE obtained by Green and Zwanzig. By doing this, we will see that a very rich structure, known as the GENERIC structure, is revealed.

Note that when the Hamiltonian cannot be expressed as a function of the rest of CG variables (so we include it in the list of CG variables), we will also have a GENERIC structure because then the Hamiltonian is expressible in terms of *one* function in the list, the Hamiltonian itself. In this case, the energy function $E(a)$ is just the identity function. However, in this case the GENERIC structure is somehow trivially hidden. We note in passing that those levels of descriptions that do not allow to express the Hamiltonian as in (3.6.1) typically describe *isothermal* situations. As we will see, the GENERIC framework is a powerful realization in the Theory of Coarse-Graining that allows to fully describe the role of energy conservation in the Fokker-Planck equation.

When (3.6.1) holds, the time derivative of the CG variables admit the following form

$$i\mathcal{L}\hat{A}_{\mu}(z) = \frac{\partial \hat{A}_{\mu}}{\partial z} \cdot J \cdot \frac{\partial \hat{H}}{\partial z} = \frac{\partial \hat{A}_{\mu}}{\partial z} \cdot J \cdot \frac{\partial \hat{A}_{\nu}}{\partial z}(z) \frac{\partial E}{\partial a_{\nu}}(\hat{A}(z)) = \{\hat{A}_{\mu}, \hat{A}_{\nu}\} \frac{\partial E}{\partial a_{\nu}}(\hat{A}(z)) \tag{3.6.3}$$

where we have introduced the Poisson bracket through (1.4.4). When we insert this form of the time derivative in the drift term, involving a conditional expectation we arrive at

$$V_{\mu}(a) = \langle i\mathcal{L}\hat{A}_{\mu} \rangle^{a} = L_{\mu\nu}(a) \frac{\partial E}{\partial a_{\nu}}(a) \tag{3.6.4}$$

where we have introduced the **reversible matrix** as the conditional expectation of the Poisson bracket of the CG variables, this is

$$L_{\mu\nu}(a) \equiv \left\langle \frac{\partial \hat{A}_{\mu}}{\partial z} \cdot J \cdot \frac{\partial \hat{A}_{\nu}}{\partial z} \right\rangle^{a} = \left\langle \{\hat{A}_{\mu}, \hat{A}_{\nu}\} \right\rangle^{a} \tag{3.6.5}$$

By inserting the drift term into the FPE (3.4.37) we have

$$\partial_t P(a,t) = -\frac{\partial}{\partial a} \cdot \left[L(a) \cdot \frac{\partial E}{\partial a} + D(a) \cdot \frac{\partial S_B}{\partial a}(a) \right] P(a,t) + \frac{\partial}{\partial a} \cdot \left[D(a) \cdot \frac{\partial}{\partial a} P(a,t) \right] \quad (3.6.6)$$

3.6.1 What is the GENERIC structure?

GENERIC is an acronym that stands for General Equation for Non-Equilibrium Reversible Irreversible Coupling. The GENERIC framework is a phenomenological thermodynamic approach for the description of non-equilibrium dynamics of systems. It was introduced in two seminal papers by Öttinger and Grmela [?, ?] although the equations themselves were already introduced earlier by Kaufman [?]. A full description of the framework with many applications in the field of complex fluids is given in the textbook by Öttinger [?]. We here present a very brief account of the framework and refer the reader to the mentioned textbook for further reference.

The GENERIC framework postulates that the evolution of the state a of a system when fluctuations are irrelevant evolves according to a universal equation of the form

$$\frac{da}{dt} = L(a) \cdot \frac{\partial E}{\partial a}(a) + M(a) \cdot \frac{\partial S}{\partial a}(a) \quad (3.6.7)$$

Here, the four building blocks of GENERIC are the energy $E(a)$, the entropy $S(a)$, the reversible matrix $L(a)$, and is the friction matrix $M(a)$. These building blocks satisfy a number of properties that strongly restrict the possible structure of the dynamic equation:

$$L^T(a) = -L(a) \qquad \text{the reversible matrix is antisymmetric}$$
$$M^T(a) = M(a) \qquad \text{the friction matrix is symmetric}$$
$$M(a) \geq 0 \qquad \text{the friction matrix is positive semi-definite}$$

$$L(a) \cdot \frac{\partial S}{\partial a}(a) = 0 \qquad \text{degeneracy condition}$$
$$M(a) \cdot \frac{\partial E}{\partial a}(a) = 0 \qquad \text{degeneracy condition} \quad (3.6.8)$$

In addition, it is assumed that the Poisson structure generated by the reversible part of the dynamics satisfies the Jacobi identity. The last two degeneracy properties give rise to the name of the framework, as they reflect the form of the reversible and irreversible coupling. With the above properties it is straightforward application of the chain rule to show that the dynamic equations imply

$$\frac{dE}{dt}(a) = 0$$
$$\frac{dS}{dt}(a) \geq 0 \quad (3.6.9)$$

and, therefore, the GENERIC equation embodies the First and Second Laws of Thermodynamics automatically. The stringent properties on the four building blocks are

usually a powerful guide for *constructing* new dynamic equations for the description of non-equilibrium processes in complex systems. It is also very useful to *check the thermodynamic consistency* of existing dynamic equations for complex systems.

3.6.2 Statistical Mechanics basis of GENERIC

We can now return to the FPE (3.6.6) and appreciate, in the drift term containing only first derivatives the same structure as in GENERIC except for the trivial redefinitions of the friction matrix $M(a) = D(a)/k_B$ and the use of a dimensional bare entropy $S_B(a) = k_B \ln \Omega(a)$ in the GENERIC framework. The matrix $M(a)$ is usually referred to as the friction matrix, although in many applications it has the meaning of a diffusion matrix. To encompass all the different applications, we refer to the matrix $D(a)$ simply as the dissipative matrix.

Given that all the elements of the FPE (3.6.6) have microscopic definitions, it is a natural question to ask wether these elements fullfill the properties postulated in the GENERIC framework.

$L(a)$ is antisymmetric

This is an immediate consequence of the antisymmetry of the Poisson bracket in the definition (1.4.4).

The Jacobi identity

If there was no dissipation in the sense that the dissipative matrix $D_{\mu\nu}(a) = 0$, the FPE (3.4.37) would become a Liouville equation, corresponding to a Hamiltonian motion of the form $\dot{a} = L \cdot \partial E / \partial a$. The Hamiltonian form suggests to introduce a Poisson bracket at the CG level, as follows. Consider two arbitrary functions of the CG variables $F(a), G(a)$ and construct the Poisson bracket corresponding to the reversible matrix, this is

$$\{F, G\}(a) \equiv \frac{\partial F}{\partial a_\mu}(a) L_{\mu\nu}(a) \frac{\partial G}{\partial a_\nu}(a) \tag{3.6.10}$$

In this way, the reversible dynamics takes the form for

$$\dot{a} = \{a, E\} \tag{3.6.11}$$

An important property of the Poisson bracket is the Jacobi identity

$$\{F, \{G, K\}\} + \{G, \{K, F\}\} + \{K, \{F, G\}\} = 0 \tag{3.6.12}$$

for arbitrary functions $F(a), G(a), K(a)$. The Jacobi identity implies the time invariance of the Poisson bracket structure, this is

$$\frac{d}{dt}\{F, G\} = \{\dot{F}, G\} + \{F, \dot{G}\}. \tag{3.6.13}$$

Does the Poisson bracket defined in (3.6.10) satisfies the Jacobi identity? To answer to this question, note that by using the form (3.6.5) we have

$$\{F, G\} = \frac{\partial F}{\partial a_\mu}(a) \left\langle \{\hat{A}_\mu, \hat{A}_\nu\} \right\rangle^a \frac{\partial G}{\partial a_\nu}(a) = \left\langle \frac{\partial F}{\partial a_\mu}(\hat{A})\{\hat{A}_\mu, \hat{A}_\nu\}\frac{\partial G}{\partial a_\nu}(\hat{A}) \right\rangle^a = \left\langle \{\hat{F}, \hat{G}\} \right\rangle^a$$

(3.6.14)

where we have introduced the phase functions $\hat{F}(z) = F(\hat{A}(z))$ and $\hat{G}(z) = G(\hat{A}(z))$. It proves convenient to introduce the notation

$$\widehat{\{F, G\}}(z) = \left\langle \{\hat{F}, \hat{G}\} \right\rangle^{\hat{A}(z)}$$

(3.6.15)

Now consider $\{K, \{F, G\}\}$ which, from (3.6.14) is given by

$$\{K, \{F, G\}\} = \left\langle \{\hat{K}, \widehat{\{F, G\}}\} \right\rangle^a$$

(3.6.16)

Now, assume for a moment that the following identity holds

$$\widehat{\{F, G\}}(z) = \{\hat{F}, \hat{G}\}(z)$$

(3.6.17)

Then we have that (3.6.16) is

$$\{K, \{F, G\}\}(a) = \left\langle \{\hat{K}, \{\hat{F}, \hat{G}\}\} \right\rangle^a$$

(3.6.18)

In this case, by simple linearity of the average, the Jacobi identity at the CG level is ensured by the Jacobi identity at the microscopic level. However, (3.6.17) is not true in general for arbitrary functions $\hat{F}, \hat{G}$. When is it true? Note that (3.6.17) is

$$\left\langle \frac{\partial}{\partial z}F(A)\cdot J\cdot\frac{\partial}{\partial z}G(A) \right\rangle^{\hat{A}(z)} = \frac{\partial}{\partial z}F(\hat{A}(z))\cdot J\cdot\frac{\partial}{\partial z}G(\hat{A}(z))$$

$$\frac{\partial F}{\partial a_\mu}(\hat{A}(z)) \left\langle \frac{\partial \hat{A}_\mu}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\nu}{\partial z} \right\rangle^{\hat{A}(z)} \frac{\partial G}{\partial a_\mu}(\hat{A}(z)) = \frac{\partial F}{\partial a_\mu}(\hat{A}(z))\frac{\partial \hat{A}_\mu}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\nu}{\partial z}\frac{\partial G}{\partial a_\mu}(\hat{A}(z))$$

(3.6.19)

Therefore, it seems that we may ensure the Jacobi identity whenever the CG variables satisfy the identity

$$\left\langle \frac{\partial \hat{A}_\mu}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\nu}{\partial z} \right\rangle^{\hat{A}(z)} = \frac{\partial \hat{A}_\mu}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\nu}{\partial z}(z)$$

(3.6.20)

One way to ensure this equality is by assuming that all the Poisson brackets $\{\hat{A}_\mu, \hat{A}_\nu\}$ are functions of the CG variables themselves, this is

$$\{\hat{A}_\mu, \hat{A}_\nu\}(z) = K_{\mu\nu}(\hat{A}(z)) \tag{3.6.21}$$

This is a condition to be requested to the CG variables in order for the reversible drift to be consistent with the Jacobi identity. Unfortunately, this sufficient condition for the validity of the Jacobi identity is not satisfied in almost none level of description that we consider in this book **[Are you sure???, CHECK in Vol2]**. In summary, the Jacobi identity has not yet been obtained for the microscopic definition of the matrix L, but it is believed that it is a generally valid property that restricts the possible forms of the reversible matrix L [?]. **[Could Jacobi be ensured by the time reversibility of the CG variables?]**

[Can Jacobi be demonstrated in some sort of steepest descent limit of very small fluctuations?]

The degeneracy condition of $L(a)$

An important relation is satisfied by the reversible matrix. We start with the following identity

$$\frac{\partial}{\partial a_\nu} \left[L_{\mu\nu}(a)\Omega(a) \right] = \frac{\partial}{\partial a_\nu} \mathrm{Tr} \left[\hat{\Psi}_a \frac{\partial \hat{A}_\mu}{\partial z} \cdot J \cdot \frac{\partial \hat{A}_\nu}{\partial z} \right] = -\mathrm{Tr} \left[\frac{\partial \hat{\Psi}_a}{\partial z} \cdot J \cdot \frac{\partial \hat{A}_\mu}{\partial z} \right]$$

$$= \mathrm{Tr} \left[\hat{\Psi}_a J : \frac{\partial^2 \hat{A}_\mu}{\partial z \partial z} \right] = 0 \tag{3.6.22}$$

where in the last identity we are using that the trace of a symmetric matrix and anti-symmetric matrix vanishes. The above identity implies that

$$L_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu} = -\frac{\partial}{\partial a_\nu} L_{\mu\nu}(a) \tag{3.6.23}$$

where we have introduced the bare entropy (3.4.39).

In many applications the entropy is an extensive quantity that scales with the size of the system while $L_{\mu\nu}$ is independent of the size of the system. In these cases, the right hand side of (3.6.23) is much "bigger" than the left hand side and one may neglect the latter. In this case, one has the approximate degeneracy property

$$L_{\mu\nu}(a)\frac{\partial S}{\partial a_\nu}(a) \simeq 0 \tag{3.6.24}$$

When the entropy scales with the size of the system, fluctuations are typically small, and (3.6.24) is accurate. However, when fluctuations are important, we cannot assume (3.6.24). In fact, the correct form (3.6.23) will ensure the proper equilibrium distribution.

A property closely related to (3.6.23) of the drift term is the following

$$\frac{\partial}{\partial a_\mu} V_\mu(a)\Omega(a) = \frac{\partial}{\partial a_\mu} \int dz \rho_N^0 \delta(\hat{A}(z) - a) i\mathcal{L}\hat{A}_\mu(z) = -\int dz \rho_N^0 i\mathcal{L}\delta(\hat{A}(z) - a)$$
$$= 0 \tag{3.6.25}$$

where we have used the definition of the reversible drift (3.4.38). We could have, alternatively, have demonstrated this identity by using the form (3.6.4) for the drift in (3.6.25), leading to (3.6.23). The identity (3.6.25) which is a reformulation of (3.6.23) will ensure the correct equilibrium distribution.

$D(a)$ is symmetric

This cannot be ensured from the microscopic expression (**??**). Although only the symmetric part of $D(a)$ appears in the second derivative terms of the FPE (3.6.6), because any skew-symmetric part would dissapear upon contraction with the two derivatives, the possibility of an antisymmetric contribution in the dissipative drift part cannot be ruled out. Note, that an antisymmetric contribution would not contribute to the production of the entropy.

The dissipative matrix satisfies Onsager-Casimir reciprocity

Although the dissipative matrix is not symmetric in general, it does fullfill the Onsager-Casimir reciprocity. **[The Onsager-Casimir reciprocity will be derived in the probabilistic flavour when we consider external forcing]**

The degeneracy condition of $D(a)$

The dissipative matrix (**??**) turns out to be degenerate, a fact that is not very apparent in Green's and Zwanzig's approach, but that is manifest when the Hamiltonian is expressed in terms of the CG variables as in (**??**). We first note that we have the following identity, which is nothing else than the conservation of the energy $i\mathcal{L}\hat{H} = 0$, i.e.

$$i\mathcal{L}\hat{A}_\mu(z)\frac{\partial E}{\partial a_\mu}(\hat{A}(z)) = \frac{\partial E}{\partial a_\mu}(\hat{A}(z))\{\hat{A}_\mu, \hat{A}_\nu\}\frac{\partial E}{\partial a_\nu}(\hat{A}(z)) = 0 \tag{3.6.26}$$

In a similar way, the reversible drift satisfies,

$$\frac{\partial E}{\partial a_\mu} V_\mu = \frac{\partial E}{\partial a_\mu} L_{\mu\nu} \frac{\partial E}{\partial a_\nu} = 0 \tag{3.6.27}$$

due to the antisymmetry of the reversible matrix $L_{\mu\nu}$.

Next, consider the combination

$$\delta(\hat{A}(z) - a)\left[i\mathcal{L}\hat{A}_\mu(z) - V_\mu(a)\right]\frac{\partial E}{\partial a_\nu}(a)$$
$$= \delta(\hat{A}(z) - a)i\mathcal{L}\hat{A}_\mu(z)\frac{\partial E}{\partial a_\nu}(\hat{A}(z)) - \delta(\hat{A}(z) - a)V_\mu(a)\frac{\partial E}{\partial a_\nu}(a) = 0 \tag{3.6.28}$$

on due account of (3.6.26) and (3.6.27). Therefore, when we consider the action of the dissipative matrix on the gradient of the energy we have

$$D_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a) = \int_0^{\Delta t} d\tau \text{Tr}\left[\frac{\hat{\Psi}_a}{\Omega(a)}\left(i\mathcal{L}\hat{A}_\nu - V_\nu(a)\right)\frac{\partial E}{\partial a_\nu}(a)\exp\left\{i\mathcal{L}\tau\right\}\left(i\mathcal{L}\hat{A}_\mu - V_\mu(a)\right)\right]$$
$$= 0 \tag{3.6.29}$$

where we have used (3.6.28). Therefore, the dissipative matrix has an eigenvector, the gradient of the energy function, that has a zero eigenvalue. The dissipative matrix is degenerate.

In order to make contact with Grabert's textbook we note that, on account of the GENERIC properties (3.6.23),(3.6.29), the FPE (3.6.6) can be written in the compact form given by Grabert

$$\partial_t P(a,t) = -\frac{\partial}{\partial a}\left[L^*(a)\cdot\left[\frac{\partial F}{\partial a} - \frac{\partial}{\partial a}\right]P(a,t)\right] \tag{3.6.30}$$

where the bare free energy and bare transport coefficients are

$$F^*(a) = E(a) + k_B T S_B(a)$$
$$L^*(a) = L(a) + \frac{1}{k_B T}D(a) \tag{3.6.31}$$

where we can use any value for the thermal energy $k_B T$. We prefer the form (3.6.6) of the Fokker-Planck Equation because it displays a much richer structure that strongly constraints the possible models that we may adopt for the reversible and irreversible operators.

3.6.3   GENERIC is hidden in Green and Zwanzig's approach

What is the GENERIC structure of the FPE given by Green and Zwanzig when the total energy *is not* expressible in terms of the rest of CG variables? In this case, we should include the total energy as an additional CG variable. Assume that the only dynamic invariant is the Hamiltonian $\hat{H}(z)$, and let $\hat{B}(z)$ be the set of relevant variables that are not dynamic invariants. In this way, the set of CG variables are $\hat{A}_0(z) = \hat{H}(z)$ and $\hat{A}_\mu(z) = B_\mu(z)$, with $\mu = 1, \cdots, M$. We start first by the FPE (3.4.37) given by Green and Zwanzig and realize that because the dynamic invariants like the energy $\hat{H}(z)$ satisfy $i\mathcal{L}\hat{H}(z) = 0$, a number of terms in the FPE (3.4.37) vanish. With this new notation the drift term (3.4.38) will have now the form $V_\mu(E, b)$

$$V_0(E,b) = \left\langle i\mathcal{L}\hat{H}\right\rangle^{Eb} = 0$$
$$V_\mu(E,b) = \left\langle i\mathcal{L}\hat{B}_\mu\right\rangle^{Eb} \tag{3.6.32}$$

The dissipative matrix $D_{\mu\nu}(E,b)$ in Eq. (3.4.27), again due to the fact that $i\mathcal{L}\hat{H}(z) = 0$, will vanish if $\mu = 0$ or $\nu = 0$. For the rest of components will have the form

$$D_{\mu\nu}(E,b) = \int_0^{\Delta t} d\tau \left\langle (i\mathcal{L}\hat{B}_\nu - V_\nu(E,b))) \exp\{i\mathcal{L}\tau\} (i\mathcal{L}\hat{B}_\mu - V_\mu(E,b)) \right\rangle^{Eb} \quad (3.6.33)$$

As a result, all terms in the Fokker-Planck equation that multiply derivatives with respect to E will vanish and will not appear in the FPE. The Fokker-Planck (3.4.37) when we single out the dynamical invariants become

$$\partial_t P(E,b,t) = -\frac{\partial}{\partial b_\mu} \left[V_\mu(E,b) + D_{\mu\nu}(E,b) \frac{\partial S_B}{\partial b_\nu}(E,b) \right] P(E,b,t)$$
$$+ \frac{\partial}{\partial b_\mu} D_{\mu\nu}(E,b) \frac{\partial}{\partial b_\nu} P(E,b,t) \quad (3.6.34)$$

where only derivatives with respect to the non-conserved variables b appear. If we restrict the discussion to the b variables, it looks like the dissipative matrix is non-degenerate. However, by recalling that the list of variables is the set E,b, the matrix $D_{\mu\nu}(E,b)$ is degenerate (as it has its first row and column filled with zeros).

Consider now the reversible drit term, as specifyed by GENERIC , when we single out the energy in the set of variables $a = E,b$. Then the GENERIC reversible drift takes the form

$$L \cdot \frac{\partial E}{\partial a} = \left(\begin{array}{cc} \left\langle \{\hat{H},\hat{H}\} \right\rangle^{Eb} & \left\langle \{\hat{H},\hat{B}\} \right\rangle^{Eb} \\ \left\langle \{\hat{B},\hat{H}\} \right\rangle^{Eb} & \left\langle \{\hat{B},\hat{B}\} \right\rangle^{Eb} \end{array} \right) \left(\begin{array}{c} \frac{\partial E}{\partial E} \\ \frac{\partial E}{\partial b} \end{array} \right)$$
$$= \left(\begin{array}{cc} 0 & \left\langle \{\hat{H},\hat{B}\} \right\rangle^{Eb} \\ \left\langle \{\hat{B},\hat{H}\} \right\rangle^{Eb} & 0 \end{array} \right) \left(\begin{array}{c} 1 \\ 0 \end{array} \right) = \left(\begin{array}{c} 0 \\ V(E,b) \end{array} \right) \quad (3.6.35)$$

where we have used the definition (3.6.32) of the drift term. This identity thus shows that the FPE (3.6.34) has the GENERIC structure. Therefore, in both cases, either if the Hamiltonian is a function of the CG variables (and therefore not included in the list) or if it is not (in which case we should include it in the list of CG variables) the FPE in the probabilistic flavour has the GENERIC structure.

3.6.4   The Stochastic Differential Equation

The Ito SDE corresponding to the FPE (3.6.6) is

$$da = L(a) \cdot \frac{\partial E}{\partial a}(a) dt + D(a) \cdot \frac{\partial S_B}{\partial a}(a) dt + \frac{\partial \cdot D^S}{\partial a}(a) dt + d\tilde{a} \quad (3.6.36)$$

The divergence of the dissipative matrix term is referred to as the **spurious drift**. This term is linked to the (Ito) stochastic interpretation of the SDE. The additional term in (3.6.36), which is the random force $d\tilde{a}$ is a linear combination of Wiener processes that

satisfy the Fluctuation-Dissipation theorem

$$d\tilde{a}d\tilde{a}^T = 2dtD^S(a)$$ (3.6.37)

In addition the random noise satisfy the orthogonality condition

$$d\tilde{a}^T \cdot \frac{\partial E}{\partial a} = 0$$ (3.6.38)

in order for the Fluctuation-Dissipation theorem (3.6.37) to be consistent with the degeneracy condition of the friction (3.6.8).

Eq. (3.6.38) has a nice geometrical interpretation. The gradient of the energy is a vector which is perpendicular to the submanifold of constant energy. Therefore, the orthogonality tells us that the random kicks to the state a are tangent to the submanifold of constant energy. In other words, the kicks leave the state in that submanifold and, therefore, they conserve energy. Now, it is easily shown by using Ito calculus that the SDE (3.6.36) conserves energy thanks to (3.6.38). The stochastic differential of the energy $E(a)$ function is, according to Ito calculus

$$dE(t) = \left(\frac{\partial E}{\partial a}\right)^T \cdot da + \frac{1}{2}\frac{\partial^2 E}{\partial a \partial a} : dada.$$ (3.6.39)

Observe the presence of the second order terms which is a consequence of the peculiarities of Ito calculus. By substituting the SDE (3.6.36) in (3.6.39) and by using we arrive at

$$dE(t) = \frac{\partial}{\partial a} \cdot \left[M(a) \cdot \frac{\partial E}{\partial a}\right] = 0,$$ (3.6.40)

Therefore, the GENERIC SDE (3.6.36) conserves energy exactly.

3.7 Another view to Green-Kubo from the SDE

If we write the SDE as

$$dx_t = F(x_t)dt + B(x)dW_t$$ (3.7.1)

we can write

$$B(x_t)\frac{dW_t}{dt} = \frac{dx_t}{dt} - F(x_t)$$ (3.7.2)

Mutiply with $B(x_0)dW_0$ and integrate to get

$$\int_0^\infty dt B(x_t)\frac{dW_t}{dt}B(x_0)\frac{dW_0}{dt} = \int_0^\infty dt\left(\frac{dx_t}{dt}-F(x_t)\right)\left(\frac{dx_0}{dt}-F(x_0)\right)$$

$$\int_0^\infty dt B(x_t)B(x_0)\delta(t) = \int_0^\infty dt\left(\frac{dx_t}{dt}-F(x_t)\right)\left(\frac{dx_0}{dt}-F(x_0)\right)$$

$$B(x_t)B(x_0) = 2\int_0^\infty dt\left(\frac{dx_t}{dt}-F(x_t)\right)\left(\frac{dx_0}{dt}-F(x_0)\right) \qquad (3.7.3)$$

3.8   Einstein formula for equilibrium fluctuations

One obvious question to ask about the FPE (3.4.37) or (3.6.6) is what is its stationary solution. Because of the identity (3.6.25), it looks like

$$P^{\text{eq}}(a) = \Omega(a) = \exp\{S_B(a)\} \qquad \text{(Incorrect)} \qquad (3.8.1)$$

is the equilibrium solution of the FPE (3.4.37) as can be explicitly checked by substitution. Eq (3.8.1) looks like the Einstein form for equilibrium fluctuations [**?**]. Unfortunately, the expression (3.8.1) does not gives a normalized distribution because

$$\int da P^{\text{eq}}(a) = \int da \int dz \rho_N^0 \delta(\hat{A}(z)-a) = \rho_N^0 \int dz = \infty \qquad (3.8.2)$$

and we have to be more careful in obtaining the equilibrium solution of the FPE (3.4.37). The clue to the solution is the explicit consideration of the dynamic invariants. It is easily shown by using the degeneracy properties (3.6.8) that any probabilty of the form

$$P^{\text{eq}}(a) = \Phi(E(a))\exp\{S_B(a)\} \qquad (3.8.3)$$

is an equilibrium solution of the FPE, where $\Phi(E)$ is, in principle, an arbitrary funcion. The functional form of $\Phi(E)$ is, however, fixed by the following argument. The probability $P(E,t)$ that the system has energy E at time t is, by definition

$$P(E,t) = \int da P(a,t)\delta(E(a)-E) \qquad (3.8.4)$$

By using the definition of $P(a,t)$ in terms of the microscopic ensemble we will have

$$P(E,t) = \int da\left[\int dz \rho_t(z)\delta(\hat{A}(z)-a)\right]\delta(E(a)-E) = \int dz \rho_t(z)\delta(\hat{H}(z)-E) \quad (3.8.5)$$

This probability is clearly time-independent as can be seen by taking the time derivative and using Liouville's theorem. Therefore we have that the distribution of energy is

identical at all times to the distribution of energy at the initial time, this is

$$P(E,t) = P(E,0) \equiv P_0(E) \tag{3.8.6}$$

The invariant distribution $P_0(E)$ needs to be produced also by the equilibrium solution (3.8.3)

$$P_0(E) = \int da P^{\mathrm{eq}}(a)\delta(E(a) - E) = \int da \Phi(E(a)) \exp\{S(a)\}\delta(E(a) - E)$$

$$= \Phi(E) \int da \exp\{S(a)\}\delta(E(a) - E) \tag{3.8.7}$$

This equation allows to determine $\Phi(E)$ from the known $P_0(E)$. The equilibrium solution of the GENERIC FPE (3.6.6) is

$$P^{\mathrm{eq}}(a) = \frac{P_0(E(a)) \exp\{S(a)\}}{\int da' \exp\{S(a')\}\delta(E(a') - E(a))} \qquad \text{(Correct)} \tag{3.8.8}$$

This is the **Einstein formula** for fluctuations [**?**], suitably modified to include dynamic invariants [**?**].

3.9   The Second Law in the probabilistic flavour

The Second Law in the probabilistic flavour is automatically satisfied because *any* Fokker-Planck Equation has an H-theorem [?]. Let us describe the details when we have the particular form (3.4.37) of the FPE.

The entropy of the probabilistic flavour is given by the entropy functional (2.7.28). We take, for simplicity a system with a fixed number of particles and the entropy functional is

$$S[P] = - \int da\, P(a,t) \ln \frac{P(a,t)}{\Omega(a)} \qquad (3.9.1)$$

Note that the entropy functional is always negative and has a global maximum when the functions $P(a,t) = \Omega(a)$. **[Warning $\Omega(a)$ is not normalized!]**

We show now that the entropy functional has a time derivative that is always positive or zero. By using the chain rule we have

$$\frac{d}{dt} S[P(a,t)] = \int da \frac{\delta S}{\delta P(a,t)} \partial_t P(a,t)$$
$$= - \int da \left[\ln \frac{P(a,t)}{\Omega(a)} + 1 \right] \frac{\partial}{\partial a_\mu} \left[V_\mu(a) P(a,t) + D_{\mu\nu}(a)\Omega(a) \frac{\partial}{\partial a_\nu} \frac{P(a,t)}{\Omega(a)} \right] \qquad (3.9.2)$$

By integrating by parts (and assuming that any surface term vanishes), we have

$$\frac{d}{dt} S[P(a,t)] = \int da \frac{P^{\mathrm{eq}}(a)}{P(a,t)} \frac{\partial}{\partial a_\mu} \frac{P(a,t)}{\Omega(a)} \left[-V_\mu(a) P(a,t) + D_{\mu\nu}(a)\Omega(a) \frac{\partial}{\partial a_\nu} \frac{P(a,t)}{\Omega(a)} \right] \qquad (3.9.3)$$

The reversible term does not change the entropy functional because

$$\int da \frac{\Omega(a)}{P(a,t)} \frac{\partial}{\partial a_\mu} \frac{P(a,t)}{\Omega(a)} V_\mu(a) P(a,t) = \int da \Omega(a) \frac{\partial}{\partial a_\mu} \frac{P(a,t)}{\Omega(a)} V_\mu(a)$$
$$= - \int da \frac{P(a,t)}{\Omega(a)} \frac{\partial}{\partial a_\mu} V_\mu(a)\Omega(a) = 0 \qquad (3.9.4)$$

where in the last equality we have integrated by parts and have used (3.6.25). Therefore, we have

$$\frac{d}{dt} S[P(a,t)] = \int da \frac{\Omega^2(a)}{P(a,t)} \frac{\partial}{\partial a_\mu} \frac{P(a,t)}{\Omega(a)} D_{\mu\nu}(a) \frac{\partial}{\partial a_\nu} \frac{P(a,t)}{\Omega(a)}$$
$$= \int da\, P(a,t) \left[\frac{\partial}{\partial a_\mu} \ln \frac{P(a,t)}{\Omega(a)} \right] D_{\mu\nu}(a) \left[\frac{\partial}{\partial a_\nu} \ln \frac{P(a,t)}{\Omega(a)} \right] \geq 0 \qquad (3.9.5)$$

This time derivative is always positive because the dissipative matrix is positive semi-definite and the probability density $P(a,t)$ is also positive. This means that the entropy functional (3.9.1) always increases. As it has a unique global maximum, this shows that

the probability distribution $P(a,t)$ becomes identical to $\Omega(a)$ in the long time limit and, therefore, $\Omega(a)$ is the unique equilibrium solution of the FPE.

But wait! We already saw in (3.8.1) that $\Omega(a)$ cannot be the equilibrium solution. How can it be that the entropy functional always increases, heading towards its global maximum, but never attaining it? The solution to this apparent paradox lies, again in the consideration of the dynamic invariants (the energy, for the sake of simplicity). Note that if we single out the energy from the list of variables, the evolution of $P(E,a,t)$ is subject to the constraint

$$\int dE da E P(E,a,t) = ctn \qquad (3.9.6)$$

Therefore, the evolution does not go, in the functional space of probabilities, towards the global maximum of the entropy, but rather at a maximum constrained by (3.9.6). An alternative to appreciate this is in the expression (3.9.5). Note that, when we single out the total energy, many elements of the dissipative matrix vanish. This means that any probability of the form (??) inserted in (??) will give a vanishing time derivative. Therefore, the system stops evolving once reached the distribution (??).

The bare entropy does not obey a Second Law

It is instructive to look now at the stochastic evolution of the bare entropy $S_B(a)$ introduced in (3.4.39) according to the SDE (3.4.41). As we discuss in the Appendix 24, in a stochastic setting, we need to go to second order in a Taylor expansion

$$dS_B(a(t)) = \frac{\partial S_B}{\partial a_\mu}(a(t))da_\mu(t) + \frac{1}{2}\frac{\partial^2 S_B}{\partial a_\mu \partial a_\nu}(a(t))da_\mu(t)da_\nu(t) \qquad (3.9.7)$$

By using (??) and (3.4.44) we have (omit the arguments for clarity)

$$dS_B(t) = \frac{\partial S_B}{\partial a_\mu}\left[V_\mu dt + D_{\mu\nu}\frac{\partial S_B}{\partial a_\mu \partial a_\nu}D_{\mu\nu}\right]dt \qquad (3.9.8)$$

where we have neglected terms higher than dt. Regrouping terms we have

$$dS_B(t) = \frac{\partial S_B}{\partial a_\mu}V_\mu dt + \frac{\partial S_B}{\partial a_\mu}D_{\mu\nu}\frac{\partial S_B}{\partial a_\nu}dt + \frac{\partial}{\partial a_\mu}\left[S_B\frac{\partial D_{\mu\nu}}{\partial a_\nu}\right]dt + \frac{\partial S_B}{\partial a_\mu}d\tilde{a}_\mu(t) \qquad (3.9.9)$$

Note that (3.6.25) in terms of the bare entropy is

$$\frac{\partial S_B}{\partial a_\mu}V_\mu dt = -\frac{\partial V_\mu}{\partial a_\mu} \qquad (3.9.10)$$

then,

$$dS_B(t) = \frac{\partial S_B}{\partial a_\mu}D_{\mu\nu}\frac{\partial S_B}{\partial a_\nu}dt + \frac{\partial}{\partial a_\mu}\left[-V_\mu + S_B\frac{\partial D_{\mu\nu}}{\partial a_\nu}\right]dt + \frac{\partial S_B}{\partial a_\mu}d\tilde{a}_\mu(t) \qquad (3.9.11)$$

This equation is stochastic. The first term satisfies

$$\frac{\partial S_B}{\partial a_\mu} D_{\mu\nu} \frac{\partial S_B}{\partial a_\nu} \geq 0 \tag{3.9.12}$$

that shows that $S_B(t)$ is "typically" growing in time. However, neither the second term nor the third stochastic term have definite character and, therefore, the function $S_B(t)$ is not ensured to increase in time. It is for this reason that we do not call $S_B(t)$ an entropy, which is supposed to satisfy a Second Law. We will see in the next section that when fluctuations are negligible, the second and last terms of (3.9.11) are small and, then, the index of probability $S_B(a(t))$ satisfies approximately the Second Law.

3.10   The limit of small fluctuations

The right hand side of the FPE (3.4.37) has two terms, a reversible drift term that involves only first derivatives with respect to the CG variables and a diffusion term that contains second derivatives of the probability distribution function. The second derivative term controls the *broadening* of the probability distribution as time proceeds. To see this, we will assume for a moment that we may neglect the diffusion term and show that then there is no broadening. When we neglect the diffusion term, the FPE has the form of a Liouville equation with only first derivative terms, i.e.

$$\partial_t P(a,t) = -\frac{\partial}{\partial a} \cdot \left[V(a) + D(a) \cdot \frac{\partial S_B}{\partial a}(a) \right] P(a,t) \qquad (3.10.1)$$

If we know with precision the initial value a_0 of the CG variables the initial condition is

$$P(a,0) = \delta(a - a_0) \qquad (3.10.2)$$

It is simple to prove that the probability distribution

$$P(a,t) = \delta(a(t,a_0) - a) \qquad (3.10.3)$$

is a solution of (3.10.1) with initial condition (3.10.2) provided that $a(t,a_0)$ is a solution of

$$\dot{a}(t) = V(a(t)) + D(a(t)) \cdot \frac{\partial S_B}{\partial a}(a(t)) \qquad (3.10.4)$$

with initial condition a_0. This equation (3.10.4) will be referred to as the **deterministic limit** of the SDE (3.4.41). We see that the probability remains always peaked at the solution of (3.10.4) and it is obvious that any broadening of $P(a,t)$ with time is due to the diffusion term involving the diffusion matrix $D_{\mu\nu}(a)$. When the diffusion term cannot be neglected, the distribution function will broaden, from the initial condition (3.10.2) until it reaches the equilibrium distribution $P^{\text{eq}}(a)$. Of course, if the equilibrium distribution is highly peaked, then we expect that the whole evolution of the distribution will be highly peaked too.

Typically, fluctuations are small when the CG variables are **extensive** and, in addition, certain scaling behaviour is obeyed by all the objects in the Fokker-Planck Equation. By extensive CG variables we usually mean phase functions that are made of additive contributions of "many" particles. Therefore, what determines the importance of fluctuations is, essentially, how broad is the equilibrium distribution itself. The equilibrium probability is given by Einstein formula (**??**). When should we expect that the probability distribution is highly peaked and the fluctuations be small? When the bare $S_B(a)$ is "very large", the equilibrium probability will be a highly peaked distribution. Of course, this depends on the functional form of the CG variables and, therefore, the possibility of neglecting fluctuations is closely related to this functional form.

As a rule of thumb, CG variables (i.e. they are *extensive*) will show small fluctuations. To see when we may expect small fluctuations, let us imagine that we may associate a dimensionless parameter $\hat{N}$ to the CG variables, this is, the CG variables are of the form

$\hat{A}(z) = N\hat{X}(z)$ with numerical values also satisfying $a = Nx$, where x is an *intensive* variable. We further assume that the different objects of the FPE (3.4.37) display the following scaling

$$
\begin{aligned}
S_B(a) &= N s_B(x) \\
V(a) &= N v(x) \\
D(a) &= N d(x)
\end{aligned}
\tag{3.10.5}
$$

From the microscopic definition of the drift (3.4.38) the assumed scaling for $V(a)$ is rather natural due to the appearance of $i\mathcal{L}\hat{A}$. A bit more problematic seems to be the diffusive matrix, defined by the Green-Kubo formula (**??**). The Green-Kubo expression contains two terms of the form $i\mathcal{L}\hat{A}$ and would, therefore, seems that the dissipative matrix should scale as N^2. However, what happens usually is that $i\mathcal{L}\hat{A}_\mu$ is a function with N terms that is correlated with just a few number of the N terms of the current $i\mathcal{L}\hat{A}_\nu$. Hence the plausible scaling as N of the dissipative matrix.

We write now the FPE for the probability density $p(x,t) = P(Nx,t)$ of the intensive variables x. A simple use of the chain rule and the scaling assumptions (3.10.5) In terms of the intensive variables, the FPE equation (3.4.37) takes the form

$$
\partial_t p(x,t) = -\frac{\partial}{\partial x} \cdot \left[v(x) + d(x) \cdot \frac{\partial s_B}{\partial x}(x) \right] p(x,t) + \frac{1}{N} \frac{\partial}{\partial x} \cdot d(x) \cdot \frac{\partial}{\partial x} p(x,t)
\tag{3.10.6}
$$

In these variables, the diffusion term appears scaled with a factor N^{-1}. When N is large, this term is small. Therefore, we see that for extensive CG variables when the scaling relations (3.10.5) hold, we expect small fluctuations.

One may try to perform a perturbation calculation in terms of the small parameter $\epsilon = N^{-1}$. However, this perturbation is singular because setting $\epsilon = 0$, which multiplies second order derivatives, changes the order of the partial differential equation. The proper way to do this perturbation is presented in Gardiner's textbook [**?**]. Nevertheless, we know that in the limit $N \to \infty$ the solution of the FPE (3.4.37) will be very similar to the Dirac delta function (3.10.3).

3.10.1 The entropy

What is the form of the entropy functional (2.7.28) in the limit of small fluctuations? A first estimate is to plug (3.10.3) into (2.7.28) with the formal result

$$
\begin{aligned}
\mathcal{S}[P] &= -\int da\,\delta(a(t,a_0) - a)) \ln \frac{\delta(a(t,a_0) - a)}{P_0 \exp\{S_B(a)\}} \\
&= S_B(a(t,a_0)) - \int da\,\delta(a(t,a_0) - a) \ln \frac{\delta(a(t,a_0) - a)}{P_0}
\end{aligned}
\tag{3.10.7}
$$

By a change of variables $x = a(t,a_0) - a$ the last integral is simply

$$
\int dx\,\delta(x) \ln \frac{\delta(x)}{P_0} = \ln \frac{\delta(0)}{P_0}
\tag{3.10.8}
$$

although scaringly singular, the important thing is that this quantity is just a number, independent of time[3]. As we know that the bare is defined up to an irrelevant constant, we may simply absorb this into the definition of the bare entropy.

In summary, in the limit of small fluctuations the bare entropy coincides with the entropy functional. Because the entropy functional $\mathcal{S}[P]$ satisfies the H-theorem (3.9.5), even in the limit of highly peaked distributions, we expect that the bare entropy evaluated at the deterministic solution, $S_B(a(t, a_0))$, will have a positive time derivative. We may check this explicitly by computing

$$\frac{dS_B}{dt} = \frac{\partial S_B}{\partial a_\mu} \dot{a}_\mu = \frac{\partial S_B}{\partial a_\mu} \left[V_\mu + D_{\mu\nu} \frac{\partial S_B}{\partial a} \right] \tag{3.10.9}$$

Note that the property (3.6.25) implies

$$V_\mu(a) \frac{\partial S_B}{\partial a_\mu}(a) = \frac{\partial}{\partial a_\mu} V_\mu(a) \tag{3.10.10}$$

Therefore we have

$$\frac{dS_B}{dt} = \frac{\partial S_B}{\partial a_\mu} \dot{a}_\mu = \frac{\partial S_B}{\partial a_\mu} D_{\mu\nu} \frac{\partial S_B}{\partial a} - \frac{\partial}{\partial a_\mu} V_\mu(a) \tag{3.10.11}$$

Note that for extensive variables as we are assuming the last term scales as N^0, while the first term scales as N^1. Therefore, we may neglect the last term in front of the first one. Because the first term is always positive, we obtain that, indeed, the bare entropy is always an increasing function of time, as advanced.

3.10.2 Comparing the averages and probabilistic flavours

The deterministic evolution (3.10.4) has the same structure as the transport equation for the average values (**??**). However, the particular expressions are different in general. For example, the bare entropy (3.4.39) is a function of the CG variables which is different from the entropy (2.7.13), which is the function that emerges after evaluating the relevant ensemble at the Gibbs-Jaynes entropy functional. There is no reason to expect that they are the same function in general.

However, we expect that in the limit of highly peaked distribution functions, the average values $a(t)$ will coincide with the most probable value of the distribution, which obeys the deterministic equation (3.10.4). Therefore, when fluctuations are negligible we should be able to relate all the terms of the equation (**??**) with all the terms of (3.10.4).

As a first step, let us introduce the following notation for the averages

$$\langle \cdots \rangle_{\text{can}}^a = \int dz \frac{1}{Z(\lambda(a))} \exp\{-\lambda(a)\hat{A}(z)\} \cdots$$

$$\langle \cdots \rangle_{\text{mic}}^a = \int dz \frac{1}{\Omega(a)} \delta(\hat{A}(z) - a) \cdots \tag{3.10.12}$$

---

[3]We can always regularize the integral in a way similar to what we have already done in Sec **??**.

By introducing the identity

$$1 = \int da \delta(\hat{A}(z) - a) \tag{3.10.13}$$

inside the canonical average we obtain the exact relationship between both averages

$$\langle \cdots \rangle_{\text{can}}^a = \int da' P_a(a') \langle \cdots \rangle_{\text{mic}}^{a'} \tag{3.10.14}$$

where we have introduced the probability density that the CG variables take the value a' when the microscopic probability is given by the generalized canonical relevant ensemble determined by a, this is

$$P_a(a') = \int dz \frac{1}{Z(\lambda(a))} \exp\{-\lambda(a) \cdot \hat{A}(z)\} \delta(\hat{A}(z) - a') \tag{3.10.15}$$

which can be written as

$$P_a(a') = \frac{1}{Z(\lambda(a))} \exp\{-\lambda(a) \cdot a'\} \Omega(a') \tag{3.10.16}$$

Note that this probability takes the following form for extensive variables, in which the scaling (3.10.5) holds.

$$P_a(a') = \frac{1}{Z(\lambda(a))} \exp\{N(s_B(x') - \lambda(a) \cdot x'\} \tag{3.10.17}$$

When $N \to \infty$ this is a highly peaked distribution $P_a(a') \approx \delta(a - a')$. Note that by using (2.7.13) and (2.7.10) this probability can be expressed as

$$P_a(a') = \exp\{S_B(a') - S(a) + \lambda(a)(a - a')\} \tag{3.10.18}$$

The maximum of this probability occurs for the value a' which is the solution of

$$\frac{\partial S_B}{\partial a'}(a') = \lambda(a) \tag{3.10.19}$$

and in the limit $N \to \infty$ we have

$$S_B(a) \approx S(a) \tag{3.10.20}$$

Therefore, in the case of extensive variables, when the distribution (3.10.16) becomes highly peaked, the transport equation (**??**) and the deterministic equation (3.10.4) are the same.

3.10.3   Approximating conditional expectations near equilibrium

In many cases, the calculation of the diffusion matrix is a difficult task due to the need of accounting for the conditional expectations. Under the assummption that fluctuations are

small *and* that the system is near equilibrium, we may assume that the diffusion matrix takes a value very close to (**??**) with the averages taken over the equilibrium ensemble instead of the conditional ensemble. Near equilibrium, the states a_0 will be close to the maximum of a_0^* of the equilibrium probability $P^{\mathrm{eq}}(a)$. Because the fluctuations are small, then $P^{\mathrm{eq}}(a) \simeq \delta(a - a_0^*)$ and the equilibrium and conditional expectations differ by small terms that are of the size of the fluctuations of CG variables.

3.11   Summary

This has been a long chapter but, at the end of the day, the results are very rapidly summarized. The basic assumption is that the CG variables have a two-scale evolution, one on a short time scale τ_{mic} stemming from "rapid and small events" that cummulative build the variation on a larger scale τ_{mac}. When this separation of time scales exists, one can model the dynamics of the CG variables with a stochastic diffusion process that is governed with a Fokker-Planck Equation

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[V_\mu(a) + D_{\mu\nu}(a)\frac{\partial S_B}{\partial a}(a)\right]P(a,t) + \frac{\partial}{\partial a_\mu}\,D^S_{\mu\nu}(a)\frac{\partial}{\partial a_\nu}P(a,t) \quad (3.11.1)$$

where $D^S(a) = (D(a) + D^T(a))/2$ is the symmetric part of the dissipative matrix. Here $P(a,t)$ is the probability distribution of CG variables. The reversible drift is

$$V_\mu(a) = \left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a \tag{3.11.2}$$

where the conditional expectation of an arbitrary phase function $\hat{F}(z)$ is

$$\langle\hat{F}\rangle^a = \frac{1}{\Omega(a)}\int dz\,\hat{\Psi}_a(z)\hat{F}(z) \tag{3.11.3}$$

where the measure of the CG variables is

$$\Omega(a) = \int dz\,\hat{\Psi}_a(z) \tag{3.11.4}$$

Its logarithm is the generalized Boltzmann's entropy

$$S_B(a) \equiv \ln\Omega(a) \tag{3.11.5}$$

The drift term (6.3.2) takes also the GENERIC form

$$V_\mu(a) = L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a) \tag{3.11.6}$$

where the CG energy $E(a)$ is obtained from the microscopic Hamiltonian as

$$\hat{H}(z) = E(\hat{A}(z)) \tag{3.11.7}$$

and the reversible operator is defined as

$$L_{\mu\nu}(a) = \left\langle\{\hat{A}_\mu,\hat{A}_\nu\}\right\rangle^a \tag{3.11.8}$$

The reversible satisfies the reversibility condition

$$V_\mu(a)\frac{\partial S}{\partial a_\mu}(a) = -\frac{\partial V_\mu}{\partial a_\mu}(a) \tag{3.11.9}$$

closely related to the (quasi) degeneracy condition of the reversible operator $L(a)$.

$$L_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu} = -\frac{\partial}{\partial a_\nu}L_{\mu\nu}(a) \tag{3.11.10}$$

The dissipative matrix is

$$D_{\mu\nu}(a) = \int_0^{\Delta t} d\tau \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a))) \exp\{-i\mathcal{L}\tau\} (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^a \tag{3.11.11}$$

The dissipative matrix satisfies Onsager's reciprocity

$$D_{\mu\nu}(a) = \varepsilon_\mu \varepsilon_\nu D_{\mu\nu}(\varepsilon{\cdot}a) \tag{3.11.12}$$

The entropy of a probabilistic level of description is

$$S[P] = -\int da P(a) \ln \frac{P(a)}{\Omega(a)} \tag{3.11.13}$$

and it satisfies an H-theorem $\frac{d}{dt}S[P_t] \geq 0$. The H-Theorem leads to the unique equilibrium solution given by the Einstein formula

$$P^{\text{eq}}(a) = \frac{P_0(E(a))\exp\{S(a)\}}{\int da' \exp\{S(a')\}\delta(E(a') - E)} \tag{3.11.14}$$

where $P_0(E)$ is the distribution of the energy in the system at the initial time.

The SDE which is mathematically equivalent to the FPE (6.3.1) is

$$da = V(a)dt + D(a){\cdot}\frac{\partial S_B}{\partial a}(a)dt + \frac{\partial{\cdot}D^S}{\partial a}(a)dt + d\tilde{a} \tag{3.11.15}$$

The random force $d\tilde{a}$ is a linear combination of Wiener processes that satisfy the Fluctuation-Dissipation theorem

$$d\tilde{a}d\tilde{a}^T = 2dt D^S(a) \tag{3.11.16}$$

The random noise satisfy the orthogonality condition

$$d\tilde{a}^T{\cdot}\frac{\partial E}{\partial a} = 0 \tag{3.11.17}$$

4

External forcing

4.1 Introduction

The Non-Equilibrium Statistical Mechanics theory presented so far deals with isolated systems governed by a time-independent Hamiltonian. In these systems, energy is conserved and the only problems that may be addressed are related to the macroscopic decay towards the unique equilibrium state of the system. In the present chapter, we generalize the theory by including external fields. This will allow to study more general situations in which the evolution and response of a system to an external forcing is monitored.

In the classical world at which we restrict ourselves in this book, the external fields to be considered are those due to gravitation, non-inertial forces, and electromagnetism. Gravitation takes a simple time independent form in most situations of interest and can be dealt with the framework for isolated systems. Hamilton's equations are equivalent to Newton's second law and are, therefore, restricted to inertial reference systems. When we express the equations in non-intertial frames, additional non-inertial forces appear. Formally, the functional form of non-inertial forces is identical to those of particular electromagnetic fields.

In this chapter, we derive the stochastic dynamics of the CG variables in the presence of time-dependent external forces. The Fokker-Planck equation is derived and shown to have building blocks that generalize those of the non-forced FPE while still respecting the basic properties of the non-forced case. In addition, we also discuss in this chapter Fluctuation Theorems (FT) for Hamiltonian systems and Linear Response Theory. Strictly speaking, these later results have nothing to do with coarse-graining as they do not require to select CG variables or no assumption about separation of time scales is required for the CG variables. We include these results because they are very general, as in the case of the Fluctuation Theorems, or very useful, as in the case of Linear Response Theory.

4.2   The microscopic dynamics under external forcing

In this section, we generalize the results given in Chapter 1 for the microscopic dynamics when there are external fields present. There, we assumed that the Hamiltonian was autonomous, i.e. did not depend explicitly on time. Now, we will assume that the Hamiltonian is time-dependent, and of the form

$$\hat{H}(z,t) = \hat{H}^0(z) + \hat{H}^1(z,t)$$

$$\hat{H}^0(z) = \sum_i^N \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q}_1, \cdots, \mathbf{q}_N) \tag{4.2.1}$$

where the unperturbed time-independent Hamiltonian $\hat{H}^0(z)$ is the sum of the kinetic energy and the potential energies of interaction between the particles. The contribution $\hat{H}^1(z,t)$ accounts for the interaction of the system with an external field. Typically, we will assume that the explicit form of this interaction is linear in the CG variables $\hat{A}(z)$

$$\hat{H}^1(z,t) = \hat{H}^1(z,\phi_t) = -\hat{A}(z){\cdot}\phi_t \tag{4.2.2}$$

The vector ϕ_t represents a time-dependent external field which is independent of the microstate z. We will also refer to ϕ_t as the **protocol**. The fact that now the dynamics depends implicitly on the protocol, and that we will compare processes with "reversed" protocols, requires to use a somewhat cumbersome notation. We believe that this is unavoidable. Many expressions in the literature, although correct, are obscured behind words that do not always convey what is actually being done.

Hamilton's equations are now

$$\dot{z}_t = J{\cdot}\frac{\partial \hat{H}}{\partial z}(z_t, \phi_t) \tag{4.2.3}$$

This generalizes (1.3.4) to time-dependent Hamiltonians. We may write Hamilton's equations (4.2.3) in the form that generalizes (1.4.14) when we have a time-dependent external field

$$\frac{d}{dt}z_t = i\mathcal{L}(\phi_t)\hat{z}(z_t) \tag{4.2.4}$$

where we have introduced the Liouville operator

$$i\mathcal{L}(\phi_t) = -\frac{\partial \hat{H}}{\partial z}(z, \phi_t){\cdot}J{\cdot}\frac{\partial}{\partial z} \tag{4.2.5}$$

Note that this operator does not commute at different times

$$i\mathcal{L}(\phi_t)i\mathcal{L}(\phi_{t'}) \neq i\mathcal{L}(\phi_{t'})i\mathcal{L}(\phi_t) \tag{4.2.6}$$

The formal solution

We may introduce the time ordered exponential operator (see Appendix **??**) through the infinite series

$$\exp_+\left\{\int_{t_0}^t d\tau i\mathcal{L}(\phi_\tau)\right\} \equiv \sum_{n=0}^\infty \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^{\tau_{n-1}} d\tau_n i\mathcal{L}(\phi_{\tau_1}) \cdots i\mathcal{L}(\phi_{\tau_n}) \qquad (4.2.7)$$

For time independent fields, this evolution operator becomes simply the exponential of the Liouville operator. An aditional useful representation of the time ordered exponential is given through the product formula

$$\exp_+\left\{\int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau)\right\} = \lim_{N\to\infty} e^{i\mathcal{L}(\phi_{t_1})(t_1-\tau_{N-1})}e^{i\mathcal{L}(\phi_{\tau_{N-1}})(\tau_{N-1}-\tau_{N-2})} \cdots e^{i\mathcal{L}(\phi_{\tau_1})(\tau_1-t_0)}$$

$$(4.2.8)$$

where the time interval $[t_0, t_1]$ is divided in N sub intervals defined from the times $t_0, \tau_1, \tau_2, \cdots, \tau_{N-1}, t_1$. The time ordered exponential operator satisfies

$$\frac{d}{dt}\exp_+\left\{\int_{t_0}^t d\tau i\mathcal{L}(\phi_\tau)\right\} = i\mathcal{L}_t \exp_+\left\{\int_{t_0}^t d\tau i\mathcal{L}(\phi_\tau)\right\} \qquad (4.2.9)$$

By using the time ordered exponential operator (4.2.7) the formal solution of the time-dependent Hamilton's equations with initial condition $z_{t_0} = z_0$ is

$$z_t = \exp_+\left\{\int_{t_0}^t d\tau i\mathcal{L}(\phi_\tau)\right\}\hat{z}(z_0) \qquad (4.2.10)$$

as can be seen by taking the time derivative on both sides, leading to (4.2.4). It is obvious that the evolution of the microstate depends on the whole history of the external field ϕ_τ. Note that it is not necessary that $t > t_0$ in the above expression, although we will often assume so. We will also write the solution (4.2.10) of Hamilton's equations at time t with initial condition z_0 at time t_0, under the protocol ϕ_t in terms of the flow map

$$z_t = \hat{\Phi}_{[\phi,t_0,t]}(z_0) \qquad (4.2.11)$$

Note that in this chapter we use a more detailed notation for the flow map, because the protocol ϕ_t breaks the invariance under time translations. Therefore, we have to specify the initial t_0 and final times t, as well as the functional dependence on the external protocol.

Inverse evolution

As a consequence of the fact that Hamilton's equations are first order ODE, with unique solution given the initial microstate, we have that evolving the microstate from t_0 to an intermediate time t' and then from t' to t_1 gives the same result as evolving directly from
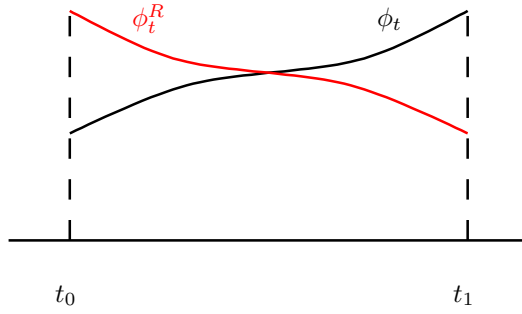
Figure 4.1: The protocol $\phi_t = \phi(t)$ and the reverse protocol $\phi_t^R = \phi(t_1 + t_0 - t)$.

t_0 to t_1, this is

$$\exp_+\left\{\int_{t_1}^{t_2} d\tau\, i\mathcal{L}(\phi_\tau)\right\}\exp_+\left\{\int_{t_0}^{t_1} d\tau\, i\mathcal{L}(\phi_\tau)\right\} = \exp_+\left\{\int_{t_0}^{t_2} d\tau\, i\mathcal{L}(\phi_\tau)\right\} \qquad (4.2.12)$$

Here $\phi_1(t)$ is the protocol from t_0 to t_1, $\phi_2(t)$ is the protocol from t_1 to t_2, $\phi_1(t) + \phi_2(t)$ is the yuxtaposition of both protocols, that leads from t_0 to t_2.

The product representation (4.2.8) for a very large but finite value of N shows that the inverse operator is given by

$$\left[\exp_+\left\{\int_{t_0}^{t_1} d\tau\, i\mathcal{L}(\phi_\tau)\right\}\right]^{-1} = e^{i\mathcal{L}(\phi_{\tau_1})(t_0-\tau_1)}\cdots e^{i\mathcal{L}(\phi_{\tau_{N-1}})(\tau_{N-2}-\tau_{N-1})}e^{i\mathcal{L}(\phi_{t_1})(\tau_{N-1}-t_1)}$$

$$= \exp_+\left\{\int_{t_1}^{t_0} d\tau\, i\mathcal{L}(\phi_\tau)\right\} \qquad (4.2.13)$$

where the order of the exponentials is reversed, as well as the sign of the argument of the exponentials. Indeed, by mutiplying (4.2.13) with (4.2.8) (either from the right or from the left), every little exponential cancel one by one. We could have arrived at (4.2.13) by choosing $t_2 = t_0$ in (4.2.12).

It is convenient to introduce the reversed protocol, see Fig 4.1

$$\phi^R(t) = \phi(t_1 + t_0 - t) \qquad (4.2.14)$$

We may express the forward evolution operator (4.2.7) in terms of the reversed protocol. By performing the change of variables $\tau_n \to \tau_n' = t_1 + t_0 - \tau_n$, in (4.2.7)

$$\exp_+\left\{\int_{t_0}^{t_1} d\tau\, i\mathcal{L}(\phi_\tau)\right\} = \sum_{n=0}^{\infty} \int_{t_0}^{t_1} d\tau_1 \cdots \int_{t_1}^{t_1+t_0-\tau_{n-1}} -d\tau_n'\, i\mathcal{L}(\phi_{\tau_1})\cdots i\mathcal{L}(\phi_{\tau_n'}^R) \quad (4.2.15)$$

By doing this change for all n gives

$$\exp_+ \left\{ \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau) \right\} = \sum_{n=0}^{\infty} (-1)^n \int_{t_1}^{t_0} d\tau_1' \cdots \int_{t_1}^{\tau_{n-1}'} d\tau_n' i\mathcal{L}(\phi_{\tau_1}^R) \cdots i\mathcal{L}(\phi_{\tau_n}^R)$$

$$= \exp_+ \left\{ - \int_{t_1}^{t_0} d\tau i\mathcal{L}(\phi_\tau^R) \right\} \tag{4.2.16}$$

On the other hand, the inverse evolution in terms of the reversed protocol is obtained from the product representation of (4.2.13) which is, with $\Delta\tau > 0$

$$\left[\exp_+ \left\{ \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau) \right\} \right]^{-1} = e^{-i\mathcal{L}(\phi_{\tau_1})\Delta\tau} e^{-i\mathcal{L}(\phi_{\tau_2})\Delta\tau} \cdots e^{-i\mathcal{L}(\phi_{\tau_N})\Delta\tau}$$

$$= e^{-i\mathcal{L}(\phi_{\tau_N}^R)\Delta\tau} e^{-i\mathcal{L}(\phi_{\tau_2}^R)\Delta\tau} \cdots e^{-i\mathcal{L}(\phi_{\tau_1}^R)\Delta\tau}$$

$$= \exp_+ \left\{ - \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau^R) \right\} \tag{4.2.17}$$

In summary, while the direct evolution can be expressed as

$$z_1 = \exp_+ \left\{ \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau) \right\} z_0 = \exp_+ \left\{ - \int_{t_1}^{t_0} d\tau i\mathcal{L}(\phi^R(\tau)) \right\} z_0 \tag{4.2.18}$$

the inverse evolution can be expressed as

$$z_0 = \left[\exp_+ \left\{ \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau) \right\} \right]^{-1} z_1 = \exp_+ \left\{ \int_{t_1}^{t_0} d\tau i\mathcal{L}(\phi_\tau) \right\} z_1$$

$$= \exp_+ \left\{ - \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau^R) \right\} z_1 \tag{4.2.19}$$

Note that within a time ordered exponential

$$\exp_+ \left\{ \int_{t_1}^{t_0} d\tau i\mathcal{L}(\phi_\tau) \right\} z_1 \neq \exp_+ \left\{ - \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau) \right\} z_1 \tag{4.2.20}$$

4.2.1   Time reversibility

We consider now the issue of time reversibility of the microscopic dynamics 1n the presence of external forcing. In the absence of external forcing we have already discussed this in Sec 1.5.3. In this section we present a slightly different derivation based on the exponential operator representation of the dynamics.

The time reversal operator $\mathcal{T}$ is introduced in (1.5.14). The time reversal operator $\mathcal{T}$ and the derivative operator $\frac{\partial}{\partial z_j}$ do not commute because

$$\frac{\partial}{\partial z_j}\mathcal{T}\hat{F}(z) = \frac{\partial}{\partial z_j}\hat{F}(\epsilon{\cdot}z) = \hat{F}_k(\epsilon{\cdot}z)\epsilon_{kj} = \epsilon_{jk}\mathcal{T}\hat{F}_k(z) = \epsilon_{jk}\mathcal{T}\frac{\partial}{\partial z_k}\hat{F}(z) \qquad (4.2.21)$$

where we follow the notation $\hat{F}_k(\epsilon z)$ for the partial derivative of $\hat{F}(z)$ with respect to z_k evaluated at ϵz. Because $\hat{F}(z)$ is arbitrary, we have

$$\frac{\partial}{\partial z_j}\mathcal{T} = \epsilon_{jk}\mathcal{T}\frac{\partial}{\partial z_k} \qquad (4.2.22)$$

where sum over repeated indices is implied.

The unperturbed Hamiltonian is invariant under the time reversal operator $\hat{H}^0(z) = \hat{H}^0(\epsilon z) = \mathcal{T}\hat{H}^0(z)$. We will assume that the coupling phase functions $\hat{A}(z)$ have a definite parity, this is $\hat{A}(\epsilon{\cdot}z) = \varepsilon{\cdot}\hat{A}(z)$ where ε is a diagonal matrix of equal dimension as the number of fields ϕ_t and such that the diagonal elements are ± 1 according to the parity of each component of $\hat{A}(z)$. The full Hamiltonian then satisfies

$$\hat{H}(\epsilon{\cdot}z, \varepsilon{\cdot}\phi_t) = \hat{H}(z, \phi_t) \qquad (4.2.23)$$

The time reversal behaviour of the derivatives of the Hamiltonian are easily obtained by applying the operator $\frac{\partial}{\partial z_i}$ to both sides of (4.2.23). We have

$$\frac{\partial}{\partial z_i}\hat{H}(\epsilon{\cdot}z, \varepsilon{\cdot}\phi_t) = \frac{\partial}{\partial z_i}\hat{H}(z, \phi_t)$$
$$\epsilon_{ik}\hat{H}_k(\epsilon{\cdot}z, \varepsilon{\cdot}\phi_t) = \hat{H}_i(z, \phi_t) \qquad (4.2.24)$$

We may express (4.2.23) and (4.3.38) in terms of the reversal operator. On one hand, from (4.2.23) we have that the Hamiltonian transforms as

$$\mathcal{T}\hat{H}(z, \phi_t) = \hat{H}(z, \varepsilon{\cdot}\phi_t) \qquad (4.2.25)$$

The action of the time reversal operator on the Hamiltonian is just to time reverse the external field. On the other hand, the derivatives of the Hamiltonian, transform according to

$$\mathcal{T}H_i(z, \phi_t) = \epsilon_{ij}H_j(z, \varepsilon{\cdot}\phi_t) \qquad (4.2.26)$$

We now consider the effect of time reversal on the Liouville operator. By applying

the Liouville operator on a time reversed function we have

$$i\mathcal{L}(\phi_t)\mathcal{T}f(z) = -\hat{H}_i(z,\phi_t)J_{ij}\frac{\partial}{\partial z_j}f(\epsilon{\cdot}z) = -\hat{H}_i(z,\phi_t)J_{ij}\epsilon_{jk}f_k(\epsilon{\cdot}z) = \hat{H}_i(z,\phi_t)\epsilon_{ij}J_{jk}f_k(\epsilon{\cdot}z)$$

$$(4.2.27)$$

In the last equality in (4.2.27) we have used the anti-commutativity (1.5.19). By inserting (4.3.38) into (4.2.27) gives

$$i\mathcal{L}(\phi_t)\mathcal{T}f(z) = \hat{H}_j(\epsilon{\cdot}z,\varepsilon{\cdot}\phi_t)J_{jk}f_k(\epsilon{\cdot}z) = -\mathcal{T}i\mathcal{L}(z,\varepsilon{\cdot}\phi_t)f(z) \qquad (4.2.28)$$

Because the function $f(z)$ is arbitrary,

$$i\mathcal{L}(\phi_t)\mathcal{T} = -\mathcal{T}i\mathcal{L}(\varepsilon{\cdot}\phi_t) \qquad (4.2.29)$$

This shows that the Liouville operator and the time reversal operator anticommute, provided that we change accordingly the parity of the protocol.

With this result we can now see what is the effect of evolving a time reversed initial state. We proceed term by term in the evolution operator. Consider the first term of the time ordered exponential, when applied to an arbitrary function $\hat{F}(z)$

$$\int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau)\hat{F}(\epsilon{\cdot}z) = \int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau)\mathcal{T}\hat{F}(z) = -\mathcal{T}\int_{t_0}^{t_1} d\tau i\mathcal{L}(\varepsilon{\cdot}\phi_\tau)\hat{F}(z) \qquad (4.2.30)$$

where we have used the anticonmutativity property (4.2.29). Next, consider the second order term

$$\int_{t_0}^{t_1} d\tau_1 i\mathcal{L}(\phi_{\tau_1})\int_{t_0}^{\tau_1} d\tau_2 i\mathcal{L}(\phi_{\tau_2})\hat{F}(\epsilon{\cdot}z) = -\int_{t_0}^{t_1} d\tau_1 i\mathcal{L}(\phi_{\tau_1})\mathcal{T}\int_{t_0}^{\tau_1} d\tau_2 i\mathcal{L}(\varepsilon{\cdot}\phi_{\tau_2})\hat{F}(z)$$

$$= \mathcal{T}\int_{t_0}^{t_1} d\tau_1 i\mathcal{L}(\varepsilon{\cdot}\phi_{\tau_1})\int_{t_0}^{\tau_1} d\tau_2 i\mathcal{L}(\varepsilon{\cdot}\phi_{\tau_2})\hat{F}(z)$$

$$(4.2.31)$$

By induction we obtain the evolution of a time reversed phase function

$$\exp_+\left\{\int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau)\right\}\mathcal{T}\hat{F}(z) = \mathcal{T}\exp_+\left\{-\int_{t_0}^{t_1} d\tau i\mathcal{L}(\varepsilon{\cdot}\phi_\tau)\right\}\hat{F}(z) \qquad (4.2.32)$$

This is the fundamental expression that conveys, at the level of evolution operators, the notion of time reversibility of Hamilton's equations.

By applying this expression to $\hat{F} = \hat{z}$ gives

$$\exp_+\left\{\int_{t_0}^{t_1} d\tau i\mathcal{L}(\phi_\tau)\right\}\epsilon{\cdot}\hat{z} = \mathcal{T}\exp_+\left\{-\int_{t_0}^{t_1} d\tau i\mathcal{L}(\varepsilon{\cdot}\phi_\tau)\right\}\hat{z} \qquad (4.2.33)$$

This property of time reversibility may be more transparent if we introduce the flow map $\Phi_{[\phi,t_0,t_1]}(z)$ that gives the evolution of the microstate z from time t_0 to t_1 under the
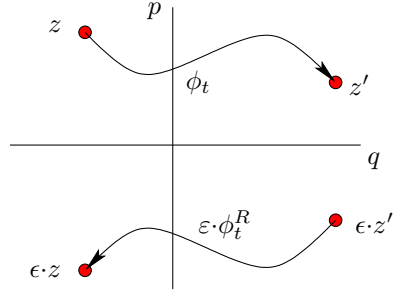
Figure 4.2: Consider a microstate z that evolves forward under the protocol ϕ_t starting at t_0 and reaching at t_1 the microstate $z' = \hat{\Phi}_{[\phi,t_0,t_1]}(z)$. If we change the sign of the momenta to produce $\epsilon \cdot z'$ and evolve this microstate *forward* in time from t_0 to t_1 under the reversed protocol $\varepsilon \cdot \phi_t^R$ we will arrive, according to microscopic reversibility, at $\hat{\Phi}_{[\varepsilon \cdot \phi_t^R,t_0,t_1]}(\epsilon \cdot z') = \epsilon \cdot z$. If we record on a film the evolution from $z \to z'$ (the upper trajectory) and run the film backwards, we will observe the lower trajectory from $\epsilon \cdot z' \to \epsilon \cdot z$.

protocol $\phi(t)$. By definition it is given by

$$\Phi_{[\phi,t_0,t_1]}(z) = \exp_+ \left\{ \int_{t_0}^{t_1} d\tau i \mathcal{L}(\phi_\tau) \right\} \hat{z}(z) \tag{4.2.34}$$

Now, consider the inverse evolution (4.2.35) that we quote again in the form

$$\left[\exp_+ \left\{ \int_{t_0}^{t_1} d\tau i \mathcal{L}(\phi_\tau) \right\} \right]^{-1} = \exp_+ \left\{ \int_{t_1}^{t_0} d\tau i \mathcal{L}(\phi_\tau) \right\}$$

$$= \exp_+ \left\{ - \int_{t_0}^{t_1} d\tau i \mathcal{L}(\phi_\tau^R) \right\} \tag{4.2.35}$$

$$\exp_+ \left\{ - \int_{t_0}^{t_1} d\tau i \mathcal{L}(\varepsilon \cdot \phi_\tau) \right\} = \left[\exp_+ \left\{ \int_{t_0}^{t_1} d\tau i \mathcal{L}(\varepsilon \cdot \phi_\tau^R) \right\} \right]^{-1} \tag{4.2.36}$$

Therefore, the reversibility condition (4.2.33) in terms of the flow map is

$$\epsilon \cdot \Phi_{[\phi,t_0,t_1]}(z) = \Phi^{-1}_{[\varepsilon \cdot \phi^R,t_0,t_1]}(\epsilon \cdot z) \tag{4.2.37}$$

In an equivalent way this is

$$z' = \Phi_{[\phi,t_0,t_1]}(z)$$
$$z = \epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R,t_0,t_1]}(\epsilon \cdot z') \tag{4.2.38}$$

In order to shorten the notation, we introduce the forward and backward trajectories

$$z^F(z) = \Phi_{[\phi,t_0,t_1]}(z)$$
$$z^B(z) = \epsilon{\cdot}\Phi_{[\epsilon\cdot\phi^R,t_0,t_1]}(\epsilon{\cdot}z) \qquad (4.2.39)$$

Whenever an F or B superscript appear in a microstate, we should really understand the two flow maps (4.2.39) from t_0 to t_1 under the corresponding external forcings. In terms of these two trajectories, the microscopic reversibility property states that

$$z^B(z^F(z)) = z \qquad (4.2.40)$$

In Fig. 4.2 we explain in words the meaning of this identity expressing microscopic reversibility in the presence of an external forcing. However simple the statement (4.2.40) may look, the full blown up statement of microscopic reversibility is given by (4.2.37).

4.2.2 Liouville's equation with external fields

We show in this section that the structure of the Liouville's equation is left affected by the presence of external fields. Recall that the initial conditions are always subject to uncertainty, which is described with a probability distribution $\rho_0(z)$ over initial microstates. This uncertainty about initial conditions is propagated with the deterministic Hamilton's equations and produce a probability density $\rho_t(z)$ of obtaining the system in the vicinity of z at time t. The Liouville's equation describes how $\rho_t(z)$ is related to $\rho_0(z)$.

We first note that the flow in phase space is incompressible, even in the case of external forcing, as is easily seen from (4.2.3)

$$\frac{\partial}{\partial z_i}\hat{v}_i(z) = J_{ij}\frac{\partial^2 \hat{H}}{\partial z_i \partial z_j}(z, \phi_t) \qquad (4.2.41)$$

and the fact that the trace of the product of an antisymmetric and a symmetric matrix vanishes. The incompressibility of the Hamiltonian flow field implies that the Jacobian of the transformation of coordinates from z to $\hat{\Phi}_{[\phi,0,t]}(z)$ is always one.

Let M be a region of not vanishing measure in phase space and $M' = \hat{\Phi}_{[\phi,t_0,t_1]}(M)$ the region resulting from the evolution of each point of M according to Hamilton's equations. It is obvious that the probability that the system is in the region M at $t = t_0$ is *identical* to the probability of being in M' at $t = t_1$. For this reason,

$$\int_M \rho_{t_0}(z)dz = \int_{M'} \rho_{t_1}(z')dz'. \qquad (4.2.42)$$

We now perform a change of variables $z' = \hat{\Phi}_{[\phi,t_0,t_1]}(z)$, that has unit Jacobian. The right hand side of (4.2.42) becomes

$$\int_{M'} \rho_{t_1}(z')dz' = \int_M \rho_{t_1}(\hat{\Phi}_{[\phi,t_0,t_1]}(z))dz \qquad (4.2.43)$$

This is true for any region M and, therefore, the integrand of the right hand side of (4.2.42) and the left hand side of (4.2.43) should be equal, i.e.

$$\rho_{t_0}(z) = \rho_{t_1}\left(\hat{\Phi}_{[\phi,t_0,t_1]}(z)\right),\tag{4.2.44}$$

or, by a simple renaming of variables,

$$\rho_{t_1}(z) = \rho_{t_0}\left(\hat{\Phi}_{[\phi,t_0,t_1]}^{-1}(z)\right)\tag{4.2.45}$$

This is the way the density in phase space evolves in time. It simply says that the probability density of microstate z at the final time is the same as the one that had the initial condition of z at the initial time.

In summary, the Liouville theorem has the same form for external forcing as without forcing. This is, of course, a consequence that the flow is incompressible even with external forcing, as seen in (4.2.41). The Liouville's equation (1.6.7) becomes

$$\partial_t \rho_t(z) = -i\mathcal{L}(\phi_t)\rho_t(z)\tag{4.2.46}$$

We will always assume that the unperturbed Hamiltonian $H_0(z)$ has a well defined equilibrium state described with an equilibrium ensemble which is just a function of $H_0(z)$ (and any other dynamic invariant) and, hence, a stationary solution of the unperturbed Liouville equation. However, in the presence of time-dependent external forcing, the Hamiltonian $\hat{H}(z, \phi_t)$ no longer generates a stationary probability density solution of the Liouville's equation in general. Under time-dependent external forcing there is no equilibrium ensemble.

4.3   The CG dynamics under external forcing

In this section, we derive the exact equation for the evolution of the probability $P(a,t)$ of the CG variables when the system is subject to an external forcing. The probability distribution $P(a,t)$ that the phase functions $\hat{A}(z)$ take particular values a is defined by

$$P(a,t) = \int dz \rho_t(z) \hat{\Psi}_a(z) = \mathrm{Tr}[\rho_t \hat{\Psi}_a] \qquad (4.3.1)$$

where we have introduced the notation for the M-dimensional Dirac delta function

$$\hat{\Psi}_a(z) = \delta(\hat{A}(z) - a) \qquad (4.3.2)$$

This notation emphasizes that the probability $P(a,t)$ is just the average of the distribution (4.3.2). We also used in (4.3.1) the "trace" notation to denote an integration over phase space. If we take the time derivative of the probability distribution we obtain

$$\partial_t P(a,t) = \mathrm{Tr}\left[\hat{\Psi}_a \partial_t \rho_t\right] = \mathrm{Tr}\left[\rho_t i \mathcal{L}(\phi_t) \hat{\Psi}_a\right] \qquad (4.3.3)$$

where we have used the Liouville's equation (4.2.46) and an integration by parts. Eq. (4.4.10) is not closed because the right hand side is not expressed in terms of the $P(a,t)$.

In order to obtain a closed equation for the evolution of the probability $P(a,t)$, one introduces the relevant ensemble

$$\overline{\rho}_t(z) = \frac{P(\hat{A}(z),t)}{\Omega(\hat{A}(z))} \qquad (4.3.4)$$

where

$$\Omega(a) = \int dz \delta(\hat{A}(z) - a) \qquad (4.3.5)$$

is the measure of the submanifold of phase space defined implicitly through the equation $\hat{A}(z) = a$.

Note that the initial ensemble $\rho_0(z)$ that captures the statistical knowledge about the initial microstate needs to be specified in Statistical Mechanics. We assume that at the initial time all what we know is the macroscopic information, which is characterized by the probability of the CG variables at the initial time $P(a,0)$. By using the Principle of Maximum Entropy again, the least biased ensemble which is consistent with this macroscopic information is given by

$$\rho_0(z) = \frac{P(\hat{A}(z),0)}{\Omega(\hat{A}(z))} \qquad (4.3.6)$$

This assumption implies that the initial ensemble is of the relevant type (4.3.4).

4.3.1   An exact evolution equation

Of course, the CG variables are slow but are not strict constants of the motion and the real ensemble $\rho_t(z)$ is different from the relevant ensemble $\overline{\rho}_t(z)$. Nevertheless, a perturbative scheme suggests itself in which one considers

$$\rho_t(z) = \overline{\rho}_t(z) + \delta\rho_t(z) \tag{4.3.7}$$

where $\delta\rho_t(z)$ is the so called irrelevant part of the ensemble. The objective is now to express the irrelevant part $\delta\rho_t(z)$ in terms of the relevant part $\overline{\rho}_t(z)$ and, eventually, to obtain a closed equation for the relevant ensemble $\overline{\rho}_t(z)$. As the relevant ensemble is determined by the CG probability distribution $P(a,t)$, one will end up with a closed equation for $P(a,t)$ itself.

As a first step, note that the relevant ensemble can be obtained from the action of a projection operator $\mathcal{P}$ on the real ensemble, this is

$$\overline{\rho}_t(z) = \mathcal{P}\rho_t(z) \tag{4.3.8}$$

where the time-independent projector operator is defined by

$$\mathcal{P}\cdots = \int dz' \frac{\delta(\hat{A}(z) - \hat{A}(z'))}{\Omega(\hat{A}(z))}\cdots \tag{4.3.9}$$

This operator is self-adjoint in the sense that

$$\int dz \hat{F}(z)\mathcal{P}\hat{G}(z) = \int dz \hat{G}(z)\mathcal{P}\hat{F}(z) \tag{4.3.10}$$

for two arbitrary phase functions $\hat{F}(z), \hat{G}(z)$. We introduce also the complementary operator $\mathcal{Q} = 1 - \mathcal{P}$ that gives the irrelevant part of the ensemble

$$\delta\rho_t(z) = \mathcal{Q}\rho_t(z) \tag{4.3.11}$$

We can now follow the standard derivation in which one obtains two equations, one for the time derivatives of the relevant part

$$\begin{aligned}
\partial_t \overline{\rho}_t(z) = \partial_t \mathcal{P}\rho_t &= -\mathcal{P}i\mathcal{L}(\phi_t)\rho_t \\
&= -\mathcal{P}i\mathcal{L}(\phi_t)(\mathcal{P} + \mathcal{Q})\rho_t \\
&= -\mathcal{P}i\mathcal{L}(\phi_t)\overline{\rho}_t - \mathcal{P}i\mathcal{L}(\phi_t)\delta\rho_t
\end{aligned} \tag{4.3.12}$$

and a similar equation for the irrelevant part

$$\begin{aligned}
\partial_t \delta\rho_t(z) = \partial_t \mathcal{Q}\rho_t &= -\mathcal{Q}i\mathcal{L}(\phi_t)\rho_t \\
&= -\mathcal{Q}i\mathcal{L}(\phi_t)(\mathcal{P} + \mathcal{Q})\rho_t \\
&= -\mathcal{Q}i\mathcal{L}(\phi_t)\overline{\rho}_t - \mathcal{Q}i\mathcal{L}(\phi_t)\delta\rho_t
\end{aligned} \tag{4.3.13}$$

The formal solution of the differential equation (4.3.13) is given by

$$\delta\rho_t(z) = \exp_+\left\{-\int_0^t d\tau \, \mathcal{Q}i\mathcal{L}(\phi_\tau)\right\}\delta\rho_0(z) - \int_0^t dt' \exp_+\left\{-\int_{t'}^t d\tau \, \mathcal{Q}i\mathcal{L}(\phi_\tau)\right\}\mathcal{Q}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}$$

(4.3.14)

The assumption that the initial ensemble $\rho_0(z)$ is of the relevant form as in Eq. (4.3.6), implies in Eq. (4.3.7) that $\delta\rho_0(z) = 0$. Therefore, Eq. (4.3.14) becomes

$$\delta\rho_t(z) = -\int_0^t dt' \, \mathcal{U}_{[\phi,t,t']}\mathcal{Q}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}$$

(4.3.15)

where we have introduced the notation

$$\mathcal{U}_{[\phi,t,t']} = \exp_+\left\{-\int_{t'}^t d\tau \, \mathcal{Q}i\mathcal{L}(\phi_\tau)\right\}$$

(4.3.16)

for the projected evolution operator. This operator is a functional of the external field depending on all its history. Note that this operator satisfies

$$\mathcal{U}_{[\phi,t,t']} = \mathcal{Q}\mathcal{U}_{[\phi,t,t']}$$

(4.3.17)

as can be seen from the series expansion of the time ordered exponential.

By using (4.3.15) in (4.3.7) we have

$$\rho_t(z) = \overline{\rho}_t(z) - \int_0^t dt' \, \mathcal{U}_{[\phi,t,t']}\mathcal{Q}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}(z)$$

(4.3.18)

This remarkable identity shows that, when the initial ensemble is of the relevant form, the real ensemble $\rho_t(z)$ can be fully expressed in terms of the relevant ensemble $\overline{\rho}_t(z)$ at present and past times. We now use the decomposition (4.4.9) in (4.4.10)

$$\partial_t P(a,t) = \underbrace{\text{Tr}\left[\overline{\rho}_t(z)i\mathcal{L}(\phi_t)\hat{\Psi}_a\right]}_{(1)} - \underbrace{\int_0^t dt' \text{Tr}\left[\left(i\mathcal{L}(\phi_t)\hat{\Psi}_a\right)\mathcal{U}_{[\phi,t,t']}\mathcal{Q}_{t'}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}(z)\right]}_{(2)}$$

(4.3.19)

In this expression we will use two results. The first one gives the action of the Liouville operator on the distribution $\hat{\Psi}_a(z)$ which, after using the chain rule, is given by

$$i\mathcal{L}(\phi_t)\hat{\Psi}_a(z) = -\frac{\partial\hat{\Psi}_a}{\partial a_\mu}(z)i\mathcal{L}(\phi_t)\hat{A}_\mu(z)$$

(4.3.20)

The second result is the relevant ensemble (4.3.4) expressed in the form

$$\overline{\rho}_t(z) = \int da \frac{P(a,t)}{\Omega(a)}\hat{\Psi}_a(z)$$

(4.3.21)

The reversible drift

By using Eqs. (4.4.12), (4.4.13), the first term in (4.4.11) becomes

$$(1) = -\frac{\partial}{\partial a_\mu} V_\mu(a, \phi_t) P(a, t) \tag{4.3.22}$$

where the reversible drift is defined as

$$V_\mu(a, \phi_t) \equiv \left\langle i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a \tag{4.3.23}$$

Here, the conditional expectation is defined as

$$\langle \cdots \rangle^a \equiv \frac{1}{\Omega(a)} \mathrm{Tr}[\hat{\Psi}_a \cdots] \tag{4.3.24}$$

The reversible drift (4.3.23) is the conditional expectation of the time derivative of the CG variables.

The reversible drift admits a nice form when we assume that the time-dependent Hamiltonian can be expressed in terms of the selected CG variables, this is

$$\hat{H}_t(z) = E^0(\hat{A}(z)) + E^1(\hat{A}(z), \phi_t) = E(\hat{A}, \phi_t) \tag{4.3.25}$$

where the total CG energy function

$$E(a, t) = E^0(a) + E^1(a, \phi_t) \tag{4.3.26}$$

is the sum of the energy function $E^0(a)$ of the unperturbed system and the energy function $E^1(a, \phi_t)$ describing the action of the time-dependent external forcing. Note that it is always possible to write the Hamiltonian in the form (4.4.2) because, if not, we may include the unperturbed Hamiltonian $\hat{H}^0(z)$ in the list of CG variables. Note that in many cases of interest the perturbation is linear in the CG variables, this is $E^1(a, t) = -a \cdot \phi_t$. We will refer to ϕ_t or to $E^1(a, \phi_t)$ itself as *the protocol*.

One consequence of (4.4.2) is that the gradient of the Hamiltonian in phase space is given from the chain rule as

$$\frac{\partial \hat{H}_t}{\partial z}(z) = \frac{\partial E}{\partial a_\mu}(\hat{A}(z), \phi_t)\frac{\partial \hat{A}_\mu}{\partial z}(z) \tag{4.3.27}$$

This implies that the action of the Liouville operator (4.5.3) on the CG variables is

$$\begin{aligned}
i\mathcal{L}(\phi_t)\hat{A}_\mu(z) &= \frac{\partial \hat{A}_\mu}{\partial z}(z) \cdot J \cdot \frac{\partial H_t}{\partial z}(z) \\
&= \frac{\partial \hat{A}_\mu}{\partial z}(z) \cdot J \cdot \frac{\partial \hat{A}_\nu}{\partial z}(z) \frac{\partial E}{\partial a_\nu}(\hat{A}(z), \phi_t) \\
&= \{\hat{A}_\mu, \hat{A}_\nu\}\frac{\partial E}{\partial a_\nu}(\hat{A}(z), \phi_t)
\end{aligned} \tag{4.3.28}$$

where we have introduced the Poisson bracket of the CG variables

$$\{\hat{A}_\mu, \hat{A}_\nu\} = \frac{\partial \hat{A}_\mu}{\partial z}(z) \cdot J \cdot \frac{\partial \hat{A}_\nu}{\partial z}(z) \tag{4.3.29}$$

By using (4.3.28) in the reversible drift (4.3.23), and the fact that within the conditional expectation we may substitute $\hat{A}(z)$ with a, we get the following expression for the reversible drift

$$V_\mu(a, \phi_t) = L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a, \phi_t) \tag{4.3.30}$$

where the reversible operator defined as

$$L_{\mu\nu}(a) \equiv \left\langle \{\hat{A}_\mu, \hat{A}_\nu\} \right\rangle^a \tag{4.3.31}$$

is the conditional expectation of the Poisson bracket of the CG variables.

The memory term

We now consider the second term in Eq. (4.4.11). By using (4.4.13), and the properties (4.3.17), (4.3.10), the second term in (4.4.11) becomes

$$(2) = -\int_0^t dt' \text{Tr}\left[(i\mathcal{L}(\phi_t)\hat{\Psi}_a)\mathcal{U}_{[\phi,t,t']}\mathcal{Q}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}\right]$$

$$= -\int_0^t dt' \int da' \frac{P(a', t')}{\Omega(a')}\text{Tr}\left[(\mathcal{Q}i\mathcal{L}(\phi_t)\Psi_a)\mathcal{U}_{[\phi,t,t']}\left(\mathcal{Q}i\mathcal{L}(\phi_{t'})\hat{\Psi}_{a'}\right)\right] \tag{4.3.32}$$

Consider the term

$$\mathcal{Q}i\mathcal{L}(\phi_t)\Psi_a = -\mathcal{Q}\frac{\partial\hat{\Psi}_a}{\partial a_\mu}(z)i\mathcal{L}(\phi_t)\hat{A}_\mu(z)$$

$$= -\frac{\partial\hat{\Psi}_a}{\partial a_\mu}(z)\left[i\mathcal{L}(\phi_t)\hat{A}_\mu(z) - V_\mu(\hat{A}(z), t)\right]$$

$$= -\frac{\partial\hat{\Psi}_a}{\partial a_\mu}(z)\left[i\mathcal{L}(\phi_t)\hat{A}_\mu(z) - V_\mu(a, \phi_t)\right] \tag{4.3.33}$$

where we have used (4.4.12), the explicit form of the projection operator (4.3.9), and the definition (4.3.23) of the reversible drift. By using (4.3.33) into (4.3.32) we obtain

$$(2) = -\int_0^t dt' \int da' \frac{P(a', t')}{\Omega(a')}\frac{\partial}{\partial a_\mu}\frac{\partial}{\partial a'_\nu}\text{Tr}\left[\hat{\Psi}_a\left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t)\right]\mathcal{U}_{[\phi,t,t']}\hat{\Psi}_{a'}\left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'})\right]\right]$$

$$= \int_0^t dt' \int da' \frac{\partial}{\partial a_\mu}\Omega(a)\left\langle\left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t)\right]\mathcal{U}_{[\phi,t,t']}\hat{\Psi}_{a'}\left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'})\right]\right\rangle^a \frac{\partial}{\partial a'_\nu}\frac{P(a', t')}{\Omega(a')}$$

$$= \frac{\partial}{\partial a_\mu}\int_0^t dt' \int da' \Omega(a)D_{\mu\nu}(\phi, a, a', t, t')\frac{\partial}{\partial a'_\nu}\frac{P(a', t')}{\Omega(a)} \tag{4.3.34}$$

where we have defined the memory kernel

$$D_{\mu\nu}(\phi, a, a', t, t') \equiv \left\langle \left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t) \right] \mathcal{U}_{[\phi, t, t']} \hat{\Psi}_{a'} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \right\rangle^a$$

$$(4.3.35)$$

This memory kernel is a functional of the whole history of the external forcing ϕ_t.

Eqs. (4.4.16) and (4.3.34) in (4.4.11) give the generalized FPE in the presence of external forcing

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} L_{\mu\nu}(a) \frac{\partial E}{\partial a_\nu}(a, \phi_t) P(a, t)$$

$$+ \frac{\partial}{\partial a_\mu} \int_0^t dt' \int da' \Omega(a) D_{\mu\nu}(\phi, a, a', t, t') \frac{\partial}{\partial a'_\nu} \frac{P(a', t')}{\Omega(a)} \qquad (4.3.36)$$

This is an exact closed equation for the probability $P(a, t)$ of the CG variables. It is a complicated integro-differential equation which is not very useful because the memory kernel $D_{\mu\nu}(\phi, a, a', t, t')$ defined in (4.4.19) is a quite formal object.

4.3.2 Time reversal behaviour of the different terms of the exact equation

We consider in this section the time-reversal properties of the diferent objects appearing in the exact Eq. (4.4.20). We will assume that the CG variables have a well-defined parity under time reversal, this is

$$\hat{A}(\epsilon \cdot z) = \varepsilon \cdot \hat{A}(z) \tag{4.3.37}$$

were ε is a diagonal matrix with ± 1 in the diagonal, according to the parity of the CG variable. The Hamiltonian (4.2.1) has the parity (4.2.23)

$$\hat{H}(\epsilon \cdot z, \varepsilon \cdot \phi_t) = \hat{H}(z, \phi_t) \tag{4.3.38}$$

Consider the measure $\Omega(a)$ and evaluate

$$\Omega(\varepsilon \cdot a) = \int dz \delta \left(\hat{A}(z) - \varepsilon \cdot a \right) = \int dz \delta \left(\varepsilon \cdot \hat{A}(z) - a \right) = \int dz \delta \left(\hat{A}(\epsilon \cdot z) - a \right) = \Omega(a) \tag{4.3.39}$$

where in the last equation we have performed a change of variable $z' = \epsilon \cdot z$ with unit Jacobian. Note that the conditional expectations transform according to

$$\left\langle \hat{F} \right\rangle^{\varepsilon \cdot a} = \frac{1}{\Omega(\varepsilon \cdot a)} \int dz \delta \left(\hat{A}(z) - \varepsilon \cdot a \right) \hat{F}(z) = \frac{1}{\Omega(a)} \int dz \delta \left(\hat{A}(\epsilon \cdot z) - a \right) \hat{F}(z) = \left\langle \mathcal{T}\hat{F} \right\rangle^{a} \tag{4.3.40}$$

The reversible matrix $L_{\mu\nu}(a)$ transforms according to

$$L(\varepsilon \cdot a) = \left\langle \mathcal{T} \left\{ \hat{A}, \hat{A} \right\} \right\rangle^{a} = -\varepsilon \cdot L(a) \cdot \varepsilon \tag{4.3.41}$$

where we have made use of the anticommutativity of ϵ and the symplectic matrix J appearing in the Poisson bracket.

The drift term satisfies

$$V(\varepsilon \cdot a, \varepsilon \cdot \phi_t) = \left\langle i\mathcal{L}(\varepsilon \cdot \phi_t)\hat{A} \right\rangle^{\varepsilon \cdot a} = \left\langle \mathcal{T} i\mathcal{L}(\varepsilon \cdot \phi_t)\hat{A} \right\rangle^{a} = -\left\langle i\mathcal{L}(\phi_t)\mathcal{T}\hat{A} \right\rangle^{a} = -\varepsilon \cdot V(a, \phi_t) \tag{4.3.42}$$

Consider the memory kernel (4.4.19), evaluated at the time-reversed variables

$$D_{\mu\nu}(\varepsilon\cdot\phi, \varepsilon\cdot a, \varepsilon\cdot a', t, t')$$

$$= \left\langle \left[i\mathcal{L}(\varepsilon\cdot\phi_t)\hat{A}_\mu - V_\mu(\varepsilon\cdot a, \varepsilon\cdot\phi_t) \right] \mathcal{U}_{[\varepsilon\cdot\phi, t, t']} \hat{\Psi}_{\varepsilon\cdot a'} \left[i\mathcal{L}(\varepsilon\cdot\phi_{t'})\hat{A}_\nu - V_\nu(\varepsilon\cdot a', \varepsilon\cdot\phi_{t'}) \right] \right\rangle^{\varepsilon\cdot a}$$

$$= \varepsilon_\mu \varepsilon_\nu \left\langle \mathcal{T} \left(\left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t) \right] \mathcal{U}_{[\varepsilon\cdot\phi, t, t']} \left(\mathcal{T}\hat{\Psi}_{a'} \right) \mathcal{T} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \right) \right\rangle^a$$

$$= \varepsilon_\mu \varepsilon_\nu \left\langle \left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t) \right] \mathcal{T}\mathcal{U}_{[\varepsilon\cdot\phi, t, t']}\mathcal{T}\hat{\Psi}_{a'} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \right\rangle^a \quad (4.3.43)$$

Let us consider the action of the time-reversible operator $\mathcal{T}$ on the projected dynamics operator $\mathcal{U}_{[\varepsilon\cdot\phi, t, t']}$. From its definition (4.3.16) we have

$$\mathcal{T}\mathcal{U}_{[\varepsilon\cdot\phi, t, t']} = \mathcal{T} \exp_+ \left\{ - \int_{t'}^t d\tau \, \mathcal{Q}i\mathcal{L}(\varepsilon\cdot\phi_\tau) \right\}$$

$$= \mathcal{T} \sum_{n=0}^\infty (-1)^n \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^{\tau_{n-1}} d\tau_n \, \mathcal{Q}i\mathcal{L}(\varepsilon\cdot\phi_{\tau_1}) \cdots \mathcal{Q}i\mathcal{L}(\varepsilon\cdot\phi_{\tau_n})$$

$$= \sum_{n=0}^\infty \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^{\tau_{n-1}} d\tau_n \, \mathcal{Q}i\mathcal{L}(\phi_{\tau_1}) \cdots \mathcal{Q}i\mathcal{L}(\phi_{\tau_n})\mathcal{T} \quad (4.3.44)$$

where we have made use that the projector $\mathcal{P}$ (and hence $\mathcal{Q}$) defined in (4.3.9) and the time-reversal operator commute $\mathcal{P}\mathcal{T} = \mathcal{T}\mathcal{P}$. Therefore, we have that

$$\mathcal{T}\mathcal{U}_{[\varepsilon\cdot\phi, t, t']} = \mathcal{U}_{[\phi, t, t']}^{-1}\mathcal{T} \quad (4.3.45)$$

[Check!] We arrive finally at the following time-reversal symmetry of the memory kernel

$$D_{\mu\nu}(\varepsilon\cdot\phi, \varepsilon\cdot a, \varepsilon\cdot a', t, t')\Omega(a') = \varepsilon_\mu \varepsilon_\nu D_{\nu\mu}(\phi, a', a, t', t)\Omega(a) \quad (4.3.46)$$

[Finish the proof!]

4.3.3   The Markovian approximation

Note that the memory kernel (4.4.19) involves the conditional correlation of the fluctuations of the time derivatives of the CG variables. Under the assumption that the time derivatives decay in a time scale much shorter than the overall evolution of the CG variables themselves, we may take the so called Markovian approximation. This assumption allows one to approximate

$$\mathcal{U}_{[\phi, t, t']}\hat{\Psi}_{a'} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \simeq \hat{\Psi}_{a'}\mathcal{U}_{[\phi, t, t']} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \quad (4.3.47)$$

this is, the CG variables hardly evolve in the time scale of the memory and, therefore, the Dirac delta function $\hat{\Psi}_{a'}$ is unaffected by the evolution operator. Under this approximation, and because $\hat{\Psi}_a(z)\hat{\Psi}_{a'}(z) = \hat{\Psi}_a(z)\delta(a - a')$, the memory kernel (4.4.19) becomes

diagonal in the space of CG variables

$$
D_{\mu\nu}(\phi, a, a', t, t') \simeq \delta(a - a') \left\langle \left[i\mathcal{L}(\phi_t)\hat{A}_\mu - V_\mu(a, \phi_t) \right] \mathcal{U}_{[\phi, t, t']} \left[i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a', \phi_{t'}) \right] \right\rangle^a
$$
(4.3.48)

Slow external forcing

Note that the dissipative kernel is a functional of the whole history of the external forcing ϕ_t. Some simplification occurs in the dissipative matrix when we assume that the external forcing is slow. Slow forcing means that the time scale of variation of the external forcing ϕ_t is much larger than the time scale of the memory kernel. In this case, we may take the approximation $i\mathcal{L}(\phi_\tau) \simeq i\mathcal{L}(\phi_t)$ inside the time-ordered exponential (4.2.7)

$$
\begin{aligned}
\mathcal{U}_{[\phi, t, t']} &= \exp_+ \left\{ - \int_{t'}^t d\tau \, \mathcal{Q} i\mathcal{L}(\phi_\tau) \right\} \\
&\simeq \sum_{n=0}^\infty (-1)^n \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^{\tau_{n-1}} d\tau_n \, \mathcal{Q} i\mathcal{L}(\phi_t) \cdots \mathcal{Q} i\mathcal{L}(\phi_t) \\
&= \exp\{ -\mathcal{Q} i\mathcal{L}(\phi_t)(t - t') \}
\end{aligned}
$$
(4.3.49)

In this way, by approximating the projected dynamics with the unprojected dynamics we have

$$
\mathcal{U}^\dagger(t, t') = \exp_+ \left\{ -i\mathcal{L}(\phi_t)(t - t') \right\}
$$
(4.3.50)

This is the evolution operator of the system under a *constant in time* external forcing in whih the particular value of the constant external field is fixed to ϕ_t. In this adiabatic approximation the dissipative matrix becomes

$$
D_{\mu\nu}(a, \phi_t) = \int_0^{\Delta t} dt' \left\langle \delta_t \hat{A}_\nu \exp\{ i\mathcal{L}(\phi_t)t' \} \delta_t \hat{A}_\mu \right\rangle^a
$$
(4.3.51)

where we have introduced the fluctuations with respect to the reversible drift as

$$
\delta_t \hat{A}_\mu \equiv i\mathcal{L}(\phi_t)\hat{A}_\nu - V_\nu(a, t)
$$
(4.3.52)

We want to express this dissipative matrix in terms of the dissipative of the unperturbed system. Consider the Liouville operator that has the form

$$
i\mathcal{L}(\phi_t) = -\frac{\partial(\hat{H}^0 + \phi_\nu(t)\hat{A}_\nu)}{\partial z} \cdot J \cdot \frac{\partial}{\partial z} = i\mathcal{L}^0 - \phi_\nu(t)\frac{\partial \hat{A}_\nu}{\partial z} \cdot J \cdot \frac{\partial}{\partial z}
$$
(4.3.53)

Therefore,

$$
i\mathcal{L}(\phi_t)\hat{A}_\mu = i\mathcal{L}_0\hat{A}_\mu + \{\hat{A}_\mu, \hat{A}_\nu\}\phi_\nu(t)
$$
(4.3.54)

If we substitute this expression in the dissipative matrix (4.3.64) we will obtain four terms

$$D_{\mu\nu}(a,\phi_t) = D^0_{\mu\nu}(a,\phi_t) + \phi_{\mu'}(t)D^a_{\mu'\mu\nu}(a,\phi_t) + D^b_{\mu\nu\nu'}(a,\phi_t)\phi_{\nu'}(t)$$
$$+ \phi_{\mu'}(t)D^b_{\mu'\mu\nu\nu'}(a,\phi_t)\phi_{\nu'}(t) \tag{4.3.55}$$

with obvious, albeit cumbersome, definitions. The first term is closely related to the dissipative matrix of the unperturbed system

$$D^0_{\mu\nu}(a,\phi_t) = \int_0^{\Delta t} dt' \left\langle (i\mathcal{L}^0\hat{A}_\nu - V^0_\nu(a))\exp\{i\mathcal{L}(\phi_t)t'\}(i\mathcal{L}^0\hat{A}_\mu - V^0_\mu(a))\right\rangle^a \tag{4.3.56}$$

where $V^0_\mu(a) = \langle i\mathcal{L}^0\hat{A}_\mu\rangle^a$. If $\phi_t = 0$ this is exactly the dissipative matrix. Therefore, we see that the effect of the external slow forcing is to modify the dissipative matrix by adding terms linear and quadratic in the external field to the dissipative matrix. In addition, the dynamics with which one has to compute the correlations is with a constant external field set to the value ϕ_t.

Of course, the implicit assumption in this slow forcing approximation is that the correlations involved in the dissipative matrix do actually exist. This may impose conditions on the type of external forcing that is permitted. In particular, it seems natural to request that the system under a constant external forcing has a well-defined equilibrium state. Think, for example, of gas in a container with a piston over which an oscillating external force is acting. While there is an equilibrium state when the constant force is in the direction of compressing the gas, there is no equilibrium when the constant force is in the direction of expansion of the gas: the piston just accelerates without bounds. A modulation of the force, without changing the direction, on the other hand, is acceptable.

Under the assumption that in the time scale in which the memory kernel decays, the probability of the CG variables is practically constant, we may approximate $P(a',t') \simeq P(a',t)$ in Eq. (4.4.20). These two approximations are equivalent to approximate the memory kernel (4.4.19) as

$$D_{\mu\nu}(\phi,a,a',t,t') \simeq \delta(a-a')\delta^+(t-t')k_B M_{\mu\nu}(a,\phi_t) \tag{4.3.57}$$

where we have introduced the friction matrix

$$M_{\mu\nu}(a,\phi_t) \equiv \frac{1}{k_B}\int_0^{\Delta t} dt' \left\langle \left(\delta i\mathcal{L}(\phi_{t'})\hat{A}_\nu\right)\mathcal{U}^\dagger(t,t')\left(\delta i\mathcal{L}(\phi_t)\hat{A}_\mu\right)\right\rangle^a \tag{4.3.58}$$

where k_B is the Botzmann constant. Note the ordering of the indices and the use of the adjoint operator $\mathscr{U}^\dagger(t,t')$. The friction matrix has an explicit dependence on time inherited from the time dependence of the external forcing. In fact, the friction matrix is a functional of the whole history of the external forcing, because the protocol is buried inside the evolution operator $\mathcal{U}_{[\phi,t,t']}$. Even though the Markov approximation is not rigorous, we will assume that the integral of the correlation in (4.4.22) exists, in some sense. Note that we have used an upper limit $\Delta t < \infty$ in order to avoid the plateau problem [?].

By inserting (4.4.21) into the generalized FPE (4.4.20) we obtain the FPE under external forcing

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a,\phi_t) + M_{\mu\nu}(a,\phi_t)\frac{\partial S}{\partial a_\nu}(a)\right]P(a,t)$$
$$+ k_B\frac{\partial}{\partial a_\mu}M_{\mu\nu}^S(a,\phi_t)\frac{\partial}{\partial a_\nu}P(a,t) \qquad (4.3.59)$$

where we have introduced the entropy as a natural generalization of Boltzmann entropy, this is

$$S(a) \equiv k_B \ln \Omega(a) \qquad (4.3.60)$$

where the measure $\Omega(a)$ is defined in (28.3.3). The symmetric part of the friction matrix is

$$M_{\mu\nu}^S = \frac{1}{2}(M_{\mu\nu} + M_{\nu\mu}) \qquad (4.3.61)$$

The FPE (4.4.23) is one of the main results of the present paper. When the external forcing is time-independent this FPE reduces to the one obtained by Öttinger [?]. The comparison of (4.4.23) with the FPE obtained by Öttinger reveals that the reversible operator $L_{\mu\nu}(a)$ is unmodified but the energy picks up the natural dependence on the time-dependent external field. On the other hand, the entropy is unmodified by the time-dependent external field, but the friction matrix $M_{\mu\nu}(a,\phi_t)$ acquires a complex dependence on the external field.

4.3.4 Slow external forcing

Some simplification occurs in the dissipative matrix when we assume that the external forcing is slow. For concreteness, assume that the external forcing energy is of the form $E^1(a,t) = a\cdot\phi_t$. Slow forcing means that the time scale of variation of the external forcing ϕ_t is much larger than the time scale of the memory kernel. In this case, we may take the approximation $i\mathcal{L}_\tau \simeq i\mathcal{L}_t$ inside the time-ordered exponential (3.5.4)

$$\exp_+\left\{\int_{t_0}^t d\tau i\mathcal{L}(\phi_\tau)\right\} \approx \sum_{n=0}^\infty \int_{t_0}^t d\tau_1 \cdots \int_{t_0}^{\tau_{n-1}} d\tau_n i\mathcal{L}(\phi_t) \cdots i\mathcal{L}(\phi_t) = \exp\{i\mathcal{L}(\phi_t)t\}$$
$$(4.3.62)$$

In this way, by approximating the projected dynamics with the unprojected dynamics we have

$$\mathcal{U}^\dagger(t,t') = \exp\left\{-i\mathcal{L}(\phi_t)(t-t')\right\} \qquad (4.3.63)$$

This is the evolution operator of the system under a *constant in time* external forcing in whih the particular value of the constant external field is fixed to ϕ_t. In this adiabatic

approximation the dissipative matrix becomes

$$M_{\mu\nu}(a, \phi_t) = \int_0^{\Delta t} dt' \left\langle \delta_t \hat{A}_\nu \exp\{i\mathcal{L}(\phi_t)t'\}\delta_t \hat{A}_\mu \right\rangle^a \qquad (4.3.64)$$

where we have introduced the fluctuations with respect to the reversible drift as

$$\delta_t \hat{A}_\mu \equiv i\mathcal{L}(\phi_t)\hat{A}_\nu - V_\nu(a, t) \qquad (4.3.65)$$

We want to express this dissipative matrix in terms of the dissipative of the unperturbed system. Consider the Liouville operator that has the form

$$i\mathcal{L}(\phi_t) = -\frac{\partial(\hat{H}^0 - \phi_\nu(t)\hat{A}_\nu)}{\partial z}\cdot J\cdot\frac{\partial}{\partial z}$$

$$= i\mathcal{L}^0 + \phi_\nu(t)\frac{\partial \hat{A}_\nu}{\partial z}\cdot J\cdot\frac{\partial}{\partial z} \qquad (4.3.66)$$

Therefore,

$$i\mathcal{L}(\phi_t)\hat{A}_\mu = i\mathcal{L}_0\hat{A}_\mu - \{\hat{A}_\mu, \hat{A}_\nu\}\phi_\nu(t) \qquad (4.3.67)$$

If we substitute this expression in the dissipative matrix (4.3.64) we will obtain four terms

$$M_{\mu\nu}(a, \phi_t) = M_{\mu\nu}^0(a, \phi_t) + \phi_{\mu'}(t)M_{\mu'\mu\nu}^a(a, \phi_t) + M_{\mu\nu\nu'}^b(a, \phi_t)\phi_{\nu'}(t) + \phi_{\mu'}(t)M_{\mu'\mu\nu\nu'}^c(a, \phi_t)\phi_{\nu'}(t) \qquad (4.3.68)$$

with obvious, albeit cumbersome, definitions. The first term is closely related to the dissipative matrix of the unperturbed system

$$M_{\mu\nu}^0(a, \phi_t) = \int_0^{\Delta t} dt' \left\langle (i\mathcal{L}^0\hat{A}_\nu - V_\nu^0(a))\exp\{i\mathcal{L}(\phi_t)t'\}(i\mathcal{L}^0\hat{A}_\mu - V_\mu^0(a)) \right\rangle^a \qquad (4.3.69)$$

where $V_\mu^0(a) = \langle i\mathcal{L}^0\hat{A}_\mu\rangle^a$. If $\phi_t = 0$ this is exactly the dissipative matrix. Therefore, we see that the effect of the external slow forcing is to modify the dissipative matrix by adding terms linear and quadratic in the external field to the dissipative matrix. In addition, the dynamics with which one has to compute the correlations is with a constant external field set to the value ϕ_t.

   Of course, the implicit assumption in this slow forcing approximation is that the correlations involved in the dissipative matrix do actually exist. This may impose conditions on the type of external forcing that is permitted. In particular, it seems natural to request that the system under a constant external forcing has a well-defined equilibrium state. Think, for example, of gas in a container with a piston over which an oscillating external force is acting. While there is an equilibrium state when the constant force is in the direction of compressing the gas, there is no equilibrium when the constant force is in the direction of expansion of the gas: the piston just accelerates without bounds. A modulation of the force, without changing the direction, on the other hand, is acceptable.

4.3.5 Properties of the building blocks in the FPE

We now prove that, quite remarkably, the GENERIC properties described in the case of time-independent external forcing, still hold in the case that there is a time-dependent external forcing. In particular, the first degeneracy condition of GENERIC , involving the reversible operator and the entropy, that reads

$$L_{\mu\nu}(a)\frac{\partial S}{\partial a_\nu}(a) + k_B \frac{\partial}{\partial a_\nu} L_{\mu\nu}(a) = 0 \tag{4.3.70}$$

is a consequence of the following identity

$$-\frac{\partial}{\partial a_\mu} V_\mu(a, \phi_t)\Omega(a) = -\frac{\partial}{\partial a_\mu} \text{Tr}[\hat{\Psi}_a i\mathcal{L}(\phi_t)\hat{A}_\mu]$$

$$= \text{Tr}[i\mathcal{L}(\phi_t)\hat{\Psi}_a] = 0 \tag{4.3.71}$$

By using the form (4.4.16) of the reversible drift in (4.3.71) we easily arrive at the identity

$$\frac{\partial}{\partial a_\mu} \left[L_{\mu\nu}(a)\frac{\partial S}{\partial a_\nu} \right] = 0 \tag{4.3.72}$$

which gives (4.3.70).

The second degeneracy condition is

$$M_{\mu\nu}(a, \phi_t)\frac{\partial E}{\partial a_\nu}(a, \phi_t) = 0$$

$$\frac{\partial E}{\partial a_\mu}(a, \phi_t) M_{\mu\nu}(a, \phi_t) = 0 \tag{4.3.73}$$

The proof of the second identity in (4.3.73) relies on the definition (4.4.22) of the dissipative matrix and the identity (??). In fact,

$$M_{\mu\nu}(a, \phi_t)\frac{\partial E}{\partial a_\nu}(a, \phi_t) = \frac{\partial E}{\partial a_\nu}(a, \phi_t) \int_0^{\Delta t} dt' \left\langle \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a, \phi_t') \right) \mathcal{U}^\dagger(t, t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a$$

$$= \int_0^{\Delta t} dt' \frac{\partial E}{\partial a_\nu}(a, \phi_t') \left\langle \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a, \phi_t') \right) \mathcal{U}^\dagger(t, t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a \tag{4.3.74}$$

Note that we have substituted t by t' in the energy derivative when entering inside the time integral. This is consistent with the Markovian approximation. Then

$$M_{\mu\nu}(a, \phi_t)\frac{\partial E}{\partial a_\nu}(a, \phi_t) = \int_0^{\Delta t} dt' \left\langle \frac{\partial \hat{H}}{\partial a_\nu}(\hat{A}, t) \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a, \phi_t') \right) \mathcal{U}^\dagger(t, t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a$$

$$= \int_0^{\Delta t} dt' \left\langle \underbrace{i\mathcal{L}(\phi_t)\hat{H}(\hat{A}, t)}_{=0} \mathcal{U}^\dagger(t, t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a = 0 \tag{4.3.75}$$

On the other hand, the third identity in Eqs. (4.3.73) is

$$\frac{\partial E}{\partial a_\mu}(a, \phi_t) M_{\mu\nu}(a, \phi_t) = 0 \tag{4.3.76}$$

where we left multiply the gradient, does not require to use the Markov assumption.

Note that the reversible drift term satisfies the following two identities

$$V_\mu(a, \phi_t)\frac{\partial E}{\partial a_\mu}(a, \phi_t) = 0$$

$$V_\mu(a, \phi_t)\frac{\partial S}{\partial a_\mu}(a) = 0 \tag{4.3.77}$$

Both identities are proved by using the form (4.4.16) of the drift term in terms of the reversible operator. With the combination of two equations (4.4.29) we can prove

$$\frac{\partial}{\partial a_\mu}(V_\mu(a, \phi_t)\Phi(E(a, \phi_t))\Omega(a)) = 0 \tag{4.3.78}$$

where $\Phi(E)$ is an arbitrary function of the energy. This identity will be useful below.

4.3.6   First Law

Let us consider the time-dependent probability density $P(E, t)$ of the total energy. By definition it is given by

$$P(E, t) = \int da P(a, t)\delta(E(a, \phi_t) - E) \tag{4.3.79}$$

The time derivative of this probability takes the form

$$\partial_t P(E, t) = -\frac{\partial}{\partial E}\langle\partial_t E\rangle_t^E P(E, t) \tag{4.3.80}$$

where the conditional expectation is defined as

$$\langle\partial_t E\rangle_t^E = \frac{1}{P(E, t)}\int da P(a, t)\delta(E(a, \phi_t) - E)\partial_t E(a, \phi_t) \tag{4.3.81}$$

For a time-independent external forcing $\partial_t E(a, \phi_t) = 0$ and we have that $\partial_t P(E, t) = 0$. The probability $P(E, t)$ becomes time independent.

Consider the average of the energy $E(a, \phi_t)$ performed with the solution $P(a, t)$ of the FPE (4.4.23)

$$\overline{E}_t = \int da E(a, \phi_t)P(a, t) \tag{4.3.82}$$

The time derivative of the average energy is given by

$$\frac{d}{dt}\overline{E}_t = \int da\, P(a,t)\frac{\partial E}{\partial t}(a,\phi_t) + \int da\, E(a,\phi_t)\frac{\partial P}{\partial t}(a,t) \tag{4.3.83}$$

The second term turns out to vanish exactly. To see this, use the FPE (4.4.23) and integrate by parts

$$\int da\, E(a,\phi_t)\frac{\partial P}{\partial t}(a,t) = \int da\,\frac{\partial E}{\partial a_\mu}\left[\left[L_{\mu\nu}\frac{\partial E}{\partial a_\nu} + M_{\mu\nu}\frac{\partial S}{\partial a_\nu}\right]P - M_{\mu\nu}\frac{\partial}{\partial a_\nu}P\right] = 0 \tag{4.3.84}$$

that vanishes because of the antisymmetry of $L_{\mu\nu}$ and the degeneracy (4.3.70).

Therefore we see that the variation of the total energy of the system is given by

$$\frac{d}{dt}\overline{E}_t = \int da\, P(a,t)\frac{\partial E^1}{\partial t}(a,\phi_t) \tag{4.3.85}$$

that involves only the external forcing term in the energy. If the external forcing is time-independent, then total energy is conserved. When the external forcing is of the form $E^1(a,\phi_t) = a\cdot\phi_t$ the variation of the energy takes the form

$$\frac{d}{dt}\overline{E}_t = a(t)\dot{\phi}_t \tag{4.3.86}$$

where $a(t)$ is the average of a with the solution $P(a,t)$ of the FPE.

Note that we may be interested not in the average change of the total energy but only in the actual change of the energy of the unperturbed system, discounting the energy due to the coupling, this is, from (4.4.3)

$$\left\langle E^0\right\rangle_t = \left\langle E(t) - E^1(t)\right\rangle_t \tag{4.3.87}$$

The time derivative of this energy is given by

$$\frac{d}{dt}\left\langle E^0\right\rangle_t = \int da\, P(a,t)\frac{\partial E^1}{\partial t}(a,\phi_t) - \frac{d}{dt}\left\langle E^1(t)\right\rangle_t \tag{4.3.88}$$

where we have used (4.4.34).

4.3.7   The Second Law

The entropy of a given level of description is obtained by inserting the relevant ensemble (4.3.4) into the Gibbs-Jaynes entropy functional (??). This gives the familiar result

$$\begin{aligned}
S[P] &= -k_B \int da\, P(a,t)\ln\frac{P(a,t)}{\Omega(a)} \\
&= \int da\, S(a)P(a,t) - k_B \int da\, P(a,t)\ln P(a,t)
\end{aligned} \tag{4.3.89}$$

Under the assumption that $P(a,t)$ is a solution of the FPE (4.4.23), the time derivative of the entropy functional is

$$\frac{d}{dt}S[P(a,t)] = k_B \int da P(a,t) \left[\frac{\partial}{\partial a_\mu}\ln\frac{P(a,t)}{\Omega(a)}\right] M_{\mu\nu}(a,t)\left[\frac{\partial}{\partial a_\nu}\ln\frac{P(a,t)}{\Omega(a)}\right] \quad (4.3.90)$$

Provided that the symmetric part of the dissipative matrix $M_{\mu\nu}(a)$ is positive semi-definite, we have

$$\frac{d}{dt}S[P(a,t)] \geq 0 \quad (4.3.91)$$

The proof that the dissipative matrix $M_{\mu\nu}(a,t)$ is a positive semi-definite function seems to be difficult. However, if it is not, the whole meaning of the Fokker-Planck approach becomes dubious [?]. We will assume, for the time being, that the dissipative matrix is positive semi-definite, where the semi character is due to the degeneracy (4.3.73).

4.3.8   The long time limit

The entropy functional (4.4.38) has a global maximum over the set of normalized probabilities. This maximum satisfies

$$\frac{\delta}{\delta P(a,t)}\left[S[P] - \mu\int da P(a,t)\right] = 0 \quad (4.3.92)$$

where μ is the Lagrange multiplier that enforces that $\int da P(a,t) = 1$. By using the explicit form (4.4.38), this equation takes the form

$$S(a) - (\mu+1) + k_B \ln P^*(a,t) = 0 \quad (4.3.93)$$

or

$$P^*(a,t) = \frac{\exp\{S(a)/k_B\}}{\int da' \exp\{S(a')/k_B\}} \quad (4.3.94)$$

The Fokker-Planck evolution from $P(a,0)$ to $P(a,t)$ is such that the entropy functional always grows, but this does not imply necessarily, that the final global maximum (4.4.41) of the entropy functional will be reached at long times. In particular, note that if, as time evolves, the probability would tend towards a probability satisfying

$$\frac{\partial}{\partial a_\mu}\ln\frac{P(a,t)}{\Omega(a)} \propto \frac{\partial E}{\partial a_\mu}(a,t) \quad (4.3.95)$$

then the time derivative (4.4.39) will vanish, due to the degeneracy condition, but the final distribution function will not be the global maximum (4.4.41).

When the external forcing is time-independent, total energy is conserved. The global maximum of the entropy functional (4.4.38) in the subspace of normalized probability functions that give a prescribed (constant) distribution of energy $P(E)$ is obtained by

maximizing the unconstrained Lagrange functional

$$S[P] - \mu \int da P(a,t) - \int da \lambda(E(a,t)) P(a,t) \tag{4.3.96}$$

where $\lambda(E,t)$ is a set of Lagrange multipliers that enforce the relationship (4.3.79). The result of the maximization is

$$P^*(a,t) = \frac{\lambda(E(a,t),t) \exp\{S(a)/k_B\}}{\int da' \lambda(E(a',t)) \exp\{S(a')/k_B\}} \tag{4.3.97}$$

By substitution into (4.3.79) we find the explicit form of the Lagrange multiplier $\lambda(E,t)$ and the final result

$$P^*(a,t) = \left. \frac{P(E,t)}{\Omega(E,t)} \right|_{E=E(a,t)} \exp\{S(a)/k_B\} \tag{4.3.98}$$

where we have introduced

$$\Omega(E,t) = \int da \delta(E(a,t) - E)\Omega(a) \tag{4.3.99}$$

as the measure of phase space compatible with energy E at time t.

Note that the time derivative of the entropy functional, when evaluated at the probability of the form (4.3.98) vanishes identically

$$\frac{d}{dt}S[P^*] = 0 \tag{4.3.100}$$

In summary, the time derivative of the entropy is always positive when evaluated on the solution of the FPE, and vanishes for a probability of the form (4.3.97). If a probability of the form (4.3.97) was a solution of the FPE, this would imply that the Fokker-Planck equation woulde lead the probability at long times towards a probability of the form (4.3.98) at sufficient long times. However, (4.3.97) is not a solution of the FPE because it is an eigenfunction of the FPE operator with zero eigenvalue. Therefore, only in the time-independent case, we have that (4.3.98) is a solution.

4.4   Dynamics of probabilities under external forcing. Obsolete

In this section, we derive the exact equation for the evolution of the probability $P(a,t)$ of the CG variables when there is external forcing.

The strategy we follow is to use the fact that the probability is the average of the probabilistic CG variables (4.3.2) and use results from the averages flavour[1].

We will assume that the unperturbed Hamiltonian can be expressed in terms of the CG variables. This can always be done [2]. This is, we assume that the time-dependent Hamiltonian

$$\hat{H}_t(z) = \hat{H}^0(z) + \hat{H}^1(z,t) \tag{4.4.1}$$

where the $\hat{H}^0(z)$ is the unperturbed Hamiltonian of the system and $\hat{H}^1(z,t)$ is the effect of the external forcing, can be expressed in the form

$$\hat{H}_t(z) = E^0(\hat{A}(z)) + E^1(\hat{A}(z),t) = E(\hat{A},t) \tag{4.4.2}$$

where the total energy function

$$E(a,t) = E^0(a) + E^1(a,t) \tag{4.4.3}$$

is the sum of the energy function $E^0(a)$ of the unperturbed system and $E^1(a,t)$ is the energy function describing the action of the external forcing. Note that the Hamiltonian (4.4.2) is more general than (4.2.1), which is included as the particular case $E^1(a,t) = a{\cdot}\phi_t$. Note that, importantly, we assume that the pertubation Hamiltonian $\hat{H}_t^1(z)$ is expressed in terms of the CG variables. In general, this means that for the validity of the expressions that follow, the coupling functions with the external field and the system should be included in the list of the relevant variables.

The Liouville operator (4.5.3) becomes

$$i\mathcal{L}(\phi_t) = i\mathcal{L}^0 + i\mathcal{L}_t^1 \tag{4.4.4}$$

where the unperturbed Liouville operator is

$$i\mathcal{L}^0 = -\frac{\partial \hat{H}^0}{\partial z}(z){\cdot}J{\cdot}\frac{\partial}{\partial z}$$

$$i\mathcal{L}_t^1 = -\frac{\partial \hat{H}^1}{\partial z}(z,t){\cdot}J{\cdot}\frac{\partial}{\partial z} \tag{4.4.5}$$

---

[1] Here the direct approach seems to be not as straightforward. **[Why?]**

[2] If it cannot be done, then include the unperturbed Hamiltonian $\hat{H}^0(z)$ in the list of variables

We have the different effects of these operators on the Hamiltonians

$$i\mathcal{L}(\phi_t)\hat{H}_t(z) = 0$$
$$i\mathcal{L}^0\hat{H}^0(z) = 0$$
$$i\mathcal{L}_t^1\hat{H}_t^1(z) = 0 \qquad (4.4.6)$$

The objective is to find a closed equation for the probability $P(a,t)$ which is the average of the probabilistic CG variables (4.3.2), where the average is computed with the solution $\rho_t(z)$ of the Liouville equation under external forcing (4.2.45). To obtain the evolution equation for the averages of the CG variables, one introduces the relevant ensemble $\overline{\rho}_t(z)$ which is through the Maximum Entropy Principle and has the form given in (2.7.25). This relevant ensemble is, presumably, a good approximation to the real ensemble, and it makes sense to perform a perturbation of the form (3.1.9). One convenient way to do this perturbation is by introducing the Kawasaki-Gunton projector (??) and its adjoint (??). Eqs (??)-(??) are unaltered, except that now the equation for the non-relevant part (??) involves the time-dependent Liouville operator, this is

$$\partial_t \delta\rho_t(z) = -\mathcal{Q}_t i\mathcal{L}(\phi_t)\delta\rho_t(z) - \mathcal{Q}_t i\mathcal{L}(\phi_t)\overline{\rho}_t(z) \qquad (4.4.7)$$

This results into a redefinition of the evolution operator of the projected dynamics in (??) that now reads

$$\mathcal{U}_{[\phi,t,t']} = \exp_+ \left\{ -\int_{t'}^{t} dt'' \mathcal{Q}_{t''} i\mathcal{L}(\phi_{t''}) \right\} \qquad (4.4.8)$$

and it includes an implicit functional dependence on the external field ϕ_t. The important result is given in (??) that we quote again

$$\rho_t(z) = \overline{\rho}_t(z) - \int_0^t dt' \mathcal{U}_{[\phi,t,t']} \mathcal{Q}_{t'} i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}(z) \qquad (4.4.9)$$

as it expresses the real ensemble with the relevant ensemble at present and past times and contains, therefore, the potential for a closed equation for the CG variables. We now use this result in

$$\partial_t P(a,t) = \text{Tr}\left[\hat{\Psi}_a \partial_t \rho_t\right] = \text{Tr}\left[\rho_t i\mathcal{L}(\phi_t)\hat{\Psi}_a\right] \qquad (4.4.10)$$

where we have used the Liouville's equation (4.2.46) and an integration by parts. We now use the decomposition (4.4.9) in (4.4.10)

$$\partial_t P(a,t) = \underbrace{\text{Tr}\left[\overline{\rho}_t(z)i\mathcal{L}(\phi_t)\hat{\Psi}_a\right]}_{(1)} - \underbrace{\int_0^t dt'\text{Tr}\left[\left(i\mathcal{L}(\phi_t)\hat{\Psi}_a\right)\mathcal{U}_{[\phi,t,t']}\mathcal{Q}_{t'}i\mathcal{L}(\phi_{t'})\overline{\rho}_{t'}(z)\right]}_{(2)}$$

$$(4.4.11)$$

Now use the results

$$i\mathcal{L}\hat{\Psi}_a(z) = -\frac{\partial}{\partial a_\mu}\hat{\Psi}_a(z)i\mathcal{L}(\phi_t)\hat{A}_\mu(z) \tag{4.4.12}$$

and the form (2.7.25) of the relevant ensemble in the averages flavour in the form

$$\overline{\rho}_t = \int da P(a,t)\frac{1}{\Omega(a)}\hat{\Psi}_a(z) \tag{4.4.13}$$

The drift term in (4.4.11) becomes

$$(1) = -\frac{\partial}{\partial a_\mu}\left\langle i\mathcal{L}(\phi_t)\hat{A}_\mu\right\rangle^a P(a,t) \tag{4.4.14}$$

By using the form (4.5.3) of the Liouville operator with the Hamiltonian (4.4.2) we have, by following similar steps as those in (3.6.3)-(3.6.5),

$$(1) = -\frac{\partial}{\partial a_\mu}V_\mu(a,t)P(a,t) \tag{4.4.15}$$

where the reversible drift is given by

$$V_\mu(a,t) = \left\langle i\mathcal{L}(\phi_t)\hat{A}_\mu\right\rangle^a = L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a,t) \tag{4.4.16}$$

where

$$L_{\mu\nu}(a) = \left\langle \{\hat{A}_\mu, \hat{A}_\nu\}\right\rangle^a \tag{4.4.17}$$

Therefore, the reversible operator is not modified by the external forcing.

By using (4.4.12), (4.4.13), and the fact that the Kawasaki-Gunton projector becomes the Zwanzig projector (i.e. (??)-(??)), the second term in (4.4.11) becomes

$$\begin{aligned}
(2) &= \int_0^t dt' \int da' \frac{P(a',t')}{\Omega(a)}\text{Tr}\left[i\mathcal{L}(\phi_t)\Psi_{a'}\mathcal{Q}\mathcal{U}_{[\phi,t,t']}i\mathcal{L}(\phi_t)\hat{\Psi}_a\right]\\
&= \frac{\partial}{\partial a_\mu}\int_0^t dt' \int da' \frac{P(a',t')}{\Omega(a)}\frac{\partial}{\partial a'_\nu}\Omega(a')\left\langle \left(\mathcal{Q}^\dagger i\mathcal{L}(\phi_{t'})\hat{A}_\nu\right)\mathcal{U}^\dagger(t,t')\hat{\Psi}_a i\mathcal{L}(\phi_t)\hat{A}_\mu\right\rangle^{a'}\\
&= -\frac{\partial}{\partial a_\mu}\int_0^t dt' \int da'\Omega(a')D_{\mu\nu}(a,a't,t')\frac{\partial}{\partial a'_\nu}\frac{P(a',t')}{\Omega(a)} \tag{4.4.18}
\end{aligned}$$

where we have introduced the memory kernel

$$D_{\mu\nu}(\phi,a,a',t,t') = \left\langle \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a',\phi_{t'})\right)\mathcal{U}^\dagger(t,t')\hat{\Psi}_a i\mathcal{L}(\phi_t)\hat{A}_\mu\right\rangle^{a'} \tag{4.4.19}$$

Eqs (4.4.15) and (4.4.19) in (4.4.11) give the generalized FPE in the presence of external

forcing

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu} L_{\mu\nu}(a)\frac{\partial E_t}{\partial a_\nu}(a)P(a,t)$$
$$+\frac{\partial}{\partial a_\mu}\int_0^t dt'\int da'\Omega(a')D_{\mu\nu}(\phi,a,a',t,t')\frac{\partial}{\partial a_\nu'}\frac{P(a',t')}{\Omega(a)} \qquad (4.4.20)$$

The structure is formally identical to the generalized FPE (3.5.14) except that now the reversible drift depends on time and the memory kernel depends on both t, t' and not only on the time difference $t - t'$.

The Markovian approximation

The Markov assumption in which the memory is assumed to decay in a time scale much larger than the CG variables, is implemented in the form, already familiar from the discussion on the non-perturbed case in (3.5.23)

$$D_{\mu\nu}(\phi,a,a',t,t') = \delta(a-a')\delta^+(t-t')D_{\mu\nu}(a,t) \qquad (4.4.21)$$

where we have introduced

$$D_{\mu\nu}(a,t) = \int_0^{\Delta t} dt'\left\langle\left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a,\phi_t')\right)\mathcal{U}^\dagger(t,t')i\mathcal{L}(\phi_t)\hat{A}_\mu\right\rangle^a \qquad (4.4.22)$$

This dissipative matrix has an explicit dependence on time inherited from the time dependence of the external forcing. Of course, the Markov approximation is not rigorous and, in the presence of external fields, requires that the integral of the correlation in (4.4.22) exists, in some sense. In principle, the dissipative matrix is a functional of the external forcing ϕ_t.

By inserting (4.4.21) into the generalized FPE (4.4.20) we obtain the **FPE under external forcing**

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a,t) + D_{\mu\nu}(a,t)\frac{\partial S_B}{\partial a_\nu}(a)\right]P(a,t)$$
$$+\frac{\partial}{\partial a_\mu}D_{\mu\nu}(a,t)\frac{\partial}{\partial a_\nu}P(a,t) \qquad (4.4.23)$$

When $\phi_t = 0$, this FPE reduces to (3.6.6), that was obtained under no external forcing. The comparison of (4.4.23) and (3.6.6) reveals that the reversible drift is unmodified but the energy picks up the natural dependence on the external field. On the other hand, the bare entropy is unmodified by the external field, but the dissipative matrix acquires a complex dependence on the external field.

Note that we may also write the FPE (4.4.23) in the compact form

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[V_\mu(a,t)P(a,t) - \Omega(a)D_{\mu\nu}(a,t)\frac{\partial}{\partial a_\nu}\frac{P(a,t)}{\Omega(a)}\right] \qquad (4.4.24)$$

Properties of the FPE

The FPE (4.4.23) has time-dependent drift and diffusion and, therefore, it does not have, in general, a stationary equilibrium solution. In particular, it is no longer true that the FPE has an H-theorem that would warrant the existence of a unique equilibrium state.

We now prove that, quite remarkably, the GENERIC properties described in the case of no external forcing, still hold in the case that there is an external forcing. In particular, all the GENERIC degeneracy properties are satisfied

$$L_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu} + \frac{\partial}{\partial a_\nu}L_{\mu\nu}(a) = 0$$

$$D_{\mu\nu}(a,t)\frac{\partial E}{\partial a_\nu}(a,t) = \frac{\partial E}{\partial a_\mu}(a,t)D_{\mu\nu}(a,t) = 0 \tag{4.4.25}$$

The first equation is a consequence of (3.6.22) and (3.6.23), which are unchanged because there is no reference to the forced dynamics in them.

The proof of the second identity in (4.4.25) relies on the definition (4.4.22) of the dissipative matrix and the identity $i\mathcal{L}(\phi_t)\hat{H}(z,t) = 0$, encountered in (4.4.6). In fact,

$$\begin{aligned}
D_{\mu\nu}(a,t)\frac{\partial E}{\partial a_\nu}(a,t) &= \frac{\partial E}{\partial a_\nu}(a,t)\int_0^{\Delta t} dt' \left\langle \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a,\phi_t')\right)\mathcal{U}^\dagger(t,t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a \\
&= \int_0^{\Delta t} dt'\frac{\partial E}{\partial a_\nu}(a,\phi_t')\left\langle \left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a,\phi_t')\right)\mathcal{U}^\dagger(t,t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a
\end{aligned} \tag{4.4.26}$$

Note that we have, within the range of the Markovian approximations done to derive the dissipative matrix, substituted t by t' in the energy derivative when entering inside the time integral. Then

$$\begin{aligned}
D_{\mu\nu}(a,t)\frac{\partial E}{\partial a_\nu}(a,t) &= \int_0^{\Delta t} dt' \left\langle \frac{\partial \hat{H}}{\partial a_\nu}(\hat{A},t)\left(i\mathcal{L}(\phi_{t'})\hat{A}_\nu - V_\nu(a,\phi_t')\right)\mathcal{U}^\dagger(t,t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a \\
&= \int_0^{\Delta t} dt' \left\langle \underbrace{i\mathcal{L}(\phi_t)\hat{H}(\hat{A},t)}_{=0}\mathcal{U}^\dagger(t,t')i\mathcal{L}(\phi_t)\hat{A}_\mu \right\rangle^a = 0
\end{aligned} \tag{4.4.27}$$

On the other hand, the identity

$$\frac{\partial E}{\partial a_\mu}(a,t)D_{\mu\nu}(a,t) = 0 \tag{4.4.28}$$

where we left multiply the gradient, does not require to use the Markov assumption.

Two other important identities are also satisfied by the drift term which are general-

izations of identities encountered in Sec **??**.

$$V_\mu(a,t)\frac{\partial E}{\partial a_\mu}(a,t) = 0$$

$$\frac{\partial}{\partial a_\mu}\left(V_\mu(a,t)\Omega(a)\right) = 0 \qquad (4.4.29)$$

Both identities are proved by using the form (4.4.16) of the drift term in terms of the reversible operator. With the combination of two equations (4.4.29) we can prove

$$\frac{\partial}{\partial a_\mu}\left(V_\mu(a,t)\Phi(E(a,t))\Omega(a)\right) = 0 \qquad (4.4.30)$$

where $\Phi(E)$ is an arbitrary function of the energy. This identitiy will be useful in a moment.

First Law

Consider the average of the energy $E(a,t)$ performed with the solution $P(a,t)$ of the FPE (4.4.23)

$$\langle E(t)\rangle_t = \int da E(a,t)P(a,t) \qquad (4.4.31)$$

We now compute the time derivative which is given by

$$\frac{d}{dt}\langle E(t)\rangle_t = \int da P(a,t)\frac{\partial E}{\partial t}(a,t) + \int da E(a,t)\frac{\partial P}{\partial t}(a,t) \qquad (4.4.32)$$

The second term turns out to vanish exactly. To see this, use the FPE (4.4.23) and integrate by parts

$$\int da E(a,t)\frac{\partial P}{\partial t}(a,t) = \int da\frac{\partial E}{\partial a_\mu}\left[\left[L_{\mu\nu}\frac{\partial E}{\partial a_\nu} + D_{\mu\nu}\frac{\partial S_B}{\partial a_\nu}\right]P - D_{\mu\nu}\frac{\partial}{\partial a_\nu}P\right] = 0 \quad (4.4.33)$$

that vanishes because the antisymmetry of $L_{\mu\nu}$ and the degeneracy (4.4.25).

Therefore we see that the variation of the total energy of the system is given by

$$\frac{d}{dt}\langle E(t)\rangle_t = \int da P(a,t)\frac{\partial E^1}{\partial t}(a,t) \qquad (4.4.34)$$

that involves only the external forcing term in the energy. When the external forcing is of the form $E^1(a,t) = a\cdot\phi_t$ the variation of the energy takes the form

$$\frac{d}{dt}\langle E(t)\rangle_t = a(t)\dot{\phi}_t \qquad (4.4.35)$$

where $a(t)$ is the average of a with the solution $P(a,t)$ of the FPE.

Note that we may be interested not in the average change of the total energy but only

in the actual change of the energy of the unperturbed system, discounting the energy due to the coupling, this is, from (4.4.3)

$$\left\langle E^0 \right\rangle_t = \left\langle E(t) - E^1(t) \right\rangle_t \tag{4.4.36}$$

The time derivative of this energy is given by

$$\frac{d}{dt} \left\langle E^0 \right\rangle_t = \int da P(a,t) \frac{\partial E^1}{\partial t}(a,t) - \frac{d}{dt} \left\langle E^1(t) \right\rangle_t \tag{4.4.37}$$

where we have used (4.4.34).

[How the probability $P(E,t)$ evolves in time?]

Second Law

As we know, the entropy of the probabilistic level of description is obtained by inserting the relevant ensemble (2.7.25) of the probabilistic level into the Gibbs-Jaynes entropy functional, leading to the familiar result

$$S[P] = - \int da P(a,t) \ln \frac{P(a,t)}{\Omega(a)} \leq 0 \tag{4.4.38}$$

which is identical to the un-forced problem. The entropy functional is negative for all $P(a,t)$ and a convex functional. Therefore, if it has a maximum, it is unique. When considering the time derivative of the entropy functional, under the assumption that $P(a,t)$ is a solution of the FPE (4.4.23), we may reproduce exactly the steps in Eqs (3.9.2)-(3.9.5) whith the drift $V_\mu(a)$ and the dissipative matrix $D_{\mu\nu}(a)$ substituted by their externally forced counterparts $V_\mu(a,t), D_{\mu\nu}(a,t)$. Eq (3.9.4) is satisfyed in this case thanks to (4.4.29). Therefore, we arrive at the following time derivative of the entropy functional

$$\frac{d}{dt} S[P(a,t)] = \int da P(a,t) \left[\frac{\partial}{\partial a_\mu} \ln \frac{P(a,t)}{\Omega(a)} \right] D_{\mu\nu}(a,t) \left[\frac{\partial}{\partial a_\nu} \ln \frac{P(a,t)}{\Omega(a)} \right] \tag{4.4.39}$$

Provided that the symmetric part of the dissipative matrix $D_{\mu\nu}(a)$ is positive semi-definite, we have

$$\frac{d}{dt} S[P(a,t)] \geq 0 \tag{4.4.40}$$

There is no guarantee that the dissipative matrix $D_{\mu\nu}(a,t)$ is a positive semi-definite function. However, if it is not, the whole meaning of the Fokker-Planck approach becomes dubious. We will assume, for the time being, that the dissipative matrix is positive semi-definite.

Therefore, the entropy (4.4.38), which is negative and takes its global maximum at $\Omega(a)$, keeps increasing when evaluated at the solution of the FPE (4.4.23). However, it does not increases until reaching the global maximum at $\Omega(a)$ (which in general is not non-normalizable!) because of the degeneracy condition (4.4.25) of the dissipative

matter. In fact, this time derivative will vanish for any distribution function of the form

$$P^*(a,t) = \Phi(t, E(a,t)) \exp\{S_B(a)\} \tag{4.4.41}$$

as can be easily be shown by substituting (4.4.41) in (4.4.39) and using the degeneracy condition of the dissipative matrix. The function $\Phi(t, E)$ is restricted by the fact that $P^*(a,t)$ is normalized and is a solution of the FPE (4.4.23). Both conditions are satisfied if

$$\partial_t \Phi(t, E(a,t)) + \partial_E \Phi(t, E(a,t)) \partial_t E(a,t) = 0 \tag{4.4.42}$$

Therefore, the evolution of the probability $P(a,t)$ is such that the entropy increases all the time until it reaches a time-independent value given by

$$\begin{aligned} S[P^*] &= -\int da P^*(a,t) \ln \frac{P^*(a,t)}{\Omega(a)} \\ &= -\int da P^*(a,t) \ln \frac{\Phi(E(a,t))}{\Phi(t)} \end{aligned} \tag{4.4.43}$$

Despite the implicit time dependence it is easy to see that $S[P^*]$ does not depend on time.

[We could fix the form of $\Phi(E)$ in the unperturbed case by recoursing to conservation of energy. Can we fix it again in the same way?]

Einstein's formula in external forcing

4.5   Linear Response Theory

In this section, we consider the response of a system to an external field, by following an entirely microscopic approach.

Consider a system with a Hamiltonian $\hat{H}^0(z)$ that is subject to the action of an external field. We represent this action with a time-dependent contribution $\hat{H}_t^1(z)$ in such a way that the evolution of the system is governed by the Hamiltonian

$$\hat{H}_t(z) = \hat{H}^0(z) + \hat{H}_t^1(z) \tag{4.5.1}$$

The Liouville operator contains, therefore, two parts

$$i\mathcal{L}(\phi_t) = i\mathcal{L}^0 + i\mathcal{L}_t^1 \tag{4.5.2}$$

The formal solution of the Liouville's equation

$$\partial \rho_t(z) = -i\mathcal{L}(\phi_t)\rho_t(z) \tag{4.5.3}$$

is

$$\rho_t(z) = \exp_+ \left\{ -i \int_0^t d\tau i\mathcal{L}(\phi_\tau) \right\} \rho_0(z) = \mathcal{U}(t,0)\rho_0(z) \tag{4.5.4}$$

We want to relate this probability distribution with the solution of the unperturbed evolution,

$$\partial \rho_t^0(z) = -i\mathcal{L}^0 \rho_t^0(z) \tag{4.5.5}$$

with solution

$$\rho_t^0(z) = \exp_+ \left\{ -i\mathcal{L}^0 t \right\} \rho_0(z) \tag{4.5.6}$$

To this end, we construct an equation for the deviation $\Delta\rho_t = \rho_t^1(z) - \rho_t^0(z)$, which is easily seen to be

$$\partial_t \Delta\rho_t = -i\mathcal{L}(\phi_t)\Delta\rho_t(z) - i\mathcal{L}_t^1 \rho_t^0(z) \tag{4.5.7}$$

This inhomogeneous linear equation has the solution

$$\Delta\rho_t(z) = \mathcal{U}(t,0)\Delta\rho_0(z) - \int_0^t dt' \mathcal{U}_{[\phi,t,t']} i\mathcal{L}_{t'}^1 \rho_{t'}^0(z) \tag{4.5.8}$$

Under the assumption that at the initial time $\rho_0^1(z) = \rho_0^0(0)$ this equation has the form

$$\rho_t^1(z) = \rho_t^0(z) - \int_0^t dt' \mathcal{U}_{[\phi,t,t']} i\mathcal{L}_{t'}^1 \rho_{t'}^0(z) \tag{4.5.9}$$

and we have expressed the perturbed ensemble entirely in terms of the unperturbed ensemble at past times. This is a very formal exact result. Its interest is manifested

when we assume that the external perturbation is small and we may approximate the evolution under the perturbed Hamiltonian with the unperturbed one. To first order, we obtain $\mathcal{U}_{[\phi,t,t']} \approx \exp\{-i\mathcal{L}^0(t-t')\}$ and

$$\rho_t^1(z) = \rho_t^0(z) - \int_0^t dt' \exp\{-i\mathcal{L}^0(t-t')\}i\mathcal{L}_{t'}^1\rho_{t'}^0(z) \qquad (4.5.10)$$

In this case, from the evolution of the unperturbed system (which is assumed to be known) we obtain that of the perturbed system. The average of an arbitrary function $\hat{F}$ with (4.5.10) is

$$\begin{aligned}
\left\langle \hat{F} \right\rangle_t^1 &= \left\langle \hat{F} \right\rangle_t^0 - \int_0^t dt' \mathrm{Tr}[\hat{F}\exp\{-i\mathcal{L}^0(t-t')\}i\mathcal{L}_{t'}^1\rho_{t'}^0] \\
&= \left\langle \hat{F} \right\rangle_t^0 + \int_0^t dt' \mathrm{Tr}[\hat{F}\exp\{i\mathcal{L}^0(t-t')\}\frac{\partial \hat{H}_t^1}{\partial z}\cdot J\cdot\frac{\partial}{\partial z}\rho_{t'}^0] \\
&= \left\langle \hat{F} \right\rangle_t^0 - \int_0^t dt' \mathrm{Tr}[\rho_{t'}^0 \frac{\partial}{\partial z}\left(\exp\{i\mathcal{L}^0(t-t')\}\hat{F}\right)\cdot J\cdot\frac{\partial \hat{H}_t^1}{\partial z}] \\
&= \left\langle \hat{F} \right\rangle_t^0 - \int_0^t dt' \left\langle \left\{ \hat{F}^0(t-t'), \hat{H}_t^1 \right\} \right\rangle_{t'}^0 \qquad (4.5.11)
\end{aligned}$$

This is still formal and not extremely useful. A nicer result arises when the perturbed Hamiltonian has the structure

$$\hat{H}_t^1(z) = -\phi_\mu(t)\hat{A}_\mu(z) \qquad (4.5.12)$$

for a given set of phase functions $\hat{A}(z)$ and we assume that the unperturbed system is at equilibrium $\rho_t^0(z) = \rho^{\mathrm{eq}}(z) = \exp\{-\beta\hat{H}^0(z)\}/Z(\beta)$. In this case, we obtain the standard result in linear response theory

$$\left\langle \hat{F} \right\rangle_t^1 = \left\langle \hat{F} \right\rangle^{\mathrm{eq}} + \int_0^t dt' \chi_\nu(t-t')\phi_\mu(t') \qquad (4.5.13)$$

where the response function is

$$\chi_\mu(t-t') = \left\langle \left\{ \hat{F}^0(t-t'), \hat{A}_\mu \right\} \right\rangle^{\mathrm{eq}} \qquad (4.5.14)$$

The response function is determined entirely from the dynamics of the unperturbed system and is therefore, a property of the system. The result (4.5.13) describes how the averages of the phase functions $\hat{F}$ depart from their equilibrium values under the action of the external forcing $\phi_\mu(t)$, when this forcing is sufficiently small.

The response function (4.5.14) can be given a very interesting form

$$\chi_\mu(t) = \mathrm{Tr}\left[\rho^{\mathrm{eq}}(z)\frac{\partial \hat{F}^0(t)}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\mu}{\partial z}\right] = -\mathrm{Tr}\left[\frac{\partial \rho^{\mathrm{eq}}(z)}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\mu}{\partial z}\hat{F}^0(t)\right]$$

$$= \beta\mathrm{Tr}\left[\rho^{\mathrm{eq}}(z)\frac{\partial \hat{H}^0}{\partial z}\cdot J\cdot\frac{\partial \hat{A}_\mu}{\partial z}\hat{F}^0(t)\right] = \beta\left\langle \hat{F}^0(t)i\mathcal{L}^0\hat{A}_\mu\right\rangle^{\mathrm{eq}} \qquad (4.5.15)$$

The last form can also be expressed as

$$\chi_\mu(t) = -\theta(t)\beta\frac{d}{dt}\left\langle \hat{F}^0(t)\hat{A}_\mu\right\rangle^{\mathrm{eq}} \qquad (4.5.16)$$

where the unit step function explicitly takes into account that the response function is defined only for $t > 0$. The expression (4.5.16) relates the response function to a correlation function and it is termed **the fluctuation-dissipation theorem of the first kind**, to distinguish it from the theorem that relates the correlations of the time derivatives of the CG variables with the dissipative matrix (3.5.38), that it is usually referred as fluctuation-dissipation of the second kind or Green-Kubo formula.

4.5.1   Linear response theory from the FPE

We may treat the FPE (4.4.23) in very much the same way as we did with the microscopic Liouville (4.5.3) equation in Sec 4.5. In fact, we may write the FPE (4.4.23) in the form

$$\partial_t P(a,t) = -(\mathcal{L}^0 + \mathcal{L}_t^1)P(a,t) \qquad (4.5.17)$$

where the Fokker-Planck operators are

$$\mathcal{L}^0 = \frac{\partial}{\partial a_\mu}\left[L_{\mu\nu}(a)\frac{\partial E^0}{\partial a_\nu}(a) + D^0_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu}(a) - D^0_{\mu\nu}(a)\frac{\partial}{\partial a_\nu}\right]$$

$$\mathcal{L}_t^1 = \frac{\partial}{\partial a_\mu}\left[L_{\mu\nu}(a)\frac{\partial E^1}{\partial a_\nu}(a,t) + \Delta D^1_{\mu\nu}(a,t)\frac{\partial S_B}{\partial a_\nu}(a) - \Delta D^1_{\mu\nu}(a,t)\frac{\partial}{\partial a_\nu}\right] \qquad (4.5.18)$$

where $\Delta D^1(a,t) = D(a,t)-D^0(a)$ and $D^0(a)$ is the dissipative matrix of the unperturbed system. We can now relate the solution $P(a,t)$ of the perturbed FPE (4.4.23) with the solution $P^0(a,t)$ of the unperturbed FPE. Define $\Delta P(a,t) = P(a,t)-P^0(a,t)$ and obtain, by following identical steps as those leading to (4.5.9)

$$P(a,t) = P^0(a,t) - \int_0^t dt'\,\mathscr{U}(t,t')i\mathscr{L}_{t'}^1 P^0(a,t') \qquad (4.5.19)$$

where we have assumed that at the initial time $P(a,0) = P^0(a,0)$ and have introduced the evolution operator of the perturbed dynamics as

$$\mathscr{U}(t,t') = \exp_+\left\{-\int_t^{t'} d\tau(\mathscr{L}_\tau^0 + \mathscr{L}_\tau^1)\right\} \qquad (4.5.20)$$

This operator allows to obtain the solution of the FPE (4.4.23) as

$$P(a,t) = \mathcal{U}(t,0)P(a,0) \tag{4.5.21}$$

An integral version of this operator can be obtained in terms of the conditional probability, thanks to the relationship

$$P(a,t) = \int da_0 P(a_0,0)P(a_0,0|a,t) \tag{4.5.22}$$

The conditional probability is the solution of the FPE (4.4.23) with initial condition $P(a_0,0|a,0) = \delta(a_0 - a)$. This suggests that we may write (4.5.19) in the integral form

$$P(a,t) = P^0(a,t) + \int_0^t dt' \int da_0 P(a_0,t'|a,t)\mathcal{L}_{t'}^1 P^0(a_0,t') \tag{4.5.23}$$

Despite the possible fuzzyness in the argument, it is straightforward to show that (4.5.23) is correct by showing that the time derivative of both sides are the same, while at $t = 0$ this is also correct.

The exact equation (4.5.23) is the basis for a perturbation expansion. In linear order in the external forcing, we may approximate the transition probability $P(a_0,0|a,t)$ with the transition probability $P^0(a_0,0|a,t)$ of the unperturbed system. In this case, the left hand side of (4.5.23) is entirely determined by the dynamics of the unperturbed system.

The exact and formal (4.5.23) becomes interesting when we assume that the initial probability is the equilibrium one $P(a,0) = P^{\mathrm{eq}}(a) \propto \exp\{S_B(a)\}$. Note that

$$\mathcal{L}(\phi_t)^1 P^{\mathrm{eq}}(a) = \mathcal{L}^0\left(E_t^1(a)P^{\mathrm{eq}}(a)\right) \tag{4.5.24}$$

as it is seen from the definitions (4.5.18). Therefore, at equilibrium we have

$$P(a,t) = P^{\mathrm{eq}}(a) + \int_0^t dt' \int da_0 P^0(a_0,t'|a,t)\mathcal{L}^0\left(E_{t'}^1(a_0)P^{\mathrm{eq}}(a_0)\right) \tag{4.5.25}$$

[Grabert dice]

$$
\begin{aligned}
P(a,t) &= P^{\mathrm{eq}}(a) + \int_0^t dt' \int da_0 \partial_t P^0(a_0,t'|a,t)E_{t'}^1(a_0)P^{\mathrm{eq}}(a_0) \\
&= P^{\mathrm{eq}}(a) + \int_0^t dt' \int da_0 \partial_t P^0(a_0,t',a,t)E_{t'}^1(a_0) \\
&= P^{\mathrm{eq}}(a) - \int_0^t dt' \int da_0 \chi(a_0,a,t-t')E_{t'}^1(a_0) \tag{4.5.26}
\end{aligned}
$$

4.6   Exact Fluctuation Theorems

In this section, we consider a number of exact theorems known collectively as Fluctuation Theorems (FTs). The generality of these theorems is very large, and have been proved not only for Hamiltonian but also for other kind of reversible dynamics. We stick, however, to the Hamiltonian framework and all the discussion is performed within the realm of Hamiltonian dynamics. We do not describe nor use artificial dynamics, as for example, the Nosé-Hoover dynamics, isokinetic, or other types of dynamics that emulates a thermostated dynamics. Partidaries of such dynamics will enjoy books by Hoover [] and Evans et al. []. From a logical and conceptual point of view, a thermostate is a system that is interacting with our system in a very particular way, and whose degrees of freedom are eliminated from the description, in a coarse-graining procedure. Therefore, the way that we will deal with thermostates or, more precisely, energy reservoirs will be within the framework of the theory of coarse-graining.

The interest of the FTs lies in the fact that they are some of the few known exact results in Non-Equilibrium Statistical Mechanics. The literature on FT is quite large and a number of useful reviews exist [?],[?],[?].

We present first the Evans-Searles FT, then the Crook's identity that implies in turn the Jarsinsky identity. The Evans-Searles FT does not rely on the concept of work but on that of the dissipation function. On the other hand, Crook's identity. We also include other theorems that emerge from the use of the relevant ensemble, which is so characteristic of the coarse-graining approach that we consider in this book.

The Evans-Searles TFT has been shown to be valid for other reversible dynamics different from the Hamiltonian dynamics discussed in the present book. We refer to the interested reader to the original literature [].

4.6.1 The Transient Fluctuation Theorem

We can pose an interesting and very general question about the evolution of a Hamiltonian system subject to uncertainty in the initial conditions. What is the "degree of evolution" of the system? In the case that an equilibrium state exists (which is what we assume when there are no time-dependent external forcings) we could also make the question of how far from equilibrium is the system. We consider here the first question. One way to measure the "degree of evolution" is by quantifying how different are the initial ensemble $\rho_0(z)$ from the solution $\rho_t(z)$ of the Liouville equation with initial condition $\rho_0(z)$. One way to compare two probability distributions is with the help of an important statistical object known as the **Kullback-Leibler (KL) divergence** from $\rho_t(z)$ to $\rho_0(z)$ which is defined as

$$S[\rho_t|\rho_0] = \int dz \rho_t(z) \ln \frac{\rho_t(z)}{\rho_0(z)} \tag{4.6.1}$$

which is also referred to as the **relative entropy**. This quantity gives a measure of how dissimilar two probability distributions are. Very similar probabilities $\rho_t \sim \rho_0$ for which $\ln \rho_t/\rho_0 \sim 0$ have small values of the KL divergence. The KL divergence is always a positive quantity[3] and takes its minimum value of zero when the two distributions coincide. Note that the KL divergence is not a distance because $S[\rho_t|\rho_0] \neq S[\rho_0|\rho_t]$ (and does not satisfy the triangle inequality). Instead, as a consequence of Liouville theorem (4.2.45)

$$\rho_t(z) = \rho_0\left(\hat{\Phi}_{[\phi,0,t]}^{-1}(z)\right) \tag{4.6.2}$$

we have

$$\begin{aligned} S[\rho_0|\rho_t] &= \int dz \rho_t(z) \ln \frac{\rho_t(z)}{\rho_0(z)} = \int dz \rho_0\left(\hat{\Phi}_{[\phi,0,t]}^{-1}(z)\right) \ln \frac{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}^{-1}(z)\right)}{\rho_0(z)} \\ &= \int dz' \rho_0(z') \ln \frac{\rho_0(z')}{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z')\right)} \\ &= S\left[\rho_t(z_t)|\rho_0\right] \neq S\left[\rho_t|\rho_0\right] \end{aligned} \tag{4.6.3}$$

The KL divergence (4.6.1) can be understood as the time dependent average of the following time-dependent function of phase space

$$\hat{\eta}_t(z) \equiv \ln \frac{\rho_t(z)}{\rho_0(z)} \tag{4.6.4}$$

Gibbs in pag. 16 of his foundational work [?] introduced the terminology "index of probability" to refer to the logarithm of the ensemble density and, therefore, we may venture to call the function (4.6.4) the **relative index of probabilities**. Alternatively,

---

[3]This is a consequence of Jensen's inequality.

the last identity in (4.6.3) suggests to introduce the **dissipation function** [?]

$$\hat{\Sigma}_t^F(z) \equiv \ln \frac{\rho_0(z)}{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)} = \ln \frac{\rho_0(z)}{\rho_0\left(z_t^F\right)} \tag{4.6.5}$$

The dissipation function is fully determined by the initial ensemble $\rho_0(z)$ and the Hamiltonian dynamics $z_t = \hat{\Phi}_{[\phi,0,t]}(z)$. Note that for the dissipation function to be well-defined the argument of the logarithms needs to be strictly positive. The assumption is that the dynamics is such that for all microstates z for which $\rho_0(z) \neq 0$ (the only ones for which the dissipation function can be defined) $\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right) \neq 0$ for all times. This property is named **transitive metric??** [?].

Although for the results to come we can equaly use the relative index of probability $\hat{\eta}(z)$ or the dissipation function $\hat{\Sigma}_{[\phi,0,t]}(z)$, the latter takes a nicer expression in most cases of interest. For example, imagine that the initial ensemble $\rho_0(z)$ is given by the canonical ensemble $\rho_0(z) = \frac{1}{Z}e^{-\beta H_0(z)}$ for certain Hamiltonian $H_0(z)$. In this case, the dissipation function takes the form

$$\hat{\Sigma}_t^F(z) = \ln \frac{\rho_0(z)}{\rho_0(z_t)} = \beta\left[\hat{H}_0(z_t) - \hat{H}_0(z)\right] \tag{4.6.6}$$

that gives the change of the Hamiltonian $\hat{H}_0(z)$ under the dynamics generated by the external forcing. The index of probability, on the other hand would come with a different time ordering

$$\hat{\eta}_t(z) = -\beta\left[\hat{H}_0(z_{-t}) - \hat{H}_0(z)\right] \tag{4.6.7}$$

which is somewhat more awkward.

Obviously the average dissipation function, being a KL divergence satisfies, see last identity in (4.6.3)

$$\left\langle\hat{\Sigma}_t^F\right\rangle_0 = \int dz\rho_0(z)\ln\frac{\rho_0(z)}{\rho_0(z_t^F)} \geq 0 \tag{4.6.8}$$

Using the word "dissipation" in this inequality reads "dissipation is always positive" suggests a "demonstration of the Second Law", but this inequality tells nothing more profound than the fact that the initial ensemble differs (in KL sense) in general from the past ensemble (i.e. the ensemble that would give $\rho_0(z)$ if evolved forward).

A more interesting non-trivial result known as the **Transient Fluctuation Theorem** [?], emerges when instead of looking at the *average* of the dissipation function we look at the *probability distribution* of the dissipation function which, by definition, is

$$P^F(\Sigma) = \int dz\rho_0(z)\delta\left(\hat{\Sigma}_t^F(z) - \Sigma\right) \tag{4.6.9}$$

We also define the probability in the reversed process

$$P^B\left(\Sigma\right) = \int dz \rho_0(z) \delta\left(\hat{\Sigma}^B_t(z) - \Sigma\right) \tag{4.6.10}$$

where the dissipation in the reverse process is defined as

$$\hat{\Sigma}^B_t(z) = \ln \frac{\rho_0(z)}{\rho_0\left(z^B_t\right)} = \ln \frac{\rho_0(z)}{\rho_0\left(\epsilon\cdot\hat{\Phi}_{[\epsilon\cdot\phi,0,t]}(\epsilon\cdot z)\right)} \tag{4.6.11}$$

The essential two assumptions in the theorem are the microscopic reversibility (4.6.73) and the assumption that the initial ensemble is even under time reversal, this is

$$\rho_0(\epsilon\cdot z) = \rho_0(z) \tag{4.6.12}$$

Under these assumption, the Transient Fluctuation Theorem states that

$$\frac{P^F\left(+\Sigma\right)}{P^B\left(-\Sigma\right)} = e^{\Sigma} \tag{4.6.13}$$

In words, starting from a set of initial conditions distributed according to the initial ensemble $\rho_0(z)$, the probability to find a trajectory with a positive value of the dissipation function is exponentially larger than the probability of finding a trajectory with the time-reversed protocol, with a negative value. When there is no external forcing, this means that trajectories starting from the initial distribution $\rho_0(z)$ with positive dissipation are exponentially more probable than those with negative dissipation. This is an exact result that can be understood as a signature of the irreversibility arising from the purely reversible Hamilton's equations.

Demonstration of the theorem

By using the definition of the dissipation function (4.6.5) in (4.6.9)

$$P^F\left(\Sigma\right) = \int dz \rho_0(z) \delta\left(\ln \frac{\rho_0(z)}{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)} - \Sigma\right) = \int dz \rho_0(z) \delta\left(\ln \frac{e^{-\Sigma}\rho_0(z)}{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)}\right) \tag{4.6.14}$$

The microstates for which the argument of the Dirac delta vanishes are such that $\rho_0(z) = e^{\Sigma}\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)$, therefore,

$$P^F\left(\Sigma\right) = e^{\Sigma} \int dz \rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right) \delta\left(\ln \frac{\rho_0(z)}{\rho_0\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)} - \Sigma\right) \tag{4.6.15}$$

Use microscopic reversibility

$$P^F\left(\Sigma\right) = e^{\Sigma}\int dz \rho_0\left(\epsilon\cdot\hat{\Phi}^{-1}_{[\varepsilon\cdot\phi^R,0,t]}(\epsilon\cdot z)\right)\delta\left(\ln\frac{\rho_0(z)}{\rho_0\left(\epsilon\cdot\hat{\Phi}^{-1}_{[\varepsilon\cdot\phi^R,0,t]}(\epsilon\cdot z)\right)} - \Sigma\right) \quad (4.6.16)$$

Perform the change of variables

$$z' = \epsilon\cdot\hat{\Phi}^{-1}_{[\varepsilon\cdot\phi^R,0,t]}(\epsilon\cdot z)$$
$$z = \epsilon\cdot\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}(\epsilon\cdot z') \quad (4.6.17)$$

in such a way that

$$P^F\left(\Sigma\right) = e^{\Sigma}\int dz \rho_0\left(z'\right)\delta\left(\ln\frac{\rho_0\left(\epsilon\cdot\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}(z')\right)}{\rho_0(z')} - \Sigma\right)$$

$$= e^{\Sigma}\int dz \rho_0\left(z\right)\delta\left(\ln\frac{\rho_0(z)}{\rho_0\left(z_t^B\right)} + \Sigma\right) \quad (4.6.18)$$

Under the assumption (4.6.12) that the initial ensemble is even under time reversal and the definition (4.6.5) of the dissipation function, we have

$$P^F\left(\Sigma\right) = e^{\Sigma}\int dz \rho_0(z)\delta\left(\hat{\Sigma}_t^B + \Sigma\right) = e^{\Sigma}P^B\left(-\Sigma\right) \quad (4.6.19)$$

which is the desired result (4.6.13).

Non-equilibrium partition of the identity

An immediate consequence of (4.6.19) is obtained by considering the normalization of the probability

$$1 = \int_{-\infty}^{\infty} d\Sigma P^F(\Sigma) = \int_{-\infty}^{\infty} d\Sigma e^{\Sigma}P^B(-\Sigma) \quad (4.6.20)$$

that implies

$$\int_{-\infty}^{\infty} d\Sigma e^{-\Sigma}P^B(\Sigma) = 1 \quad (4.6.21)$$

By using the definition of the probability (4.6.9), and the fact that the protocol is arbitrary, this becomes the so called **non-equilibrium partition of the identity**

$$\left\langle \exp\{-\hat{\Sigma}_t^F\}\right\rangle_0 = 1 \quad (4.6.22)$$

The average of the exponential of the negative of the dissipation function over all the trajectories starting from the initial distribution $\rho_0(z)$ is, for all times, equal to 1.

The identity (4.6.22) seems quite remarkable but it boils down to the simple statement that the initial ensemble is normalized to unity, as can be seen from the following chain of identities

$$1 = \int dz \rho_0(z) = \int dz' \rho_0 \left(\hat{\Phi}_{[\phi,0,t]}(z') \right) = \int dz \rho_0(z) \frac{\rho_0 \left(\hat{\Phi}_{[\phi,0,t]}(z) \right)}{\rho_0(z)}$$

$$= \int dz \rho_0(z) \exp \left\{ - \ln \frac{\rho_0(z)}{\rho_0 \left(\hat{\Phi}_{[\phi,0,t]}(z) \right)} \right\} = \left\langle \exp\{ -\hat{\Sigma}_t^F \} \right\rangle_0 \qquad (4.6.23)$$

4.6.2   The CG Transient Fluctuation Theorem

As we know, an essential concept in the theory of coarse-graining is that of relevant ensemble $\overline{\rho}_t(z)$. This ensemble, which is different from the solution of the Liouville equation $\rho_t(z)$, reproduces the macroscopic information (i.e. the averages or the probabilities of the CG variables, depending on the flavour). One way to quantify the differences between the relevant and real ensembles is through the Kullback-Leibler divergence between both probability densities

$$D_{KL}[\rho_t||\overline{\rho}_t] = \int dz \rho_t(z) \ln \frac{\rho_t(z)}{\overline{\rho}_t(z)} \geq 0 \qquad (4.6.24)$$

By a simple change of variables of the form $z = \mathcal{U}_t z'$ that has unit Jacobian we have

$$D_{KL}[\rho_t||\overline{\rho}_t] = \int dz' \rho_t(\mathcal{U}_t z') \ln \frac{\rho_t(\mathcal{U}_t z')}{\overline{\rho}_t(\mathcal{U}_t z')} \geq 0 \qquad (4.6.25)$$

and because of the Liouville theorem (4.2.44) we may express the KL divergence as an average over the initial ensemble

$$D_{KL}[\rho_t||\overline{\rho}_t] = \int dz \rho_0(z) \ln \frac{\rho_0(z)}{\overline{\rho}_t(\mathcal{U}_t z)} \geq 0 \qquad (4.6.26)$$

We show that this distance can be quantified in quite physical terms. The relevant ensemble in the averages flavour takes the form

$$\overline{\rho}_{a_t}(z) = \rho^0 \frac{\exp\{-\lambda_t \hat{A}(z)\}}{Z(\lambda_t)} \qquad (4.6.27)$$

where we denote explicitly that the relevant ensemble depends on time only through the time dependence of the averages a_t of the CG variables. The conjugate parameters $\lambda_t = \lambda(a_t)$ are a function of the averages, determined by the solution of

$$a = -\frac{\partial \ln Z(\lambda)}{\partial \lambda} \qquad (4.6.28)$$

The dimensionless partition function defined as

$$Z(\lambda) \equiv \int dz \rho^0 \exp\{-\lambda \cdot \hat{A}(z)\} \qquad (4.6.29)$$

Recall that the dressed entropy (2.7.13) is given by

$$S(a) = \ln Z(\lambda(a)) + \lambda(a)a \qquad (4.6.30)$$

Under the usual assumption that the initial ensemble is of the relevant form we have

$$D_{KL}[\rho_t||\overline{\rho}_t] = \int dz \overline{\rho}_0(z) \ln \frac{\overline{\rho}_0(z)}{\overline{\rho}_t(\mathcal{U}_t z)} \geq 0 \qquad (4.6.31)$$

which is entirely expressed in terms of *relevant* ensembles at the initial and final times.

We can now use (4.6.27) and (4.6.30) and conclude with the following exact result

$$D_{KL}[\rho_t||\overline{\rho}_{a_t}] = S(a_t) - S(a_0) \geq 0 \qquad (4.6.32)$$

This expresses that the dressed entropy function evaluated at the final macrostate is always larger than the entropy evaluated at the initial macrostate. *This is a quite remarkable exact result.* It is based on the only physical assumption that the initial ensemble is of the relevant form. The physical content of this result is just that the relevant ensemble is different from the real ensemble. The dissimilarity between both ensembles is quantified by the difference of dressed entropies at the initial and final states.

Note that the probabilistic flavour is just a particular case of the average flavour in which the probability is given by the average of the CG probabilistic functions (4.3.2). In the probabilistic flavour where the relevant ensemble is given by

$$\overline{\rho}_t(z) = \frac{P(\hat{A}, t)}{\Omega(\hat{A}(z))} \qquad (4.6.33)$$

we have a similar result stating that the entropy of this flavour, given by

$$S[P_t] = - \int da P_t(a) \ln \frac{P_t(a)}{\Omega(a)} \qquad (4.6.34)$$

satisfies the counterpart of (4.6.32), which is now

$$D_{KL}[\rho_t||\overline{\rho}_t] = S[P_t] - S[P_0] \geq 0 \qquad (4.6.35)$$

Note that the KL divergence (4.6.31) can be understood as the average over initial conditions of the following **CG dissipation function**

$$\overline{\Sigma}_t^F(z) \equiv \ln \frac{\overline{\rho}_{a_0}(z)}{\overline{\rho}_{a_t}(z_t^F)} \qquad (4.6.36)$$

where $z_t^F = \hat{\Phi}_{[\phi,0,t]}(z)$ is the evolved microstate under the forced dynamics characterized with the protocol ϕ_t. This phase function *is different* from the dissipation function introduced in (4.6.5) as it uses the relevant ensemble instead of the real ensemble. By using the relevant ensemble (4.6.27) in the CG dissipation function (4.6.36), we arrive

$$\overline{\Sigma}_t^F(z) \equiv S(a_t) + \lambda_t(\hat{A}(z_t) - a_t) - S(a_0) + \lambda_0(\hat{A}(z) - a_0)$$
$$= -(\Phi(\lambda(a_t)) - \Phi(\lambda(a_0))) + \lambda_t \hat{A}(z_t) - \lambda_0 \hat{A}(z) \qquad (4.6.37)$$

where the dimensionless thermodynamic potential is $\Phi(\lambda) = -\ln Z(\lambda)$. The CG dissipation function depends on both the microscopic trajectory z_t in phase space and the macroscopic trajectory a_t on the space of averages of the CG functions.

The average of the CG dissipation function with respect to the initial ensemble is the

KL divergence (4.6.31) and we recover

$$\left\langle \overline{\Sigma}_t^F \right\rangle_0 = S(a_t) - S(a_0) \tag{4.6.38}$$

that justifies the name of CG dissipation function for $\overline{\Sigma}_t$.

Now, instead of computing the average value of the dissipation function, as in (4.6.38), we can compute the *probability density* that the CG dissipation function $\overline{\Sigma}_t$ takes a particular value Σ. Recall that the CG dissipation function is defined on trajectories. We will compare the probability of the dissipation function in the forward and the backward processes defined in (4.2.39). We define two different probabilities named the forward and backward probabilities which are defined as

$$P^F(\Sigma) = \int dz \overline{\rho}_{a_0}(z) \delta \left(\ln \frac{\overline{\rho}_{a_0}(z)}{\overline{\rho}_{a_t}\left(z_t^F\right)} - \Sigma \right) \tag{4.6.39}$$

and

$$P^B(\Sigma) = \int dz \overline{\rho}_{a_t}(z) \delta \left(\ln \frac{\overline{\rho}_{a_t}(z)}{\overline{\rho}_{a_0}\left(z_t^B\right)} - \Sigma \right) \tag{4.6.40}$$

The **CG Transient Fluctuation Theorem** relates the two probababilities by stating that

$$\frac{P^F(\Sigma)}{P^B(-\Sigma)} = e^{\Sigma} \tag{4.6.41}$$

This implies that the probability of finding a value of Σ in the forward process is exponentially larger than finding a value of $-\Sigma$ in the backward process.

Demonstration of the CG Transient Fluctuation Theorem

We now prove the Fluctuation Theorem involving this probability. The demonstration is very similar to the one leading to the Transient Fluctuation Theorem (4.6.13). As in that case, the present theorem makes use of the reversibility of the underlying Hamiltonian dynamics.

As a first step, we need to know how the relevant ensemble transforms under the time reversal map. The relevant ensemble is fully determined by the averages of the CG variables. We will assume that the CG variables have the parity (1.8.35) under time reversal. It is straightforward to show that the behaviour of the averages a, conjugate variables λ, and dressed entropy $S(a)$ behave under a time reversal operation as follows

$$\begin{aligned}
a(\varepsilon \cdot \lambda) &= \varepsilon \cdot a(\lambda) \\
\lambda(\varepsilon \cdot a) &= \varepsilon \lambda(a) \\
Z(\varepsilon \cdot \lambda) &= Z(\lambda) \\
S(\varepsilon \cdot a) &= S(a) \tag{4.6.42}
\end{aligned}$$

As a consequence the relevant ensemble transforms as

$$\overline{\rho}_a(\epsilon\cdot z) = \overline{\rho}_{\varepsilon\cdot a}(z) \tag{4.6.43}$$

Next, note that we can write the forward probability (4.6.39) as

$$P^F(\Sigma) = \int dz \overline{\rho}_{a_0}(z) \delta\left(\ln e^{-\Sigma} \frac{\overline{\rho}_{a_0}(z)}{\overline{\rho}_{a_t}\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)}\right) \tag{4.6.44}$$

Only the microstates z that make the argument of the logarithm equal to 1 will contribute to the integral. These microstates satisfy

$$e^{-\Sigma}\overline{\rho}_{a_0}(z) = \overline{\rho}_{a_t}\left(\hat{\Phi}_{[\phi,0,t]}(z)\right) \tag{4.6.45}$$

and therefore,

$$P^F(\Sigma) = \int dz \overline{\rho}_{a_t}\left(\hat{\Phi}_{[\phi,0,t]}(z)\right) \delta\left(\ln e^{-\Sigma} \frac{\overline{\rho}_{a_0}(z)}{\overline{\rho}_{a_t}\left(\hat{\Phi}_{[\phi,0,t]}(z)\right)}\right) \tag{4.6.46}$$

According to the reversal of the relevant ensemble (4.6.43) we have

$$\overline{\rho}_{a_t}(\hat{\Phi}_{[\phi,0,t]}(z)) = \overline{\rho}_{\varepsilon\cdot a_t}(\epsilon\hat{\Phi}_{[\phi,0,t]}(z)) \tag{4.6.47}$$

With further use of the reversibility property (4.2.38) this gives

$$\overline{\rho}_{a_t}(\hat{\Phi}_{[\phi,0,t]}(z)) = \overline{\rho}_{\varepsilon\cdot a_t}\left(\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}^{-1}(\epsilon\cdot z)\right) \tag{4.6.48}$$

By inserting this expression in (4.6.46) we obtain

$$P^F(\Sigma) = e^{\Sigma} \int dz \overline{\rho}_{\varepsilon\cdot a_t}\left(\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}^{-1}(\epsilon\cdot z)\right) \delta\left(\ln \frac{\overline{\rho}_{a_0}(z)}{\overline{\rho}_{\varepsilon\cdot a_t}\left(\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}^{-1}(\epsilon\cdot z)\right)} - \Sigma\right) \tag{4.6.49}$$

In this expression we now make the change of variables

$$z' = \hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}^{-1}(\epsilon\cdot z) \qquad\qquad \epsilon\cdot z = \hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}(z') \tag{4.6.50}$$

giving

$$\begin{aligned}
P^F(\Sigma) &= e^{\Sigma} \int dz' \overline{\rho}_{\varepsilon\cdot a_t}(z') \delta\left(\ln \frac{\overline{\rho}_{a_0}(\epsilon\cdot\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}(z'))}{\overline{\rho}_{\varepsilon\cdot a_t}(z')} - \Sigma\right) \\
&= e^{\Sigma} \int dz \overline{\rho}_{\varepsilon\cdot a_t}(\epsilon\cdot z') \delta\left(\ln \frac{\overline{\rho}_{a_0}(\epsilon\cdot\hat{\Phi}_{[\varepsilon\cdot\phi^R,0,t]}(\epsilon\cdot z))}{\overline{\rho}_{\varepsilon\cdot a_t}(\epsilon\cdot z)} - \Sigma\right)
\end{aligned} \tag{4.6.51}$$

where we have performed a trivial change of variables $z' = \epsilon\cdot z$. By using the properties

(4.6.43) of the relevant ensemble we end up

$$P^F(\Sigma) = e^{\Sigma} \int dz \overline{\rho}_{a_t}(z) \delta \left(\ln \frac{\overline{\rho}_{a_0}(z_t^R)}{\overline{\rho}_{a_t}(z)} - \Sigma \right) = e^{\Sigma} P^B(-\Sigma) \qquad (4.6.52)$$

which is the desired result (4.6.41).

Generalized Crook's Theorem

The CG Transient Fluctuation Theorem admits a form that can be understood as a generalization of Crook's Theorem. Crook's Theorem is a particular case when the selected CG variables consists on the total energy of a system. For this reason, it will be studied in the chapter about Thermodynamics in the second Volume. Consider the microscopic definition of the forward (4.6.39) and backward (4.6.40) probabilities and introduce the explicit form of the relevant ensemble (4.6.27). One readily arrives at

$$P^F(\Sigma) = \int dz \overline{\rho}_{a_0}(z) \delta \left(\hat{W}_t^F(z) - \Delta\Phi - \Sigma \right)$$
$$P^B(\Sigma) = \int dz \overline{\rho}_{a_t}(z) \delta \left(\hat{W}_t^B(z) + \Delta\Phi - \Sigma \right) \qquad (4.6.53)$$

where we have introduced generalized forward and backward microscopic works as

$$\hat{W}_t^F(z) = \lambda_t \hat{A}(z_t^F) - \lambda_0 \hat{A}(z)$$
$$\hat{W}_t^B(z) = \lambda_0 \hat{A}(z_t^B) - \lambda_t \hat{A}(z) \qquad (4.6.54)$$

and the difference in the thermodynamic potential as

$$\Delta\Phi = -\ln Z_t + \ln Z_0 \qquad (4.6.55)$$

which is a quantity that does not depend on the microstate. By defining the forward and backward probabilities for the generalized microscopic works

$$P_W^F(W) = \int dz \overline{\rho}_{a_0}(z) \delta \left(\hat{W}_t^F(z) - W \right)$$
$$P_W^B(W) = \int dz \overline{\rho}_{a_t}(z) \delta \left(\hat{W}_t^B(z) - W \right) \qquad (4.6.56)$$

In terms of the probabilities of the works, the probabilities of the dissipation (4.6.53) are

$$P^F(\Sigma) = P_W^F(\Sigma + \Delta\Phi)$$
$$P^B(\Sigma) = P_W^B(\Sigma - \Delta\Phi) \qquad (4.6.57)$$

Then the CG Transient Fluctuation Theorem (4.6.41) implies

$$e^{\Sigma} = \frac{P^F(\Sigma)}{P^B(-\Sigma)} = \frac{P_W^F(\Sigma + \Delta\Phi)}{P_W^B(\Sigma - \Delta\Phi)} \qquad (4.6.58)$$

or by a trivial change of variables

$$\frac{P_W^F(W)}{P_W^B(-W)} = \exp\{W + \Delta\Phi\} \tag{4.6.59}$$

Non-equilibrium CG partition of the identity

An immediate consequence of (4.6.52) is obtained by considering the normalization of the probability

$$1 = \int_{-\infty}^{\infty} d\Sigma P^B(\Sigma) = \int_{-\infty}^{\infty} d\Sigma e^{\Sigma} P^F(-\Sigma) \tag{4.6.60}$$

that implies

$$\int_{-\infty}^{\infty} d\Sigma e^{-\Sigma} P^F(\Sigma) = 1 \tag{4.6.61}$$

By using the definition of the probability (4.6.39) this becomes the so called **non-equilibrium partition of the identity**

$$1 = \int_{-\infty}^{\infty} d\Sigma e^{-\Sigma} \int dz \overline{\rho}_{a_0}(z) \delta\left(\overline{\Sigma}_t(z) - \Sigma\right) = \left\langle \exp\{-\overline{\Sigma}_t\} \right\rangle_0 \tag{4.6.62}$$

In the same way as we discussed in (4.6.23) this is a trivial consequence of the normalization of the relevant ensemble. However, in the present case we may use second identity in (4.6.37) for the CG dissipative function it order to arrive at a non-trivial result

$$1 = \left\langle \exp\{-\overline{\Sigma}_t\} \right\rangle_0 = \exp\{\Phi(a_t) - \Phi(a_0)\} \left\langle \exp\{-\lambda_t \hat{A}(t) + \lambda_0 \hat{A}\} \right\rangle_0 \tag{4.6.63}$$

or equivalently

$$\left\langle \exp\{-\lambda_t \hat{A}(t) + \lambda_0 \hat{A}\} \right\rangle_0 = \exp\{-(\Phi(a_t) - \Phi(a_0))\} \tag{4.6.64}$$

In terms of the work in the forward process defined in (4.6.54) this is

$$\left\langle \exp\{-\hat{W}_t^F\} \right\rangle_0 = \exp\{-(\Phi(a_t) - \Phi(a_0))\} \tag{4.6.65}$$

This is a generalization of **Jarzinsky's theorem** for a set of arbitrary CG variables. It relates the average over initial conditions sampled from the relevant ensemble $\rho_{a_0}(z)$ of the exponential of the "work" with a free energy difference.

CG Transient Fluctuation theorem for Markov process

Because (4.6.32) can be written as

$$\int_0^t d\tau \frac{d}{d\tau} S(a_\tau) \geq 0 \qquad\qquad (4.6.66)$$

the result (4.6.32) can be understood as a mild version of the Second Law. The Second Law is, in fact, described by the stronger result

$$\frac{d}{d\tau} S(a_\tau) \geq 0 \qquad\qquad (4.6.67)$$

which is not necesarily ensured by (4.6.66).

4.6.3 The Detailed Balance Theorem

The Detailed Balance Theorem reflects the property of microscopic reversibility of Hamilton's equation at the macroscopic level, by relating the forward and backward transition probabilities of arbitrary phase functions. We have already encountered this theorem in Sec 1.8.3 for the case of time-independent Hamiltonians when we discussed the green dog-flea model. In the present section, we show that the theorem still holds in the case of a forced time-dependent Hamiltonian. Note that while the Transient Fluctuation Theorem does not require the concept of macrostate, in the Detailed Balance Theorem jumps between macrostates are considered. The resulting theorem has a "microcanonical aroma" as compared with the CG Transient Fluctuation Theorem, that has a "canonical aroma".

Consider a collection of phase variables $\hat{A}(z)$, referred as CG variables. For simplicity in the presentation, the coupling functions appearing in the external forcing term in the Hamiltonian are among the list of CG variables. The joint probability distributions $P_{[\phi]}(a_0, t_0, a_1, t_1)$ that the CG variables $\hat{A}(z)$ take the value a_0 at time t_0 *and* the value a_1 at time t_1 is

$$P_{[\phi]}(a_0, t_0, a_1, t_1) = \int dz \rho_{t_0}(z) \delta(\hat{A}(z) - a_0) \delta(\hat{A}(z^F(z)) - a_1) \qquad (4.6.68)$$

where the microscopic forward process is defined as $\hat{z}^F(z) = \hat{\Phi}_{[\phi, t_0, t_1]}(z)$. In (11.0.15), $\rho_0(z)$ is the probability density at the initial time.

We will assume that this probability density is of the relevant type (2.7.25), this is

$$\rho_0(z) = \rho_0^N \frac{P_0(\hat{A}(z))}{\Omega(\hat{A}(z))} \qquad (4.6.69)$$

implying

$$P_{[\phi]}(a_0, t_0, a_1, t_1) = \frac{P_0(a_0)}{\Omega(a_0)} \int dz \rho_0^N \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\hat{z}^F(z) - a_1) \qquad (4.6.70)$$

The conditional probability of having a_1 at time t_1 given that the macrostate was a_0 at t_0 is

$$P_{[\phi]}(a_0, t_0 | a_1 t_1) = \frac{P(a_0 t_0, a_1 t_1)}{P_0(a_0)} = \frac{1}{\Omega(a_0)} \int dz \rho_0^N \delta(\hat{A}(z) - a_0) \delta(\hat{A}(\hat{z}^F(z) - a_1)$$

$$(4.6.71)$$

We will compare the jump from a_0 to a_1 with the jump from a_1 to a_0. The conditional probability $P_{[\phi]}(a_1, t_0 | a_0, t_1)$ of finding a_0 at time t_1 given that at time t_0 the macrostate was a_1 is, under the action of the *same* protocol ϕ_t,

$$P_{[\phi]}(a_1, t_0 | a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz \rho_0^N \delta(\hat{A}(z) - a_1) \delta(\hat{A}(\hat{z}^F(z) - a_0) \qquad (4.6.72)$$
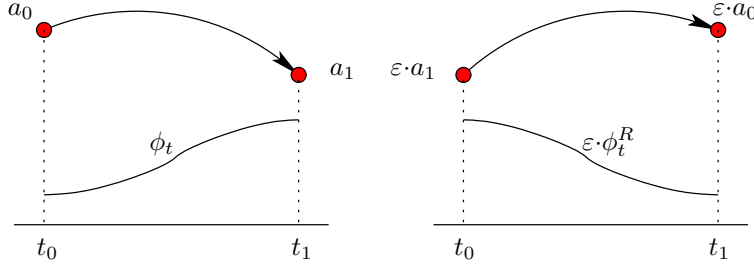
Figure 4.3: The transition probability from macrostate a_0 at time t_0 to macrostate a_1 at time t_1 under the protocol ϕ_t is $P_{[\phi]}(a_0, t_0|a_1, t_1)$ (left panel). The transition probability from $\varepsilon \cdot a_1$ at time t_0 to $\varepsilon \cdot a_0$ at time t_1 under the protocol $\varepsilon \cdot \phi_t^R$ is $P^R(\varepsilon \cdot a_0, t_0|\varepsilon \cdot a_1, t_1)$ (right panel). The ratio between both transition probabilities is given in terms of the ratio of the "number of microstates" of each macrostate. For CG variables which are even under time reversal, $\varepsilon = 1$. If the bare entropies $S_B(a_1) > S_B(a_0)$, the detailed balance theorem says that it is more probable to jump from a_0 to a_1 as shown in the left panel than the reverse jump shown in the righ panel.

If we use now the microscopic reversibility property (4.2.38) in the form

$$\hat{z}^F(z) = \hat{z}^B(z)^{-1}$$
$$\Phi_{[\phi, t_0, t_1]}(z) = \epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R, t_0, t_1]}^{-1}(\epsilon \cdot z) \tag{4.6.73}$$

we have

$$P_{[\phi]}(a_1, t_0|a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz \rho_0^N \delta(\hat{A}(z) - a_1)\delta(\hat{A}(\epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R, t_0, t_1]}^{-1}(\epsilon \cdot z)) - a_0) \tag{4.6.74}$$

Next, perform the change of variables, with unit Jacobian

$$\epsilon \cdot z' = \epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R, t_0, t_1]}^{-1}(\epsilon \cdot z)$$
$$z = \epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R, t_0, t_1]}(z') \tag{4.6.75}$$

we have

$$P_{[\phi]}(a_1, t_0|a_0, t_1) = \frac{1}{\Omega(a_1)} \int dz' \rho_0^N \delta(\hat{A}(\epsilon \cdot \Phi_{[\varepsilon \cdot \phi^R, t_0, t_1]}(z')) - a_1)\delta(\hat{A}(\epsilon \cdot z') - a_0) \tag{4.6.76}$$

Under the assumed parity behaviour of the phase functions $\hat{A}(\epsilon \cdot z) = \varepsilon \cdot \hat{A}(z)$ we arrive at our desired final expression

$$P_{[\phi]}(a_1, t_0|a_0, t_1) = \frac{\Omega(a_0)}{\Omega(a_1)} P_{[\varepsilon \cdot \phi^R]}(\varepsilon \cdot a_0, t_0|\varepsilon \cdot a_1, t_1) \tag{4.6.77}$$

With the bare entropy (3.4.39), this Detailed Balance theorem can be written also in the

form

$$\frac{P_{[\phi]}(a_0, t_0 | a_1, t_1)}{P_{[\varepsilon \cdot \phi^R]}(\varepsilon \cdot a_1, t_0 | \varepsilon \cdot a_0, t_1)} = e^{S_B(a_1) - S_B(a_0)} \tag{4.6.78}$$

This is the theorem of Detailed Balance in the presence of an external forcing. It is a generalization of the Detailed Balance theorem in the absence of external forcing (1.8.38) that we used to discuss the green dog-flea model. The theorem relates the conditional probability of jumping from macrostate a_0 to a_1 to the conditional probability of the reverse jump, occuring under the reversed protocol. It is the reflection of microscopic reversibility at the level of CG variables.

4.7 Summary

5

Near Equilibrium Theory

5.1   Introduction

5.2   Linear irreversible thermodynamics

Linear Irreversible Thermodynamics (LIT), as presented in de Groot and Mazur [?] is a particular version of the general theory presented so far in which one assumes that the dissipative matrix is not dependendent on the state $D_{\mu\nu}(a) = D_{\mu\nu}$. In this way irreversible flows

$$D_{\mu\nu}\frac{\partial S}{\partial a_\nu}(a) \qquad (5.2.1)$$

are *linear* in the thermodynamic forces defined as the derivatives of the entropy.

We expect that near equilibrium, all the state dependence in the dissipative matrix arising from the relevant ensemble will dissapear as the relevant ensemble becomes just the equilibrium ensemble. Therefore, it is expected that (5.2.1) will be valid near equilibrium. An additional assumption taken in de Groot and Mazur textbook [?] in order to obtain many of their results relies on the approximation that the entropy is a quadratic function of its arguments. This assumption can be justified also near equilibrium. Therefore, in this section we assume that the average of the CG variables are close to their equilibrium values at all times, either because we start with an initial value of the CG variables which is itself close to equilibrium, or because we consider the late stages of evolution of a system, when it is already close to equilibrium.

The entropy

In this case, we may expand the entropy function to second order

$$S(a) = S(a^{\mathrm{eq}} + (a - a^{\mathrm{eq}}))$$
$$\simeq S(a^{\mathrm{eq}}) + \frac{\partial S}{\partial a_\mu}(a^{\mathrm{eq}})(a_\mu - a_\mu^{\mathrm{eq}}) + \frac{1}{2}\frac{\partial^2 S}{\partial a_\mu \partial a_\nu}(a^{\mathrm{eq}})(a_\mu - a_\mu^{\mathrm{eq}})(a_\nu - a_\nu^{\mathrm{eq}}) \qquad (5.2.2)$$

The equilibrium values of the averages of the CG variables a^{eq} are the ones that maximize the entropy function as shown in (**??**). The linear term is then absent. On using the identity (2.7.18), this expansion is

$$S(a) = S^{\mathrm{eq}} - \frac{1}{2} \delta a_\mu C^{-1}_{\mu\nu} \delta a_\nu \qquad (5.2.3)$$

where $\delta a_\mu = a_\mu - a^{\mathrm{eq}}_\mu$ and the equilibrium covariance matrix is

$$C_{\mu\nu} = \left\langle \delta \hat{A}_\mu \delta \hat{A}_\nu \right\rangle^{\mathrm{eq}} \qquad (5.2.4)$$

where $\delta \hat{A}_\mu = \hat{A}_\mu - a^{\mathrm{eq}}_\mu$. The derivatives of the entropy, entering the transport equation are

$$\lambda(a) = -C^{-1}_{\mu\nu} \delta a_\nu \qquad (5.2.5)$$

The dissipative matrix

Let us consider now the dissipative matrix (**??**) that we quote again

$$D_{\mu\nu}(a) = \int_0^{\Delta t} dt \, \mathrm{Tr} \left[\overline{\rho}_a \left(\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\nu \right) \exp\{i\mathcal{L}t\} \left(\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu \right) \right] \qquad (5.2.6)$$

This matrix depends on the state a through the relevant ensemble $\overline{\rho}_a(z)$ that appears explicitly and inside the projection operator $\mathcal{Q}_a$. In turn, the relevant ensemble (3.1.4) depends on a through the conjugate variables $\lambda(a)$. If we single out the dynamic invariants that are included in the list of CG variables, the relevant ensemble (3.1.4) takes the form

$$\overline{\rho}_a(z) = \frac{\exp\{-\lambda_0 \hat{A}_0(z) - \lambda_\mu(a)\hat{A}_\mu(z)\}}{\mathrm{Tr}[\exp\{-\lambda_0 \hat{A}_0 - \lambda_\mu(a)\hat{A}_\mu\}]} \qquad (5.2.7)$$

If the energy is the only dynamic invariant we usually write $\lambda_0 \hat{A}_0(z) = \beta \hat{H}(z)$. We see that when the conjugate variables $\lambda_\mu(a)$ of the non-conserved CG variables tend to zero, the relevant ensemble tends to the canonical ensemble. Therefore, it makes sense to study near equilibrium situations by expanding to first order in λ. By taking into account the expansion of the normalization, to first order in λ the relevant ensemble becomes

$$\overline{\rho}_a(z) = \rho^{\mathrm{eq}}(z) \left(1 - \lambda_\mu(a)\delta \hat{A}_\mu \right) + \cdots \qquad (5.2.8)$$

where the equilibrium ensemble is given by the canonical ensemble (**??**). The derivative of the relevant ensemble, needed below, is

$$\frac{\partial}{\partial a_\mu} \overline{\rho}_a(z) = C^{-1}_{\mu\nu} \delta \hat{A}_\nu + \cdots \qquad (5.2.9)$$

where we have used (5.2.5). The adjoint of the Kawasaki-Gunton projector in the dissipative matrix (5.2.6) is the complementary of $\mathcal{P}_t^\dagger$ given in (**??**), that we quote again

$$\mathcal{P}_t^\dagger \hat{F}(z) = \text{Tr}[\overline{\rho}_t \hat{F}] + (\hat{A}(z) - a(t))\text{Tr}\left[\frac{\partial \overline{\rho}_t}{\partial a(t)}\hat{F}\right] \tag{5.2.10}$$

When we use the relevant ensemble (5.2.8) together with (5.2.9) this becomes

$$\mathcal{P}_t^\dagger \hat{F}(z) = F^{\text{eq}} + \left\langle \hat{F}\delta\hat{A}_\mu \right\rangle^{\text{eq}} C_{\mu\nu}^{-1}\delta\hat{A}_\nu(z) \tag{5.2.11}$$

This projector is no longer time dependent. It is known as the **Mori projector** and takes any phase function and converts it into a linear combination of the CG variables. It is straightforward to show that the properties (**??**)-(**??**) are satisfied by the Mori projector.

We need the projected currents $\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu(z)$ in the dissipative matrix (5.2.6). They are

$$\begin{aligned}
\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu(z) &= \mathcal{L}\hat{A}_\mu(z) - \mathcal{P}_a^\dagger i\mathcal{L}\hat{A}_\mu(z) \\
&= \mathcal{L}\hat{A}_\mu(z) - \left\langle \left(i\mathcal{L}\hat{A}_\mu\right)\delta\hat{A}_{\mu'} \right\rangle^{\text{eq}} C_{\mu'\nu'}^{-1}\delta\hat{A}_{\nu'}(z)
\end{aligned} \tag{5.2.12}$$

where we have used that $\left\langle i\mathcal{L}\hat{A}_\mu \right\rangle^{\text{eq}} = 0$. We now introduce the **collective frequencies**

$$\Omega_{\mu\nu} \equiv -C_{\nu\mu'}^{-1}\left\langle \delta\hat{A}_{\mu'} i\mathcal{L}\hat{A}_\mu \right\rangle^{\text{eq}} \tag{5.2.13}$$

in such a way that

$$\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu(z) = \mathcal{L}\hat{A}_\mu(z) + \Omega_{\mu\nu}\delta\hat{A}_\nu(z) \tag{5.2.14}$$

In order to asses the physical content of the last term, consider the reversible drift term (**??**) in the approximation in which the relevant ensemble is (5.2.8)

$$v_\mu(a) = \left\langle \delta\hat{A}_\sigma i\mathcal{L}\hat{A}_\mu \right\rangle^{\text{eq}} C_{\sigma\nu}^{-1}\left(a_\nu(t) - a_\nu^{\text{eq}}\right) = -\Omega_{\mu\nu}\delta a_\nu \tag{5.2.15}$$

Therefore, the projected current is just the fluctuation with respect to the reversible drift evaluated at the instantaneous values of the CG variables,

$$\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu(z) = \mathcal{L}\hat{A}_\mu(z) - v_\mu(\hat{A}) \tag{5.2.16}$$

Finally, the dissipative matrix becomes, to zero order in $\lambda(a)$ (i.e. at equilibrium)

$$D_{\mu\nu} = \int_0^{\Delta t} dt \left\langle \left(i\mathcal{L}\hat{A}_\nu - v_\nu(\hat{A})\right)\exp\{i\mathcal{L}t\}\left(i\mathcal{L}\hat{A}_\mu - v_\mu(\hat{A})\right) \right\rangle^{\text{eq}} \tag{5.2.17}$$

Summary

The general transport equation (**??**) becomes, when the dissipative matrix is approximated by its near equilibrium state independent form, the usual evolution equation of

LIT

$$\frac{d}{dt}a_\mu(t) = v_\mu(a(t)) + D_{\mu\nu}\frac{\partial S}{\partial a_\nu}(a(t)) \tag{5.2.18}$$

where the dissipative matrix is given by (5.2.17). If, in addition, we assume that the entropy is a quadratic function (5.2.3) and that the reversible drift is also very close to equilibrium (5.2.15), the transport equation in the averages flavour becomes the **fully linearized** transport equations governing the evolution of small departures of the averages from its equilibrium value

$$\frac{d}{dt}a_\mu(t) = -\Omega_{\mu\nu}\delta a_\nu(t) - D_{\mu\nu}C_{\mu\nu}^{-1}\delta a_\nu(t) \tag{5.2.19}$$

5.3   Mori theory

In the previous section we have considered the transport equation in the averages flavour in situations near equilibrium. In order to discuss the probabilistic flavour near equilibrium, instead of approximating equations already encountered, we derive from scratch *exact* evolution equations, as we did in Sec 3.5, by using the Mori projector instead of Zwanzig projector. This is useful not only for historical reasons, but also because it gives some results on equilibrium time correlations that cannot be obtained from the non-linear Zwanzig theory.

5.3.1   Exact evolution equation

Our starting point are the two equations (3.5.3) and (3.5.4), that combined give

$$\partial_t \hat{\Psi}_a(\mathcal{U}_t z) = \exp\{i\mathcal{L}t\}\mathcal{P}^\dagger i\mathcal{L}\hat{\Psi}_a(z)$$
$$+ \int_0^t dt' \exp\{i\mathcal{L}t'\}\mathcal{P}^\dagger i\mathcal{L}\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a(z)$$
$$+ \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger i\mathcal{L}\hat{\Psi}_a(z) \tag{5.3.1}$$

By multiplying this equation with a_μ and integrating over a, we obtain an evolution equation for the CG variables themselves,

$$\frac{d}{dt}\hat{A}_\mu(\mathcal{U}_t z) = \underbrace{\exp\{i\mathcal{L}t\}\mathcal{P}^\dagger i\mathcal{L}\hat{A}_\mu(z)}_{(1)} + \underbrace{\int_0^t dt' \exp\{i\mathcal{L}t'\}\mathcal{P}^\dagger i\mathcal{L}\exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu(z)}_{(2)}$$
$$+ \underbrace{\exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu(z)}_{(3)} \tag{5.3.2}$$

This equation is valid for any projection operator $\mathcal{P}^\dagger$. We will use Mori's projector (5.2.11). In the probabilistic flavour the equilibrium average is consistently defined in terms of the microcanonical equilibrium ensemble instead of the canonical ensemble, although this is not essential in the thermodynamic limit. By using (5.2.14) we have for

the first term in (5.3.2)

$$(1) = -\Omega_{\mu\nu} \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\delta\hat{A}_\mu(z) = -\Omega_{\mu\nu}\delta\hat{A}_\mu(\mathcal{U}_t z) \tag{5.3.3}$$

The second memory term is, upon using the Mori projector for $\mathcal{P}^\dagger$

$$\begin{aligned}
(2) &= \int_0^t dt' \exp\{i\mathcal{L}t'\}\delta\hat{A}_\nu(z)C_{\nu\nu'}^{-1} \left\langle \delta\hat{A}_{\nu'} i\mathcal{L} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu \right\rangle^{\mathrm{eq}} \\
&= -\int_0^t dt' \exp\{i\mathcal{L}t'\}\delta\hat{A}_\nu(z)C_{\nu\nu'}^{-1} \left\langle i\mathcal{L}\hat{A}_{\nu'}\mathcal{Q}^\dagger \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu \right\rangle^{\mathrm{eq}} \\
&= -\int_0^t dt' \exp\{i\mathcal{L}t'\}\delta\hat{A}_\nu(z)C_{\nu\nu'}^{-1} \left\langle \mathcal{Q}^\dagger i\mathcal{L}\hat{A}_{\nu'} \exp\{\mathcal{Q}^\dagger i\mathcal{L}(t-t')\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu \right\rangle^{\mathrm{eq}} \\
&= -\int_0^t dt' \delta\hat{A}_\nu(\mathcal{U}_{t'} z)C_{\nu\nu'}^{-1} \left\langle \hat{F}_{\nu'} \hat{F}_\mu(t-t') \right\rangle^{\mathrm{eq}} \tag{5.3.4}
\end{aligned}$$

where we have used the self-adjoint property (1.4.8) of $i\mathcal{L}$ and the fact that $\mathcal{Q}^\dagger$ is also self-adjoint with respect to the equilibrium ensemble. In the last identity we have also introduced the random force

$$(3) = \hat{F}_\mu(z,t) \equiv \exp\{\mathcal{Q}^\dagger i\mathcal{L}t\}\mathcal{Q}^\dagger i\mathcal{L}\hat{A}_\mu \tag{5.3.5}$$

In summary, with the relevant ensemble of the form (5.2.8), one obtains an exact equation for the instantaneous fluctuations with respect to equilibrium, which is given by

$$\frac{d}{dt}\delta\hat{A}_\mu(\mathcal{U}_t z) = -\Omega_{\mu\nu}\delta\hat{A}_\nu(\mathcal{U}_t z) - \int_0^t dt' K_{\mu\nu}(t-t')\delta\hat{A}_\nu(\mathcal{U}_{t'} z) + \hat{F}_\mu(z,t) \tag{5.3.6}$$

where we have introduced the memory kernel

$$K_{\mu\nu}(t) = \left\langle \hat{F}_\mu^\dagger(t)\hat{F}_{\nu'}^\dagger \right\rangle^{\mathrm{eq}} C_{\nu'\nu}^{-1} \tag{5.3.7}$$

This equation is linear in the deviations of the CG variables with respect its equilibrium average. Note that this equation is exact, and mathematically equivalent to the exact non-linear GLE (3.5.19), obtained with Zwanzig's projector. They are just two particular rewritings of Hamilton's equations. Being a linear function one would expect that (5.3.6) is simpler to treat. However, note that the whole point is whether the memory kernel is short lived in comparison with the dynamic of the CG variables themselves in order to use a Markovian approximations. The random force $\hat{F}_\mu(z,t)$ satisfies

$$\begin{aligned}
\left\langle \hat{F}_\mu(t) \right\rangle^{\mathrm{eq}} &= 0 \\
\left\langle \hat{F}_\mu(t)\hat{A}_\nu \right\rangle^{\mathrm{eq}} &= 0 \qquad\qquad t \geq 0
\end{aligned} \tag{5.3.8}$$

This implies that the average over an initial ensemble which is of the relevant type (5.2.8) also vanishes.

The evolution equation (5.3.6) is not closed in $\delta\hat{A}_\mu$ because of the presence of the

random force. If we average (5.3.6) over an ensemble of initial conditions that is of the relevant type (5.2.8) and use the fact that for the relevant ensemble near equilibrium we have

$$\text{Tr}\left[\overline{\rho}_a \delta \hat{A}_\mu\right] = \delta a_\mu(t) \tag{5.3.9}$$

we obtain the following closed equation for the deviations of the average $a(t)$ with respect to its equilibrium value

$$\frac{d}{dt}\delta a_\mu(t) = -\Omega_{\mu\nu}\delta a_\nu(t) - \int_0^t dt' K_{\mu\nu}(t-t')\delta a_\nu(t') \tag{5.3.10}$$

which is a linear closed equation for the average deviations. Because the averages are computed with an initial ensemble of the relevant type close to equilibrium, in this equation we are implicitly assuming that the system, on average, is close to equilibrium at all times.

Because the Mori theory is a linear theory we may obtain a closed equation for equilibrium correlations[1]. If we multiply (5.3.6) with $\hat{A}_\nu$ and average over the equilibrium ensemble, we obtain

$$\frac{d}{dt}C_{\mu\nu}(t) = -\Omega_{\mu\nu'}C_{\nu'\nu}(t) - \int_0^t dt' K_{\mu\nu'}(t-t')C_{\nu'\nu}(t') \tag{5.3.11}$$

where the equilibrium time correlation is defined as

$$C_{\mu\nu}(t) = \left\langle \delta\hat{A}_\mu(t)\delta\hat{A}_\nu \right\rangle^{\text{eq}} \tag{5.3.12}$$

Note that the matrix of covariances introduced in (5.2.4) is given by

$$C_{\mu\nu}(t) = C_{\mu\nu}(0) \tag{5.3.13}$$

---

[1]This is not feasible from the Zwanzig non-linear approach.

5.3.2 Markov approximation

Eq (5.3.10) is an exact equation, for which we expect that a Markov approximation will render into the form (5.2.19). We may follow identical steps as done in the Zwanzig theory in Sec 3.5.4 and approximate the memory kernel (5.3.7) with a Dirac delta function

$$K_{\mu\nu}(t) \quad \approx \quad D_{\mu\nu'}C_{\nu'\nu}^{-1}\delta^+(t) \tag{5.3.14}$$

where the dissipative matrix $D_{\mu\nu}$ is introduced as

$$D_{\mu\nu} \quad \equiv \quad \int_0^{\Delta t} \langle F_\mu^\dagger(t)F_\nu^\dagger\rangle dt \tag{5.3.15}$$

Under the Markovian approximation (5.3.14) the two equations (5.3.10), (5.3.11) for the average deviation from equilibrium and correlations, respectively, become

$$\frac{d}{dt}\delta a(t) = -\Omega{\cdot}\delta a(t) - D{\cdot}C^{-1}{\cdot}\delta a(t)$$

$$\frac{d}{dt}C(t) = -\Omega{\cdot}C(t) - D{\cdot}C^{-1}{\cdot}C(t) \tag{5.3.16}$$

The solution of these linear equations is given by the exponential matrix form. For example,

$$C(t) = \exp\{-\Lambda t\}{\cdot}C(0) \tag{5.3.17}$$

where the exponential matrix is given in terms of its Taylor series expansion. We have introduced the matrix Λ

$$\Lambda = \Omega + D{\cdot}C^{-1} \tag{5.3.18}$$

Eq (5.3.17) is a very important result as it shows that sufficiently close to equilibrium, the validity of the Markovian approximation relies on the validity of an exponential decay (in matrix sense) of the equilibrium correlations. This offers a way to validate, in practical applications, the assumption that the selected CG variables are, in fact, Markovian.

Note that the solution (5.3.17) cannot be valid at $t = 0$ or, in general at times $t < \Delta t$. To see this, take the time derivative at $t = 0$ which is given by

$$\frac{dC}{dt}(0) = -\Omega{\cdot}C(0) - D \tag{5.3.19}$$

However, by using the microscopic definition of the correlation function we know that

$$\frac{d}{dt}C_{\mu\nu}(0) = \left\langle \delta\hat{A}_\nu i\mathcal{L}\hat{A}_\mu \right\rangle^{\text{eq}} = -\Omega_{\mu\mu'}C_{\mu'\nu}(0) \tag{5.3.20}$$

where we have used the definition of the collective frequencies (5.2.13). Eqs (5.3.19), (5.3.20) implies that $D = 0$, which makes no sense. This indicates that the solution (5.3.17) is valid only after a time Δt has been elapsed.

5.3.3   Onsager reciprocity

The Liouville operator satisfies, for arbitrary phase functions $\hat{F}, \hat{G}$

$$\left\langle \hat{F} i\mathcal{L} \hat{G} \right\rangle^{\mathrm{eq}} = - \left\langle \hat{G} i\mathcal{L} \hat{F} \right\rangle^{\mathrm{eq}}$$
$$\left\langle \hat{F} \exp\{i\mathcal{L}t\} \hat{G} \right\rangle^{\mathrm{eq}} = \left\langle \hat{G} \exp\{-i\mathcal{L}t\} \hat{F} \right\rangle^{\mathrm{eq}} \tag{5.3.21}$$

The first property implies that the matrix L in (**??**) is antisymmetric and the second one implies that the correlation matrix $C(t)$ in (**??**) satisfies the stationarity condition

$$C(t) = C^T(-t) \tag{5.3.22}$$

In addition, the time reversal symmetry of Hamilton's equation implies [**?**]

$$C(t) = \epsilon C(-t)\epsilon \tag{5.3.23}$$

where ϵ is a diagonal matrix containing ± 1 in the diagonal depending on the time reversal symmetry of the corresponding CG variable. The combination of time reversal and time invariance gives

$$C(t) = \epsilon C^T(t)\epsilon \tag{5.3.24}$$

By deriving both sides of (5.3.23) with respect to time we have

$$\dot{C}(t) = -\epsilon \dot{C}(-t)\epsilon \tag{5.3.25}$$

When evaluated at $t = 0$ this gives,

$$C(0) = \epsilon C^T(0)\epsilon$$
$$\dot{C}(0) = -\epsilon \dot{C}(0)\epsilon \tag{5.3.26}$$

Onsager's reciprocity is obtained by using the time reversal property (5.3.24) together with the exponential solution (5.3.17)

$$\exp\{-\Lambda t\} {\cdot} C(0) = \epsilon {\cdot} C^T(0) {\cdot} \exp\{-\Lambda^T t\} {\cdot} \epsilon \tag{5.3.27}$$

Taking the time derivative of both sides and evaluating at $t = 0$ gives

$$\Lambda {\cdot} C(0) = \epsilon {\cdot} C^T(0) {\cdot} \Lambda^T {\cdot} \epsilon \tag{5.3.28}$$

Eqs (5.3.18) and (5.3.20) imply

$$\Lambda {\cdot} C(0) = D - \dot{C}(0) \tag{5.3.29}$$

and then (5.3.30) gives

$$D - \dot{C}(0) = \epsilon {\cdot} [D - \dot{C}(0)]^T {\cdot} \epsilon \tag{5.3.30}$$

By using the second equation (5.3.26) and the fact that $\dot{C}^T(0) = -\dot{C}(0)$, as can be seen from the antihermiticity of $i\mathcal{L}$ in (5.3.20) we finally conclude

$$D = \epsilon D^T \epsilon \qquad (5.3.31)$$

which is the Onsager reciprocal relations for the dissipative matrix D.

5.3.4   Mori's theory is a particular case of Zwanzig's theory

We now show that the Markovian SDE of Mori theory can be obtained from the SDE obtained in Zwanig's theory in the limit that the initial conditions x_0 of the SDE are close to x^{eq} that maximizes the equilibrium probability distribution $P^{\text{eq}}(x)$ and, in addition, fluctuations are small in such a way that the equilibrium probability is highly peaked, i.e. $P^{\text{eq}}(x) = \delta(x - x^{\text{eq}})$.

5.4   Summary

6

Summary of the Theory of Coarse-Graining

6.1 The microscopic level

6.2 The averages flavour

6.3 The probabilistic flavour

[Present the results for external forcing.]

The basic assumption is that the CG variables have a two-scale evolution, one on a short time scale τ_{mic} stemming from "rapid and small events" that cummulative build the variation on a larger scale τ_{mac}. When this separation of time scales exists, one can model the dynamics of the CG variables with a stochastic diffusion process that is governed with a Fokker-Planck Equation

$$\partial_t P(a,t) = -\frac{\partial}{\partial a_\mu}\left[V_\mu(a) + D_{\mu\nu}(a)\frac{\partial S_B}{\partial a}(a)\right]P(a,t) + \frac{\partial}{\partial a_\mu}D^S_{\mu\nu}(a)\frac{\partial}{\partial a_\nu}P(a,t) \quad (6.3.1)$$

where $D^S(a) = (D(a) + D^T(a))/2$ is the symmetric part of the dissipative matrix. Here $P(a,t)$ is the probability distribution of CG variables. The reversible drift is

$$V_\mu(a) = \left\langle i\mathcal{L}\hat{A}_\mu\right\rangle^a \quad (6.3.2)$$

where the conditional expectation of an arbitrary phase function $\hat{F}(z)$ is

$$\langle\hat{F}\rangle^a = \frac{1}{\Omega(a)}\int dz\,\hat{\Psi}_a(z)\hat{F}(z) \quad (6.3.3)$$

where the measure of the CG variables is

$$\Omega(a) = \int dz \,\hat{\Psi}_a(z) \tag{6.3.4}$$

Its logarithm is the generalized Boltzmann's entropy

$$S_B(a) \equiv \ln \Omega(a) \tag{6.3.5}$$

The drift term (6.3.2) takes also the GENERIC form

$$V_\mu(a) = L_{\mu\nu}(a) \frac{\partial E}{\partial a_\nu}(a) \tag{6.3.6}$$

where the CG energy $E(a)$ is obtained from the microscopic Hamiltonian as

$$\hat{H}(z) = E(\hat{A}(z)) \tag{6.3.7}$$

and the reversible operator is defined as

$$L_{\mu\nu}(a) = \left\langle \{\hat{A}_\mu, \hat{A}_\nu\} \right\rangle^a \tag{6.3.8}$$

The reversible satisfies the reversibility condition

$$V_\mu(a) \frac{\partial S}{\partial a_\mu}(a) = -\frac{\partial V_\mu}{\partial a_\mu}(a) \tag{6.3.9}$$

closely related to the (quasi) degeneracy condition of the reversible operator $L(a)$.

$$L_{\mu\nu}(a) \frac{\partial S_B}{\partial a_\nu} = -\frac{\partial}{\partial a_\nu} L_{\mu\nu}(a) \tag{6.3.10}$$

The dissipative matrix is

$$D_{\mu\nu}(a) = \int_0^{\Delta t} d\tau \left\langle (i\mathcal{L}\hat{A}_\nu - V_\nu(a))) \exp\{-i\mathcal{L}\tau\} (i\mathcal{L}\hat{A}_\mu - V_\mu(a)) \right\rangle^a \tag{6.3.11}$$

The dissipative matrix satisfies Onsager's reciprocity

$$D_{\mu\nu}(a) = \varepsilon_\mu \varepsilon_\nu D_{\mu\nu}(\varepsilon \cdot a) \tag{6.3.12}$$

The entropy of a probabilistic level of description is

$$S[P] = -\int da P(a) \ln \frac{P(a)}{\Omega(a)} \tag{6.3.13}$$

and it satisfies an H-theorem $\frac{d}{dt} S[P_t] \geq 0$. The H-Theorem leads to the unique equilib-

rium solution given by the Einstein formula

$$P^{\text{eq}}(a) = \frac{P_0(E(a)) \exp\{S(a)\}}{\int da' \exp\{S(a')\} \delta(E(a') - E)} \tag{6.3.14}$$

where $P_0(E)$ is the distribution of the energy in the system at the initial time.

The SDE which is mathematically equivalent to the FPE (6.3.1) is

$$da = V(a)dt + D(a) \cdot \frac{\partial S_B}{\partial a}(a)dt + \frac{\partial \cdot D^S}{\partial a}(a)dt + d\tilde{a} \tag{6.3.15}$$

The random force $d\tilde{a}$ is a linear combination of Wiener processes that satisfy the Fluctuation-Dissipation theorem

$$d\tilde{a}d\tilde{a}^T = 2dt D^S(a) \tag{6.3.16}$$

The random noise satisfy the orthogonality condition

$$d\tilde{a}^T \cdot \frac{\partial E}{\partial a} = 0 \tag{6.3.17}$$

7

Challenges and limitations of the Theory of Coarse-Graining

The theory of coarse-graining as presented in this book starts with the basic assumption that reality is made of atoms, governed by Classical Mechanics. Everything is made of atoms interacting with each other with relatively simple laws. However, out of this simple interactions emerge an incredible and complex organization in terms of molecules and crystals that self-assemble in ever increasing structures (from simple isolated molecules to biological cells that, in turn aggregate to form complex organisms that are alive). It is assumed that reality may be split in pieces of reality which are independent of each other, called "systems". We call different systems with different names (fluid, glass, membrane, polymer molecule, etc.) because they behaviour is distinctly different. Otherwise we would not distinguish them. But what is "behaviour"? Different behaviour give different readings of certain variables (typically forces, motion, textures, or intensity and frequency of light) that are recorded with measurement apparatus, including our own sensory system.

One of the basic assumptions of the theory is that if an isolated system governed with a Hamiltonian is left alone, it will relax towards a unique equilibrium state. This gives the whole structure of the dynamic equations that ensure the Second Law.

While this is adequate for some systems, like a simple fluid or gas, it is questionable in a vast majority of systems in this universe. Granular systems, for example, already display the difficulty of the present approach. A granular material contained in a silo is all made of atoms that organize themselves into macroscopic grains. Depending on how the silo was filled, it will adopt a configuration of the grains that will produce vastly varying forces on the walls of the silo. Stress is transmitted through the grains forming paths that vary from one otherwise identically filled silo to another. There is no such a "unique equilibrium state" towards which the system relaxes. For a particular filling of the silo, the atoms in all grains may eventually relax towards thermal equilibrium corresponding to a particular equilibrium state. This "equilibrium state" can be distinguished macroscopically from other equally valid equilibrium states, but with different mechanical properties. Identical preparation of the system (filling of the silo) leads to very different equilibrium states. Another example is a glass system, where a collection

---

Margin notes:

Reality = interacting atoms
Atoms organize

Distinguishing behaviour / structure

Broken ergodicity

of atoms suddenly cooled down adopt a "frozen" configuration that cannot be reached dynamically from some other microstates with the same energy. When ergodicity is broken in a system, the theory of coarse-graining as presented here cannot be used.

[No jumps. No discrete Markov jumps.]

7.1   Problems in computing the building blocks

Even in the case that the system evolves towards a unique equilibrium state, it usually does it in different distinct time scales. This is precisely what allows one to introduce different levels of description for a given system, each one characterized with its typical time scale. However, the theory describes the link of each of these levels of description with the microscopic level described with Hamiltonian dynamics. The transport coefficients are given by Green-Kubo formulae in terms of the time integral of correlations of projected currents. These projected currents are explicit phase functions and can be, in principle, computed with MD simulations. However, for most of the levels of descriptions except the "closer to microscopic" ones, the time scale of the projected currents is much, much larger than molecular collision times. This means that in order to compute faithfully the correlation of the projected currents we need to simulate the system for times that cannot be reached with MD simulations.

GK correlations not computable with MD

One possibility to address this problem is to proceed in a sequence of jumps. For this strategy to be applicable at all, we need to proceed in the probabilistic flavour. In this way, we may construct a first jump in which we derive the Fokker-Planck equation for the (still accessible with MD) first level of description. Let us call this first Fokker-Planck equation FPE-1. Once we have the functional form of this FPE-1 we may perform a second jump from this FPE-1 to the FPE-2 governing the dynamics of coarser variables. Obviously, in order to be able to do this jump, we need to have the explicit mapping between the CG variables of level 2 as functions of the CG variables of level 1. The drift and diffusion of FPE-2 can be obtained from stochastic simulations of the FPE-1 (but probably not from MD simulations). Recall that for this program to be feasible, it is necessary that the system has unique equilibrium state, i.e. it needs to be ergodic at the microscopic level, because any FPE has a unique equilibrium state. If the original system does not have a unique state, trying to model it with a FPE with a unique equilibrium seems to be not a very successful strategy.

Solution: jump from CG to CG

Note that this strategy relies only on the following assumptions:

- The CG dynamics is given by a FPE.

- The CG variables are explicit functions of FG variables.

- The FG variables evolve with a computable FPE.

Under these assumptions, we may explicitly compute the drift and diffusion tensor of the FPE-2 with feasible simulations of the FPE-1 through, for example, their Krames-Moyal expressions. **[We have to develop the relative entropy estimation of the drift and diffusion when the underlying microscopic dynamics is not deterministic as in Liouville, but stochastic, as in FPE]**

How to compute the terms from microscopics

Of course, the transport equation will be useful if and only if all the objects $v(a), D(a)$ and $\lambda(a) = \frac{\partial S}{\partial a}(a)$ appearing in it are known functions of the CG variables a. As the explicit form of the CG variables $\hat{A}(z)$ is known, we have also the explicit form of $i\mathcal{L}\hat{A}(z)$ and this gives most of the structure of the drift $v(a)$ and dissipative matrix $D(a)$. However, at the end of the day we still need to perform averages with respect to the relevant ensemble. This poses a severe problem when the number of CG variables $\hat{A}$ is large. In this case, the functions $v(a), D(a), \lambda(a)$ are functions in a multidimensional space. It is obvious that, even for a moderate number of CG variables, in order to explore the functional form of we would need to run an enormous number of simulations in order to sample the multidimensional state space where a lives. This problem is known as **the curse of the dimensionality**.

Even though now we could in principle use molecular dynamics to compute the correlations, we still need to face the problem of sampling the initial conditions according to the relevant ensemble and average over these initial conditions. As we will see in specific examples, when the system evolves towards equilibrium, the relevant ensemble tends towards the equilibrium ensemble. Therefore, *near equilibrium situations* we may approximate

$$D_{\mu\nu}(a(t)) \approx D_{\mu\nu}(a^{\mathrm{eq}}) = \int_0^\infty d\tau' \mathrm{Tr}\left[\rho_{\mathrm{eq}}(\mathcal{Q}_{eq}^\dagger i\mathcal{L}\hat{A}_\nu) \exp\{-i\mathcal{L}(\phi_t)\tau'\}(\mathcal{Q}_{eq}^\dagger i\mathcal{L}\hat{A}_\mu)\right] \quad (7.1.1)$$

The right hand side of this equation can be explicitly computed in a molecular dynamics simulation. Eq. (7.1.1) is the celebrated Green-Kubo formula for the transport coefficients of a given level of description. Note that in this near equilibrium approximation the dissipative matrix is no longer state-dependent, this is, it does not depend on $a(t)$ anymore. More specifically, it only depends on the dynamical invariants that characterize the equilibrium ensemble.

Of course, this problem is not that severe if we take the near equilibrium approximation. In this case, we only need to sample a much reduced state space, that of determined by the macroscopic thermodynamic parameters characterizing the macroscopic equilibrium (generally the total density and temperature, i.e. a two dimensional space state).

Precompute is difficult.

The Equation free method and the Heterogeneous Multiscale Method

Typically, a problem is called multiscale when the dynamics of the CG variables is coupled to dynamics of small scales. This is bla-bla. What Multiscale really means is that your CG model does not represent some regions of the system. For example, the crack tip or the contact line, or aggregation of colloids in flow. Then one needs to supply the values of the CG directly from the MD simulation (or any other fine scale method). The solution to address multiscale problems is, again, the Equation Free approach or the Heterogeneous multiscale model.

We need a validity test of the Markovian approximation. Can we use Bayesian inference?

Many system cannot be described with the theory of coarse-graining

- Glasses: There is no ergodic and no equilibrium state **[FALSE!]**

- Granular systems. Arches form, is not an ergodic problem.

- Active matter and galaxies. Active matter has an inner fuel that liberates energy and one needs to account for this. Galaxies are not easy to coarse-grain because energy is disappearing with radiation and appearing due to explosion of stars.

- Turbulence

- **Master equations**. We do not consider jumps, sorry.

Assets:

- Universal structure: Reversible and Irreversible dynamics
- Irreversible: "thermodynamic forces" give "fluxes"
- Not limited to near equilibrium
- Second Law
- Fluctuation-Dissipation Theorem
- Explicit connection of CG model $S(a), v(a), M(a)$ with MD.

Problems:

- You need to know the CG variables $A(z)$
- They need to be Markovian.
- The curse of dimensionality.
- Sometimes the projected currents are so long lived that MD is impractical.

What if not?

Problem: You need to know the CG variables $A(z)$.
Strategies:

- For small number of variables: machine learning tools to detect slow manifolds (Kevrekidis)

- For large number (≥ 3) of variables: **Think further! Trial and error.** Propose candidates, and then *validate* the resulting theory Farrell, Oden, Faghihi, J. Comp. Phys. (2015)

Problem: The CG variables need to be Markovian.
Strategies:

- Modeling non-Markovian is usually a no go because you need to run long MD simulations. However, think about using non-Markov without separation of time scales: example of a molecule that instead of using 20 atoms you use 3. This is, essentially the task of designing force fields. What hamiltonian of less variables gives the best signal for CG variables of the full Hamiltonian.

- **Think further!** Enlarge the set of CG variables.

Problem: How to face the curse of dimensionality? $S(a), v(a), M(a)$ are functions in $M \gg 1$ dimensional space.

- Equation free (Kevrekidis & Samaey, Annu. Rev. Phys. Chem. 2009. 60:321–44) and HMM W. E et al. / Commun. Comput. Phys., 2 (2007), pp. 367-450: Compute only for those a that are actually visited by the dynamics. Short burst (because of Markov!).

- Precompute: Propose parametric models for $S(a; \theta), v(a; \theta), M(a; \theta)$ and optimize θ to best fit the dynamics. Español and Zúñiga Phys. Chem. Chem. Phys. (2011), Dequist and Canchaya, J. Chem. Phys. (2015) Machine learning is here!

Problem: **Memory already too large for MD**
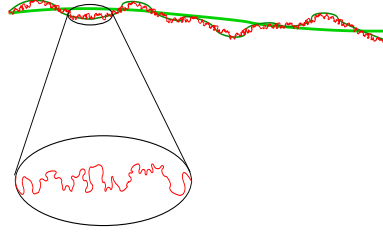
Strategy:

Extract "micro" information from intermediate levels of description PE, Phys. Rev. E (2009)

7.2 The reversible drift and the dissipative matrix from MD simulations

The drift v_μ and the dissipative matrix $D_{\mu\nu}$ in (**??**) depend on the CG variables because they are computed with the relevant ensemble $\overline{\rho}_a$ that depends implicitly on the relevant variables through the conjugate variables. When we explicitly display the dynamic invariants in the relevant ensemble, it takes the form

$$\overline{\rho}(z) = \frac{\rho_N^0}{Z(\alpha, \beta, \lambda)} \exp\{-\alpha \hat{N}(z) - \beta \hat{H}(z) - \lambda \cdot \hat{A}(z)\} \qquad (7.2.1)$$

where we have introduced α, β as the conjugate variables of the number of particles $\hat{N}(z)$ and the Hamiltonian $\hat{H}(z)$, and λ as the conjugate variables of the rest of variables $\hat{A}(z)$. The relevant ensemble can be written in the following form

$$\overline{\rho}(z) = \rho_{\alpha\beta}^{\text{mac}}(z) \frac{\exp\{-\lambda \cdot \hat{A}(z)\}}{\left\langle \exp\{-\lambda \cdot \hat{A}\} \right\rangle^{\alpha\beta}} \qquad (7.2.2)$$

where the macrocanonical equilibrium ensemble is given in (**??**) and $\langle \cdots \rangle^{\alpha\beta}$ is the average with respect to the macrocanonical equilibrium ensemble. Note that the conjugate variables α, β, λ are adjusted to ensure that the relevant ensemble (7.2.1) gives the correct values of the averages $N, E, a(t)$. This condition, first given in (**??**), takes in the time-dependent situation the form

$$\frac{\partial \Phi}{\partial \alpha}(\alpha(t), \beta(t), \lambda(t)) = N$$

$$\frac{\partial \Phi}{\partial \beta}(\alpha(t), \beta(t), \lambda(t)) = E$$

$$\frac{\partial \Phi}{\partial \lambda}(\alpha(t), \beta(t), \lambda(t)) = a(t) \qquad (7.2.3)$$

or, alternatively in terms of the entropy, as in (**??**) but now evaluated at the time-dependent averages

$$\frac{\partial S}{\partial N}(N, E, a(t)) = \alpha(t)$$

$$\frac{\partial S}{\partial E}(N, E, a(t)) = \beta(t)$$

$$\frac{\partial S}{\partial a}(N, E, a(t)) = \lambda(t) \tag{7.2.4}$$

While the averages N, E of the dynamic invariants are time independent, this does not ensure that the corresponding conjugate variables α, β are time independent themselves. In principle, α, β are functions of the time-dependent averages $a(t)$. We need to solve the set of equation (7.2.3) in order to extract $\alpha(t), \beta(t)$ and this may be seen as inconvenient from a practical point of view, given that E, N are conserved anyway.

In addition, the macrocanonical ensemble is not sampled directly with an MD simulation. In order to deal with MD situations, that samples the microcanonical ensemble, we would like to have a relevant ensemble like (7.2.2) but with an equilibrium microcanonical ensemble instead. This suggest that instead of using an average flavour in which the average values N of the number of particles and E of the energy plus the averages A of the rest of CG variables, we should use a flavour in which the macroscopic information is given by the probability distribution $P(E)$ of the dynamic invariants and the averages A of the CG variables. We have discussed in Sec. 2.7.3 what is the relevant ensemble in this flavour, given by (2.7.35)

$$\overline{\rho}(z) = \sum_{N'=0}^{\infty} \int dE P(N', E) \rho_{N'E}^{\mathrm{mic}}(z) \frac{\exp\left\{-\lambda \cdot \hat{A}(z)\right\}}{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\}\right\rangle^{NE}} \tag{7.2.5}$$

where the microcanonical equilibrium ensemble is given by

$$\rho_{N'E}^{\mathrm{mic}}(z) = \frac{\rho_N^0 \delta_{NN'} \delta(\hat{H}(z) - E)}{\Omega(E)} \tag{7.2.6}$$

and the average $\langle \cdots \rangle^{N'E}$ is a microcanonical equilibrium average with (7.2.6).

Observe that when $P(E) = \delta_{NN_0} \delta(E - E_0)$ we describe the common situation encountered in MD simulations in which we know with certainty that the system has N particles and has energy E. The relevant ensemble (7.2.5) becomes in this case

$$\overline{\rho}(z) = \rho_{NE}^{\mathrm{mic}}(z) \frac{\exp\left\{-\lambda \cdot \hat{A}(z)\right\}}{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\}\right\rangle^{NE}} \tag{7.2.7}$$

This ensemble looks formally similar to the relevant ensemble (7.2.2) obtained under the macroscopic information given by the averages N, E, A, in which the only difference would be the change of the equilibrium ensemble. However, the two relevant ensembles (7.2.1)

and (7.2.7) *are not equivalent*, despite the fact that averages of volume independent phase functions with the microcanonical and macrocanonical ensembles are equivalent when the equilibrium entropy is extensive an there exists a well-defined thermodynamic limit. The reason is that the function $\exp\left\{-\lambda\cdot\hat{A}(z_N)\right\}$ typically will scale as e^N if the entropy of the level of description is extensive. As a consequence, it is not correct to use the Laplace method as we did in (**??**), because the function g in (**??**) depends exponentially on N.

Because we are interested in MD simulations, from now on we will assume that the relevant ensemble in the averages flavour takes the form (7.2.7). Note that all properties discussed so far for the dissipative matrix (positive definite, Onsager reciprocity, etc.) are still satisfied with this ensemble. Another clear advantage of the relevant ensemble (7.2.7) over (7.2.2) is that now, we do not need to solve for the two time dependent parameters $\alpha(t), \beta(t)$, whose role is played by the dynamic invariants N, E that are known in a given MD simulation.

The average of an arbitrary phase function $\hat{F}(z)$ with the relevant ensemble (7.2.7) has the form **[What I have learned from simulations of chemical reactions is that for finite samples, the exponential function of any unlikely fluctuation does actually dominate the average. Therefore, it makes no sense whatsoever to evaluate the average of the exponential functions.]**

$$\mathrm{Tr}[\overline{\rho}\hat{F}] = \frac{\left\langle\exp\left\{-\lambda\cdot\hat{A}\right\}\hat{F}\right\rangle^{NE}}{\left\langle\exp\left\{-\lambda\cdot\hat{A}\right\}\right\rangle^{NE}} \tag{7.2.8}$$

The average of the relevant variable $\hat{A}(z)$ is precisely A and, therefore we have the identity

$$A = \frac{\left\langle\exp\left\{-\lambda\cdot\hat{A}\right\}\hat{A}\right\rangle^{NE}}{\left\langle\exp\left\{-\lambda\cdot\hat{A}\right\}\right\rangle^{NE}} \tag{7.2.9}$$

This condition can also be expressed in the form

$$A = \frac{\partial\Phi(N, E, \lambda)}{\partial\lambda} \tag{7.2.10}$$

where the thermodynamic potential is

$$\Phi(N, E, \lambda) = -\ln\left\langle\exp\left\{-\lambda\cdot\hat{A}\right\}\right\rangle^{NE} \tag{7.2.11}$$

This thermodynamic potential is a concave function of λ and, therefore, for fixed N, E the identity (7.2.9) gives the one to one connection between the conjugate parameters λ and the averages A. In this sense, λ is a function of N, E, A.

Let us now consider the drift and dissipative matrix given in (**??**). By inserting the

relevant ensemble (7.2.7) they become

$$v_\mu(a) = \frac{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\} i\mathcal{L}\hat{A}_\mu \right\rangle^{NE}}{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\} \right\rangle^{NE}}$$

$$D_{\mu\nu}(a) = \int_0^T d\tau \frac{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\} \left(\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\nu\right) \exp\{i\mathcal{L}\tau\} \left(\mathcal{Q}_a^\dagger i\mathcal{L}\hat{A}_\mu\right) \right\rangle^{NE}}{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\} \right\rangle^{NE}} \qquad (7.2.12)$$

The form of the projector

The only bit that remains is the specification of the projector $\mathcal{P}_a^+$ appearing in the dissipative matrix. Recall that the macroscopic information available in this flavour is

$$A = \sum_{N'=0}^\infty \int dz_{N'}\overline{\rho}(z_{N'})\hat{A}(z_{N'})$$

$$P(E) = \sum_{N'=0} \int dz_{N'}\overline{\rho}(z_{N'})\hat{\Psi}_{NE}(z_{N'}) \qquad (7.2.13)$$

where we have introduced the phase function

$$\hat{\Psi}_{NE}(z_{N'}) = \delta_{NN'}\delta(\hat{H}(z_{N'}) - E) \qquad (7.2.14)$$

With these variables, we may reconsider the form of the projector (**??**) in the present description

$$\mathcal{P}_t^\dagger \hat{F}(z) = \mathrm{Tr}[\overline{\rho}_t\hat{F}] + (\hat{A}(z) - a(t))\frac{\partial}{\partial a(t)}\mathrm{Tr}\left[\overline{\rho}_t\hat{F}\right]$$

$$+ \int dE(\hat{\Psi}_{NE}(z) - P(E))\frac{\delta}{\delta P(E)}\mathrm{Tr}\left[\overline{\rho}_t\hat{F}\right] \qquad (7.2.15)$$

$$\mathcal{P}_t^\dagger \hat{F}(z) = \int dE P(E)\frac{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\}\hat{F}\right\rangle^{NE}}{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\}\right\rangle^{NE}} + (\hat{A}(z) - a(t))\frac{\partial}{\partial a(t)}\mathrm{Tr}\left[\overline{\rho}_t\hat{F}\right]$$

$$+ \int dE(\hat{\Psi}_{NE}(z) - P(E))\frac{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\}\hat{F}\right\rangle^{NE}}{\left\langle \exp\left\{-\lambda\cdot\hat{A}\right\}\right\rangle^{NE}} \qquad (7.2.16)$$

The first and last terms cancel each other and we have the explicit form of the adjoint projector for this flavour

$$\mathcal{P}_t^\dagger \hat{F}(z) = \int dE \hat{\Psi}_{NE}(z) \frac{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} \hat{F} \right\rangle^{NE}}{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} \right\rangle^{NE}} + (\hat{A}(z) - a(t)) \frac{\partial}{\partial a(t)} \text{Tr}\left[\overline{\rho}_t \hat{F}\right] \quad (7.2.17)$$

When computing the projected current we will need the average of $i\mathcal{L}\hat{A}(z)$ which is just the reversible drift

$$\mathcal{Q}_t^\dagger i\mathcal{L}\hat{A}(z) = i\mathcal{L}\hat{A}(z) - \frac{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} i\mathcal{L}\hat{A} \right\rangle^{N\hat{H}(z)}}{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} \right\rangle^{N\hat{H}(z)}} - (\hat{A}(z) - a(t)) \frac{\partial}{\partial a(t)} \text{Tr}\left[\overline{\rho}_t i\mathcal{L}\hat{A}\right]$$

$$(7.2.18)$$

Because the projected current is inside the microcanonical average, we may substitute $\hat{H}(z)$ by E in the above expression. In addition, the last bit $\text{Tr}\left[\overline{\rho}_t i\mathcal{L}\hat{A}\right]$ is just the reversible drift (7.2.12). Therefore, we arrive at the final expression for the projected current

$$\mathcal{Q}_t^\dagger i\mathcal{L}\hat{A}(z) = i\mathcal{L}\hat{A}(z) - v(N, E, A) - (\hat{A}(z) - a(t)) \frac{\partial}{\partial a(t)} v(N, E, A) \quad (7.2.19)$$

where the reversible drift is

$$v(N, E, A) = \frac{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} i\mathcal{L}\hat{A} \right\rangle^{NE}}{\left\langle \exp\left\{-\lambda \cdot \hat{A}\right\} \right\rangle^{NE}} \quad (7.2.20)$$

Part II

Applications

8

Overview of applications of the general theory

In the second part of the book, we will use the general theory of coarse-graining that has been presented in the first part to particular levels of descriptions of particular systems. Because the theory is so general, the number of possible applications is very large. In order to classify somehow this large variety of applications, we proceed by looking at the categories of *systems*, *phenomena*, and *levels of description*.

We will consider systems that are solid or fluid. The distinction between what is solid and what is fluid may be intuitively obvious at first sight, but what makes a solid "solid" as compared with a fluid is not obvious in microscopic terms. We will discuss later the issue of rigidity and the concept of broken symmetry, but for the time being we will simply classify systems as those in which the particles have **bonded** interactions and those which have **unbonded** interactions. This is a concept that has a clearer microscopic meaning. A group of particles are subject to bonded interactions when they never go too far away from each other and they remain close (during the observational time!). Solids are typically subject to bonded interactions, while fluids are subject to unbonded interactions because a particle in a fluid can drift arbitrarily far away from a neighbour given sufficient time. The category of systems with bonded interactions is larger than what we think about a solid. For example, complex proteins of virus capsids are made of atoms with bonded interactions.

It should be clear that the notion of "bondedness" is not an intrinsic property of the interaction, but also of the thermodynamic conditions of the system. In fact, a solid system in which energy is being pumped in will experience eventually a melting process in which the solid atoms go into a liquid state in which the interactions (which is due to the same potential as in the solid phase!) are nonbonded. Phase change can also occur in a fluid, in which a liquid may go into the vapour phase.

A large number of applications that we will discuss relate to the transport in heterogeneous systems, ranging from the nanoscale to macroscopic everyday scales. Heterogeneous means that the system is made of subsystems that can be distinguished, usually by the phase of the system (phase in the sense of solid, liquid, gas).

Under this distinction of solid and fluid systems we may encounter the following

combinations, leading to an heterogeneous system.

- Solid-solid. Two solids are interacting with each other. If they are moving relative to each other, we may consider issues of how is the transport of momentum from one solid to the other, which will lead to the issues of tribology or friction between solids. If we are interested on how energy is transported we will

- **[Etc.]**

8.1 Recipe to coarse-grain a system

This is a checklist of the steps that need to be taken in order to obtain the coarse-grained dynamics of a system.

- **How is the system described microscopically?** In particular:

  - **How the microscopic state of the system is described?** In principle, Classical Mechanics is a theory of point particles and, therefore, it governs the motion of "atoms" that have an irreducible structure. However, sometimes objects made of many "atoms" (like complex molecules or colloidal particles) are assumed to be point particles. Of course, this suggests that a coarse graining (i.e. elimination of microscopic degrees of freedom) is already being made by the very selection of the "microscopic" variables to describe microscopically the system. This coarse-graining is, of course, assumed to be "inoffensive", this is, it is possible to assume that the dissipation associated to this coarse-graining is negligible.

  - **What is the Hamiltonian of the system?** Essentially this specifies the interactions between the microscopic degrees of freedom. As described above, if we take as point particle some compounded object, the corresponding potential energy in the Hamiltonian should be understood as an effective potential. Fore example, in some occasions, the effect of the container is simplified with an *external potential* that confine the system. Of course, in doing this we are also doing some implicit coarse-graining.

  - **What are the dynamic invariants of the system?** Usually the Hamiltonian and the number of particles are the usual invariants. Depending on whether the system is in a periodic domain, an infinite domain or in an external potential we may have conservation of total linear momentum, total angular momentum, or none.

- **What are the CG variables?** The same system can be described at different levels of description. This will be essentially dictated by the time scales that we want to explore in our system. This is where the intuition and experience enters and it is the crucial step in the construction of the dynamics. The CG variables should include the dynamic invariants of the system and should exclude redundant information. If the dynamic invariants are expressible in terms of the rest of variables (as in GENERIC) then they should be excluded from the description in order to avoid singularities.

- **What flavor do we choose?** This is determined by the statistical information that we want to retain about the relevant variables selected. Some CG variables, typically those leading to field theories and kinetic theory, only admit the averages flavor where we keep track of the averages, while others do admit a probabilistic flavor for which one keeps track of the distribution function of the CG variables. In general, we favor the probabilistic flavor whenever possible, as it gives much information of the level of description, and the effort is usually not different than obtaining the averages flavor of the same system.

Once you have decided about the above **the recipe to coarse-graining** is as follows

1. Obtain the relevant ensemble.

2. Obtain the entropy of the level of description.

3. Compute the time derivatives of the CG variables.

4. Compute the reversible drift. There are two routes to get that through $v_\mu(a) = \langle i\mathcal{L}\rangle^{\hat{a}}$ or through the formulation of the reversible operator $L_{\mu\nu}(a)$

5. Get the structure of the irreversible operator.

6. Write down the transport equation.

7. Approximate things according to your intuition and past experience about the physics of the system.

8. You are done.

The microscopic level

The CG variables

If external forcing, include the coupling functions of the external force that appear in the Hamiltonian.

The CG energy

The time derivatives of the CG variables

The projected currents

The entropy

The reversible drift term

The reversibility condition

The dissipative matrix

The Onsager's reciprocity

The spurious drift

The noise

The irreversible drift

The final SDE

Discussion

Models

Of course, once the approximate transport equations are obtained one needs to verify that they do describe the observed phenomena in the system. As mentioned in the introduction **??**, this is not the objective of the present book, that merely offers the way to obtain the equations but not offer the ways to solve them. However, the validation of the obtained equations is an obvious scientific concern. One way to validate the obtained transport equation is to run MD simulations of the microscopic system and compare the predictions for the evolution of the CG variables measured in the MD simulation and those predicted by the macroscopic approximate transport equation. This is not always possible, because the time scales of the macroscopic variables are too large, prohibiting the use of the MD technique. In this case, the ultimate validity of the obtained transport equation can only be assessed by comparing with real experiments, under the assumption that the microscopic model we started with is indeed a faithful representation of the system under experiment. In any case, the value of the CG approach is that it provides *the structure* of the equations, with few parameters (or functions) that, while they could be computed in principle with MD, they can also be fitted to get good agreement with experimental results.

Once the approximate transport equations are obtained, one can look at the phenomenology described by these equations. While this is not the primary objective in the present book, there is an issue about time scales that belongs entirely to realm of coarse-graining. As repeatedly stressed, the crucial assumption in the theory of coarse-graining is the separation of time scales. Even though the CG variables are slow, some of them may be slower than others. In this case, we may investigate whether the fast CG variables can be further eliminated from the description. The use of dimensionless forms of the equations with dimensionless numbers in it allows one to identify what terms may

be neglected and how. If one is interested on time scales corresponding to the slowest CG variables, it is possible to use the theory of coarse-graining with only those slowest variables. A usual case is one in which thermal conductivity coefficients are very large in practice, meaning that energy is transported very fast as compared to other processes, leading to an isothermal description. In this case, one can use the total energy instead of local energy density variables.

9

Thermodynamics

9.1 Introduction

The first level of description to be considered in any system is the one described by the dynamical invariants of the system. For isolated systems the time independent Hamiltonian $\hat{H}(z)$ giving the total energy of the system will be the most important invariant to be kept in the description. If the system is translationally and rotationally invariant, then the total linear and angular momentum should also be kept in the description. In MD simulations with periodic boundary conditions total linear momentum is conserved but total angular momentum is not conserved, as the system is not rotationally invariant. In general, any physical container will break these symmetries[1].

The evolution of the CG variables of this level is trivial: they do not evolve at all and, for this reason, this level may receive the name of **Thermo*statics***. The relevant ensemble $\overline{\rho}(z)$ at the level of description of Thermostatics depends only on dynamical invariants and is, therefore, an equilibrium ensemble, i.e. it is a stationary solution of the Liouville's equation. As we have discussed in Sec. 1.8 we expect that any initial distribution tends towards the equilibrium ensemble at long times. For this reason we expect that the Thermostatics level of description is appropriate for the description of system that have been left alone until they settle towards the equilibrium state. This level of description is appropriate at very long time scales when the mixing of the microscopic dynamics has reduced all information to just the one concerning to the dynamic invariants. It is therefore, the level of description of equilibrium states. The basic object in the level of description of Thermostatics is the bare entropy (3.4.39) or the dressed entropy (2.7.13) as a function of the dynamical invariants (and the parameters of the Hamiltonian like the volume of the container). The only questions that can be posed in this level of description are very simple: what is the functional form of the entropy? If we know this, then we may answer questions like how the derivatives of the entropy change with their arguments. These relationships are known as **the equations of state** of the system that tells us how the pressure of the system depends on the volume, for example. In principle, the pressure and the volume are measurable quantities in a system and,

---

[1]A cylindrical container may still respect the conservation of the component of angular momentum along the axis of the cylinder.

therefore, we could verify the equations of state that may be predicted from a model for the entropy function. Note, however, that these measurements will be made through the concept of equilibration between two subsystems, and this leads directly to the issue of obtaining the equilibrium state of several interacting systems.

In the present Chapter we explore the level of **Thermo*dynamics***. This level of description involves the interaction between two or more systems. According to Callen, the fundamental problem of Thermodynamics is the prediction of the values of the extensive variables when some internal constraint in the system is removed. We will consider this problem in the light of the Theory of Coarse-Graining. Here, we will not only answer to the question about *what* is the final distribution of extensive variables in each system, but also to the interesting question of *how* the approach towards this final state occurs, by giving the dynamic equations for the evolution of the extensive variables. Typically, the level of thermodynamics is one in which the CG variables are the energy of the systems interacting and, possibly, other macroscopic variables like the position of a piston or the length of a rubber band. In addition to consider the relaxation towards equilibrium when internal constraints are removed, we will consider situations in which the system of interest is subject to external forcing. These situations require the introduction of the concepts of work and heat. We will not make assumptions on the size of the system and, therefore, for small systems the fluctuations will be important. In recent years there has been a renewed impetus in the study of Thermodynamics in an attempt to consider smaller systems as those encountered in micro and nano applications. This has produced a new field known Thermodynamics of Small Systems [?] or Stochastic Thermodynamics [?], and with new and exact results like the Fluctuation Theorems.

Because the CG variables are the energy of the interacting systems, no spatial inhomogeneities should be present inside the systems for this level of description to be valid. Otherwise, we should use CG variables accounting for these inhomogeneities, typically in the form of CG fields. The use of the energy (and not the energy density field, for example) requires that all macroscopic motions, like those of pistons or other external fields, should be sufficiently slow, in order for the energy to be a sufficiently good CG variable.

In this Chapter, we consider several systems that illustrate how the usual concepts of Thermodynamics can be expressed in Statistical Mechanics terms. The presentation departs manifestly from usual presentation of Thermodynamics where the concept of *time* never enters although the word *process* is used profusely. This may be understandable, because one of the implicit assumptions is that the involved processes are so *slow* that the system is always in "equilibrium states" during the process. However, being slow does not means that things do not move (as an example of a reversible motor should make clear). Usual presentations of Thermodynamics never provide the Ordinary Differential Equations governing the time evolution of the different variables of the systems that are modelled. In addition, the issue of fluctuations is not discussed in usual presentations. We will discuss them in the present Chapter by providing the Stochastic Differential Equations governing the CG variables of Thermodynamics.

We will discuss a number of systems described at the level of Thermodynamics in order to illustrate the general tenet of this Chapter, that is, that Thermodynamics is a particular example of the Theory of Coarse-Graining. The systems that we consider are a piston enclosing a compressible fluid like a gas, and thermalized in contact with a thermal
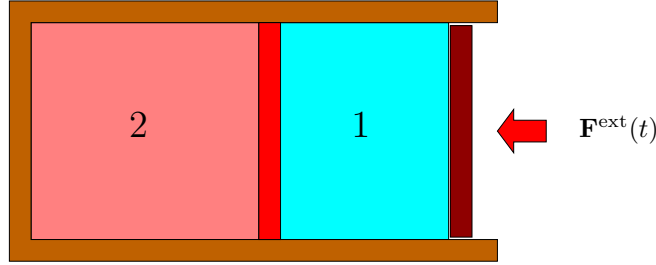
Figure 9.1: A piston encloses a gas which forms system 1. On the piston acts an external force $\mathbf{F}^{\text{ext}}(t)$ which may depend on time. System 2 interacts with system 1 and will act as a thermal bath.

bath. This example, which is at the origin of the foundations of Thermodynamics in his quest to understand stem engines by Carnot and many others, allows one to introduce the concepts of work and heat. The second system that we consider is the adiabatic piston, consisting on two chambers of gas separated by a moving wall. This problem has receved some attention in recent years.

9.2   The forced thermalized piston

Description of the system

The system that we consider in this section is a moving piston enclosing a gas inside a container 1. In contact with the container we have a second system 2 that will eventually act as a thermal bath. The piston may be subject to a time dependent force that will perform work on the system. In turn, the gas in the piston will transfer energy from and to the reservoir in the form of heat. This problem is a paradigmatic example in Thermodynamics. A cartoon of the system is shown in Fig 9.1.

The microscopic level

At the microscopic level we have N_1, N_2 atoms in each system. System A is in a gas state and it is easily compressed by a piston. The piston is made of a solid material and confined to move only in the horizontal direction with no friction with the container. The rigidity of the piston makes its thermal expansion small and its interaction with the rest of the container negligible. An external force of magnitude $F^{\text{ext}}(t)$ acts on the piston in the horizontal direction. The container will be modeled in terms of an external confining potential which is very steep and acts only on the atoms of the molecules, but not on the atoms of the piston. The microstate $z = \{z_1, z_2, X, P\}$ is given by the positions $\mathbf{q}_{i_1}$ and momenta $\mathbf{p}_{i_1}$, $z_1 = \{q_1, p_1\}$ of system A and those z_2 for system B plus the position X, P and momentum of the piston, which is just considered as another "particle". We choose the coordinate system with its origin at the wall between systems A, B.

The total Hamiltonian of the system is given by

$$\hat{H}_t(z) = \hat{H}^{(1)}(z_1, X) + \hat{H}^{(2)}(z_1, z_2) + \frac{P^2}{2M} - F^{\text{ext}}(t)\hat{X} \tag{9.2.1}$$

with

$$\hat{H}^{(1)}(z_1, X) = \sum_{i_1} \frac{\mathbf{p}_{i_1}^2}{2m_{i_1}} + \phi^{(1)}(q_1) + \sum_{i_1} \phi(x_{i_1} - X)$$

$$\hat{H}^{(2)}(z_1, z_2) = \sum_{i_2} \frac{\mathbf{p}_{i_2}^2}{2m_{i_1}} + \phi^{(2)}(q_2) + \sum_{i_1, i_2} \phi^I(\mathbf{q}_{i_1} - \mathbf{q}_{i_2}) \tag{9.2.2}$$

Here $\hat{H}^{(1)}(z_1, X)$ is the Hamiltonian of the unperturbed system A that contains the interaction potential energy $\phi^{(1)}(q_1)$ of the atoms of system A plus the interaction energy with the piston, in the last term. $\hat{H}^{(2)}(z)$ is the Hamiltonian of the thermal bath that includes in the last term the interaction potential energy between A and B.

Hamilton's equations lead to the following equations of motion

$$\dot{\mathbf{q}}_{i_1} = \mathbf{v}_{i_1}$$

$$\dot{\mathbf{p}}_{i_1} = \sum_{j_1} \mathbf{F}_{i_1 j_1} + \sum_{i_2} \mathbf{F}_{i_1 i_2} + \mathbf{F}_{i_1 X}$$

$$\dot{\mathbf{q}}_{i_2} = \mathbf{v}_{i_2}$$

$$\dot{\mathbf{p}}_{i_2} = \sum_{j_2} \mathbf{F}_{i_2 j_2} + \sum_{i_1} \mathbf{F}_{i_2 i_1}$$

$$\dot{X} = \frac{P}{M}$$

$$\dot{P} = -\sum_{i_1} \mathbf{F}_{i_1 X} \cdot \mathbf{e}_x + F^{\text{ext}}(t) \tag{9.2.3}$$

In these expressions $\mathbf{F}_{\mu\nu} = -\mathbf{F}_{\nu\mu}$ is the force that particle ν exerts on particle μ and the subindices indicate what syste the particle belongs to. The force $\mathbf{F}_{i_1 X}$ that the piston exerts on particle i_1 has only component on the horizontal direction with unit vector $\mathbf{e}_x$.

The CG variables

The CG variables that we select in this problem are

$$\hat{A}(z) \rightarrow \{\hat{X}, \hat{P}, \hat{H}^{(1)}(z), \hat{H}^{(2)}(z)\} \tag{9.2.4}$$

that take numerical values

$$a \rightarrow \{X, P, E_1, E_2\} \tag{9.2.5}$$

These variables obey the time-reversal property $\hat{A}_\mu(\epsilon \cdot z) = \varepsilon_\mu \hat{A}_\mu(z)$, where in this case, we have the signatures $\varepsilon = (1, -1, 1, 1)$.

The CG energy (6.3.7) is given by

$$E(a,t) \to E(X, P, E_1, E_2, t) = E_1 + E_2 + \frac{P^2}{2M} - F^{\text{ext}}(t)X \tag{9.2.6}$$

Of course, we could choose some other combination for these variables. For example, we could define the system A as including the piston in which case the CG energy would be

$$E'(X, P, E_1, E_2, t) = E_1' + E_2 - F^{\text{ext}}(t)X \tag{9.2.7}$$

As we know, performing a change of variables in the list of CG variables leaves invariant the FPE and, therefore, the physics is exactly the same in one or the other set of variables.

The derivatives of the CG energy are

$$\frac{\partial E}{\partial a_\mu}(a, t) \to (-F^{\text{ext}}(t), V, 1, 1) \tag{9.2.8}$$

The time derivatives of the CG variables

The time derivatives of the CG variables are the result of applying the Liouville operator on them. The Liouville operator is given by $i\mathcal{L}_t = \dot{z} \cdot \frac{\partial}{\partial z}$, where $\dot{z}$ are given by Hamilton's equations, this is

$$i\mathcal{L}_t = \mathbf{v}_{i_1} \cdot \frac{\partial}{\partial \mathbf{q}_{i_1}} + \mathbf{F}_{i_1} \cdot \frac{\partial}{\partial \mathbf{p}_{i_1}} + \mathbf{v}_{i_2} \cdot \frac{\partial}{\partial \mathbf{q}_{i_2}} + \mathbf{F}_{i_2} \cdot \frac{\partial}{\partial \mathbf{p}_{i_2}} + \hat{V} \frac{\partial}{\partial X} + \hat{F}_X \frac{\partial}{\partial P} \tag{9.2.9}$$

The action of the Liouville operator on each CG variable is

$$i\mathcal{L}_t \hat{X} = \hat{V}$$

$$i\mathcal{L}_t \hat{P} = -\sum_{i_1} \mathbf{F}_{i_1 X} \cdot \mathbf{e}_x + F^{\text{ext}}(t) = -\hat{F}_X + F^{\text{ext}}(t)$$

$$i\mathcal{L}_t \hat{H}_1 = \sum_{i_1 i_2} \mathbf{v}_{i_1} \cdot \mathbf{F}_{i_1 i_2} + \sum_{i_1} V \mathbf{e}_x \cdot \mathbf{F}_{i_1 X} \equiv \hat{Q} + \hat{V} \hat{F}_X$$

$$i\mathcal{L}_t \hat{H}_2 = -\sum_{i_1 i_2} \mathbf{v}_{i_1} \cdot \mathbf{F}_{i_1 i_2} = -\sum_{i_1} \mathbf{v}_{i_1} \cdot \mathbf{F}_{i_1} \equiv \sum_{i_\mu} Q_{i_\mu} \equiv -\hat{Q} \tag{9.2.10}$$

where we have introduced the transferred heat $\hat{Q}$ and the force $\hat{F}_X$ that the piston exerts on the gas. Note that by using these explicit expressions and the form (9.2.1) of the Hamiltonian, it is easy to check that $\mathcal{L}_t \hat{H}_t(z) = 0$, as it should.

The bare entropy

The bare entropy (6.3.5) takes the form

$$S_B(X, P, E_1, E_2) = \ln \int dz_1 dz_2 dX dP$$
$$\times \delta(\hat{X}(z_1) - X)\delta(\hat{P}(z_1) - P)\delta(\hat{H}^{(1)}(z_1, X) - E_1)\delta(\hat{H}^{(2)}(z_1, z_2) - E_2)$$
$$= \ln \int dz_1 dz_2 \delta(\hat{H}^{(1)}(z_1, X) - E_1)\delta(\hat{H}^{(2)}(z_1, z_2) - E_2) \qquad (9.2.11)$$

By multiplying and diving with $\Omega^{(2)}(E_2)$, we can express the entropy (9.2.11) in the form

$$S_B(X, P, E_1, E_2) = S^{(1)}(X, E_1, E_2) + S^{(2)}(E_2) \qquad (9.2.12)$$

where we have introduced the entropies of system 1 and 2 according to

$$S^{(1)}(X, E_1, E_2) = \int dz_1 dz_2 \frac{\delta(\hat{H}^{(2)}(z_1, z_2) - E_2)}{\Omega^{(2)}(E_2)} \delta(\hat{H}^{(1)}(z_1, X) - E_1)$$
$$S^{(2)}(E_2) = \ln \Omega^{(2)}(E_2)$$
$$\Omega^{(2)}(E_2) = \int dz_1 dz_2 \delta(\hat{H}^{(2)}(z_1, z_2) - E_2)$$

$$(9.2.13)$$

We will assume that the dependence of $\hat{H}^{(1)}(z_1, X)$ on X is through an interaction potential energy between the atoms of system 1 and the piston that is very steep. This means that the effect of the potential is just to limit the volume where particles of system 1 may move, which is $V = AX$ where A is the area of the piston, this is

$$S^{(1)}(X, E_1, E_2) = \int_V dz_1 \int dz_2 \frac{\delta(\hat{H}^{(2)}(z_1, z_2) - E_2)}{\Omega^{(2)}(E_2)} \delta(\hat{H}^{(1)}(z_1) - E_1) \qquad (9.2.14)$$

where all the dependence of X arises from the limit of integrations over the positions of atoms of system 1.

Note that the Hamiltonian $\hat{H}^{(2)}(z_1, z_2)$ depends on the degrees of freedom z_1 of system 1 because it contains the interaction energy between the two systems. If we assume that this interaction energy is much smaller than the overall energy of system 2, then we have $\hat{H}^{(2)}(z_1, z_2) \simeq \hat{H}^{(2)}(z_2)$ and arrive at the approximate form

$$S^{(1)}(X, E_1, E_2) \simeq S^{(1)}(X, E_1) = \int dz_1 \delta(\hat{H}^{(1)}(z_1, X) - E_1) \qquad (9.2.15)$$

which is the bare entropy of the system 1 as if it was isolated. The derivatives of the

entropy are

$$\frac{\partial S_B}{\partial a_\mu}(a) \rightarrow \begin{cases} \frac{\partial S_B}{\partial X} &\equiv A\frac{P_1}{T_1} = \frac{\partial S^{(1)}}{\partial X} \\[2mm] \frac{\partial S_B}{\partial P} &= 0 \\[2mm] \frac{\partial S_B}{\partial E_1} &\equiv \frac{1}{T_1} = \frac{\partial S^{(1)}}{\partial E_1} \\[2mm] \frac{\partial S_B}{\partial E_2} &\equiv \frac{1}{T_2} = \frac{\partial S^{(1)}}{\partial E_2} + \frac{\partial S^{(2)}}{\partial E_2} \end{cases} \tag{9.2.16}$$

These equations define the temperatures T_1, T_2 and the pressure P_1. These are the equations of state. Note that the temperature T_2 is not simply given by the derivative of $S^{(2)}(E_2)$ with respect to E_2, as it includes a contribution due to the dependence of the entropy of system 1 with the energy of system 2. In the weak coupling limit, this contribution vanishes.

The reversible drift term

In computing the drift term (6.3.2) we first note that the heat $\hat{Q}$ is a linear function of the velocities. Therefore, when averaging over momenta, due to the fact that the Hamiltonian is quadratic in momenta, we will have $\left\langle \hat{Q} \right\rangle^a = 0$.

The drift term (6.3.2) becomes

$$V_\mu(a) \rightarrow \begin{cases} V_X = \frac{P}{M} \\[2mm] V_P = -F_X + F^{\text{ext}}(t) \\[2mm] V_{E_1} = V F_X \\[2mm] V_{E_2} = 0 \end{cases} \tag{9.2.17}$$

where the mean force that the piston exerts on the gases is defined as

$$F_X(X, E_1, E_2) \equiv \langle \hat{F}_X \rangle^{X,P,E_1,E_2} \tag{9.2.18}$$

Note that this force is independent on momenta. The third component of the drift term has this form because $i\mathcal{L}X$ is proportional to the momentum that is, by itself, a CG variable. Note that the drift term satisfies the first identity (4.4.29) expressing energy conservation.

The reversibility condition

The reversibility condition (4.4.29) provides an explicit expression for the average force (9.2.18). In fact, the second identity (4.4.29) becomes

$$F_X \frac{\partial S}{\partial E_1} + \frac{\partial S}{\partial X} + \frac{\partial F_X}{\partial E} = 0 \qquad (9.2.19)$$

By using the derivatives (9.2.16) and neglecting the last term, which is of order $1/N$ smaller than the other terms we obtain

$$F_X = -A P_1(E_1, V_1) \qquad (9.2.20)$$

this is, the average force that the piston exerts on the gas is given by the negative of the pressure times the area of the piston.

The dissipative matrix

The dissipative matrix (6.3.11) in the present case has, in fact, the same structure (10.4.24 as in the adiabatic piston

$$D_{\mu\nu}(a) \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{PP} & D_{P1} & D_{P2} \\ 0 & D_{1P} & D_{11} & D_{12} \\ 0 & D_{2P} & D_{21} & D_{22} \end{pmatrix} \qquad (9.2.21)$$

where the vanishing elements are due to the fact that the time derivative of the position of the piston is the momentum, which is itself a relevant variable. The different elements are given in terms of Green-Kubo expressions

$$D_{\mu\nu}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{H}_\mu \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L}_t \hat{H}_\nu \right\rangle^a \qquad \mu, \nu = 1, 2$$

$$D_{\mu P}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{H}_\mu \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L} \hat{P} \right\rangle^a \qquad \mu = 1, 2$$

$$D_{PP}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{P} \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L} \hat{P} \right\rangle^a \qquad (9.2.22)$$

The fluctuations appearing in these correlations are $\delta \hat{G}(z) = \hat{G}(z) - \langle \hat{G} \rangle^a$. By using (9.2.10), these projected currents become

$$\delta i \mathcal{L}_t \hat{P} = -\delta \hat{F}_X$$
$$\delta i \mathcal{L}_t \hat{H}_1 = \hat{Q} + V \delta \hat{F}_X$$
$$\delta i \mathcal{L}_t \hat{H}_2 = -\hat{Q} \qquad (9.2.23)$$

where the fluctuation of the force that the piston exerts on the gas with respect to the average force given by (9.2.20) is $\delta \hat{F}_X = \hat{F}_X - \langle \hat{F}_X \rangle^a$. Note that the external force does not enter into the projected current and, therefore, the dissipative matrix will not depend

on the external force explicitly. The whole dependence on the external force is through the change in the microscopic dynamics $\exp\{i\mathcal{L}_t t'\}$, which is obviously affected by the external forcing.

We introduce now the following transport coefficients, through the Green-Kubo correlations of these projected currents as

$$\gamma(a) = \frac{1}{T_1} \int_0^{\Delta t} dt \left\langle \delta \hat{F}_X(t) \delta \hat{F}_X \right\rangle^a$$

$$\zeta'(a) = \int_0^{\Delta t} dt \left\langle \delta \hat{F}_X(t) \hat{Q} \right\rangle^a$$

$$\zeta(a) = \int_0^{\Delta t} dt \left\langle \hat{Q}(t) \delta \hat{F}_X \right\rangle^a$$

$$\kappa(a) = \int_0^{\Delta t} dt \left\langle \hat{Q}(t) \hat{Q} \right\rangle^a \tag{9.2.24}$$

The introduction of the prefactor T_1 is just convenient and accords with a usual definition of the friction coefficient in Brownian motion. In principle the transport coefficients (9.2.24) depend on the state $a = (X, P, E_1, E_2)$, as they are defined in terms of conditional expectations.

The elements of the dissipative matrix are

$$D_{11} = \int_0^{\Delta t} dt \left\langle (\hat{Q}(t) + V\delta\hat{F}_X(t))(\hat{Q} + V\delta\hat{F}_X) \right\rangle^a \qquad = \kappa + V(\zeta + \zeta') + \gamma T_1 V^2$$

$$D_{12} = -\int_0^{\Delta t} dt \left\langle (\hat{Q}(t) + V\delta\hat{F}_X(t))\hat{Q} \right\rangle^a \qquad = -(\kappa + V\zeta')$$

$$D_{21} = -\int_0^{\Delta t} dt \left\langle \hat{Q}(t)(\hat{Q} + V\delta\hat{F}_X) \right\rangle^a \qquad = -(\kappa + V\zeta)$$

$$D_{22} = \int_0^{\Delta t} dt \left\langle \hat{Q}(t)\hat{Q} \right\rangle^a \qquad = \kappa$$

$$D_{1P} = -\int_0^{\Delta t} dt \left\langle (\hat{Q}(t) + V\delta\hat{F}_X(t))\delta\hat{F}_X \right\rangle^a \qquad = -(\zeta + \gamma V T_1)$$

$$D_{2P} = \int_0^{\Delta t} dt \left\langle \hat{Q}(t)\delta\hat{F}_X \right\rangle^a \qquad = \zeta$$

$$D_{P1} = -\int_0^{\Delta t} dt \left\langle \hat{F}_X(t)(\hat{Q} + V\delta\hat{F}_X) \right\rangle^a \qquad = -(\zeta' + \gamma T_1 V)$$

$$D_{P2} = \int_0^{\Delta t} dt \left\langle \hat{F}_X(t)\hat{Q} \right\rangle^a \qquad = \zeta'$$

$$D_{PP} = \int_0^{\Delta t} dt \left\langle \hat{F}_X(t)\delta\hat{F}_X \right\rangle^a \qquad = T_1 \gamma \tag{9.2.25}$$

Note that under the assumption that the transport coefficients do not depend on the state, the dissipative matrix *still* depends on the state of the system through the velocity

$V = P/M$ of the piston.

Onsager's symmetry (5.3.31) states that **[Check external force sign!]**

$$D_{\mu\nu}(a) = \varepsilon_\mu \varepsilon_\nu D_{\nu\mu}(\varepsilon \cdot a) \tag{9.2.26}$$

or more explicitly

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{PP}(a) & D_{P1}(a) & D_{P2}(a) \\ 0 & D_{1P}(a) & D_{11}(a) & D_{12}(a) \\ 0 & D_{2P}(a) & D_{21}(a) & D_{22}(a) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{PP}(\varepsilon \cdot a) & -D_{1P}(\varepsilon \cdot a) & -D_{2P}(\varepsilon \cdot a) \\ 0 & -D_{P1}(\varepsilon \cdot a) & D_{11}(\varepsilon \cdot a) & D_{21}(\varepsilon \cdot a) \\ 0 & -D_{P2}(\varepsilon \cdot a) & D_{12}(\varepsilon \cdot a) & D_{22}(\varepsilon \cdot a) \end{pmatrix} \tag{9.2.27}$$

For example, the identity $D_{1P}(a) = -D_{P1}(\varepsilon \cdot a)$ leads to

$$\begin{aligned} D_{1P}(a) &= -(\zeta + \gamma T_1 V) \\ -D_{P1}(\varepsilon \cdot a) &= (\zeta' + \gamma T_1 V) \\ \zeta + \zeta' &= 0 \end{aligned} \tag{9.2.28}$$

Note that, from the definition through the Green-Kubo formula (9.2.24), the coefficient ζ involves correlations of the energy exchange between systems 1 and 2 and the fluctuation of the forces on the piston. If the system 1 is sufficiently large, these correlations are expected to be negligible. Therefore, we will neglect $\zeta \simeq 0$. The final form of the dissipative matrix is

$$D_{\mu\nu}(a) \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \gamma T_1 & -\gamma T_1 V & 0 \\ 0 & -\gamma T_1 V & \kappa + \gamma T_1 V^2 & -\kappa \\ 0 & 0 & -\kappa & \kappa \end{pmatrix} \tag{9.2.29}$$

Note that this matrix is symmetric, positive semi-definite (given that $\kappa, \gamma > 0$), where the semi comes from the degeneracy $D_{\mu\nu}\frac{\partial E}{\partial a_\nu} = 0$. This degeneracy is easily seen to be satisfyed when we multiply the dissipative matrix (10.4.32) with the gradient of the energy (9.2.8).

Scaling of transport coefficients with the area of the piston

The transport coefficients $\kappa(a), \gamma(a)$ are given in terms the Green-Kubo expressions (9.2.24) and involve the time integral of the correlations of the microscopic heat flux $\hat{Q}(t)$ and fluctuation of the force $\hat{F}_X - \left\langle \hat{F}_X \right\rangle^a$ that the piston exerts on the gas. Recall

that this microscopic quantities are defined as

$$\hat{Q} = \sum_{i_1} \mathbf{v}_{i_\mu} \cdot \left(\sum_{i_2} \hat{\mathbf{F}}_{i_1 i_2} \right) = \sum_{i_1} Q_{i_1}$$

$$\hat{F}_X = \sum_{i_1} \hat{\mathbf{F}}_{i_1 X} \tag{9.2.30}$$

where $\hat{\mathbf{F}}_{i_i i_2}$ is the force that particle i_2 of system 2 exerts on particle i_1 of system 1, and $\hat{\mathbf{F}}_{i_1 X}$ is the force that the piston exerts on particle i_1 of system 1. The transport coefficients are, therefore,

$$\kappa(a) = \sum_{i_1 j_1} \int_0^{\Delta t} dt \left\langle \hat{Q}_{i_1}(t) \hat{Q}_{j_1} \right\rangle^a$$

$$\gamma(a) = \frac{1}{T_1} \sum_{i_1 j_1} \int_0^{\Delta t} dt \left\langle \delta\hat{\mathbf{F}}_{i_1 X}(t) \delta\hat{\mathbf{F}}_{j_1 X} \right\rangle^a \tag{9.2.31}$$

We expect that the power $\hat{Q}_{i_1}$ on the particle i_1 will be correlated with that $\hat{Q}_{j_1}$ of particle j_1 only if the two particles are within a microscopic length of the order of the range σ of the microscopic potential generating the forces between particles. Also, the power $\hat{Q}_{i_1} = \sum_{i_2} \hat{\mathbf{F}}_{i_1 i_2}$ vanishes if particle i_1 is a distance larger than the range of interaction between particles i_1, i_2. The number of terms in the double sum in the thermal transfer coefficient κ is expected to be of the order of the number of particles that are within a layer of volume σA near the piston surface. If $n_1 = N_1/V_1$ is the number of particles per unit volume in the gas, we expect that the overall scaling of the thermal transfer coefficient $\kappa(a)$ will be $n_1 \sigma A$. We see that the thermal transfer coefficient scales with the area of contact between system 1 and 2.

On the other hand, the friction coefficient $\gamma(a)$ involves the correlation of the forces $\hat{\mathbf{F}}_{i_1 X}, \hat{\mathbf{F}}_{j_1 X}$ that the piston surface exerts on particles i_1, j_1. We expect that the auto-correlation of $\hat{\mathbf{F}}_{i_1 X}$ will be much larger than any other correlation of $\hat{\mathbf{F}}_{i_1 X}$ with $\hat{\mathbf{F}}_{j_1 X}$ for a different particle. Therefore, only N_1 terms in the sum will contribute. However, not all particles i_1 give a contribution, only those that are in a thin layer of volume σA near the piston[2]. Again, the number of particles in this layer is $n_1 \sigma A$ and we expect that the friction coefficient will scale with $n_1 \sigma A$.

In summary, we expect that both transport coefficients are proportional to the area of the piston.

The spurious drift

The transport coefficients $\kappa(a), \gamma(a)$ depend on the state $a = X, P, E_1, E_2$ as they are given in terms of conditional expectations. Also, from simple arguments like the one leading to the scaling with the area, we expect that these transport coefficients are proportional to $n_1 = N_1/AX$ and therefore, they should depend inversely proportional

---

[2]For the sake of the argument we assume the range of the particle-piston interaction to be the same as for particle-particle interactions.

to the volume of the piston. This makes sense because the transfer of energy, for example, will be more efficient if there are more particles in the interaction layer of volume σA. It is clear, however, that a detailed quantification of these effects is not trivial.

One possibility at hand is to approximate the dissipative matrix with its equilibrium value, that is, by averaging $D(a)$ with $P^{\text{eq}}(a)$. We will not do this for the full matrix $D(a)$, as this would wash up the important dependence on the velocity. However, we will approximate the conditional expectations in the transport coefficients (9.2.24) by ordinary equilibrium averages, where the equilibrium state is the one that would be reached if the external force was constant.

Because the dissipative matrix depends on the state only through the velocity of the piston, the spurious drift that enters the SDE (6.3.15) is given in the present problem

$$\frac{\partial D^S_{\mu\nu}}{\partial a_\nu} \rightarrow \begin{pmatrix} 0 \\[6pt] -\frac{\gamma V}{C_1} \\[6pt] \frac{\gamma V^2}{C_1} - \frac{\gamma}{M} T_1 \\[6pt] 0 \end{pmatrix} \tag{9.2.32}$$

where we have introduced the dimensionless heat capacity C_1 of system 1 through

$$\frac{1}{C_1} = \frac{\partial T_1}{\partial E_1} \tag{9.2.33}$$

The irreversible drift

The irreversible part of the dynamics is given by

$$D_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\mu} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{PP} & D_{P1} & D_{P2} \\ 0 & D_{1P} & D_{11} & D_{12} \\ 0 & D_{2P} & D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} A\frac{P_1}{T_1} \\[4pt] 0 \\[4pt] \frac{1}{T_1} \\[4pt] \frac{1}{T_2} \end{pmatrix} = \begin{pmatrix} 0 \\[4pt] -\gamma V \\[4pt] \kappa\left(\frac{1}{T_1} - \frac{1}{T_2}\right) + \gamma V^2 \\[4pt] -\kappa\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \end{pmatrix} \tag{9.2.34}$$

Note that the irreversible part of the dynamics conserves energy because

$$\frac{\partial E}{\partial a_\mu}(a,t)D_{\mu\nu}(a,t)\frac{\partial S_B}{\partial a_\nu}(a) = 0 \tag{9.2.35}$$

The noise

According the the modelling of the projected currents (9.2.23) with independent incrre-ments of the Wiener process, we define the following noise terms

$$
d\tilde{a}_\mu \rightarrow
\begin{cases}
d\tilde{X} = & 0 \\
d\tilde{P} = & (2T_1\gamma)^{1/2} \, d\tilde{F}_t \\
d\tilde{E}_1 = & -d\tilde{E}_2 - d\tilde{P} \\
d\tilde{E}_2 = & (2\kappa)^{1/2} \, d\tilde{Q}_t
\end{cases}
\tag{9.2.36}
$$

Here, $d\tilde{F}_t, d\tilde{Q}_t$ are two independent increments of the Wiener process satisfying the mnemotechnical Ito rules

$$
d\tilde{F}_t d\tilde{F}_t = dt
$$
$$
d\tilde{Q}_t d\tilde{F}_t = 0
$$
$$
d\tilde{Q}_t d\tilde{Q}_t = dt
\tag{9.2.37}
$$

The prefactors in (9.2.36) are guessed from the structure of the Green-Kubo formulaes. The independence of the two Wiener processes complies with the assumption that $\zeta = 0$. It is straightforward to show that the diadic $d\tilde{a}_\mu d\tilde{a}_\nu = 2D_{\mu\nu}(a)dt$, where the dissipative is given in (10.4.32).

The final SDE

We may now collect the reversible part (9.2.17), the irreversible part (9.2.34), the spurious drift (9.2.32), and the noise terms (9.2.36) to get the final form of the SDE for the thermalized forced piston

$$
\frac{dX}{dt} = V
$$
$$
\frac{dP}{dt} = AP_1 + F^{\text{ext}}(t) - \gamma V - \frac{\gamma V}{C_1} + (2T_1\gamma)^{1/2} \frac{d\tilde{F}_t}{dt}
$$
$$
\frac{dE_1}{dt} = -VAP_1 + \kappa\left(\frac{1}{T_1} - \frac{1}{T_2}\right) + \gamma V^2 + \frac{\gamma V^2}{C_1} - \frac{\gamma}{M}T_1 - (2\kappa)^{1/2}\frac{d\tilde{Q}_t}{dt} - (2T_1\gamma)^{1/2}\frac{d\tilde{F}_t}{dt}
$$
$$
\frac{dE_2}{dt} = -\kappa\left(\frac{1}{T_1} - \frac{1}{T_2}\right) + (2\kappa)^{1/2}\frac{d\tilde{Q}_t}{dt}
\tag{9.2.38}
$$

The equations (9.2.38) must be supplemented with the equations of state introduced in (9.2.16) which are $P_1(AX, E_1, E_2), T_1(AX, E_1, E_2), T_2(E_2)$, where AX is the volume of system 1, and A is the area of the piston. Again, these equations of state need to be computed or modelled.

Discussion

Scope

The SDE (9.2.38) describes the dynamics of the forced thermalized piston at a CG level. In the derivation, we have made no assumptions about the size of the systems. Therefore the equations should be valid for large macroscopic system for which fluctuations will be negligible, as well as for small systems of nanoscopic size.

Note that we have not assumed that the coupling of the piston with the thermal bath is weak. For a macroscopic piston, the "weakness" of the coupling can be argued from the observation that the microscopic interaction potential energy between the two systems should scale as the surface of contact between the two systems. This is because the interaction is short range and involve a few number of pairs of atoms of different systems, those that are close to the surface. For a macroscopic piston this number is exceedingly smaller than the number of pairs of atoms in the bulk, that make most of the internal energy of the two systems. However, for a microscopic or nanoscopic piston, the above argument no longer holds, and we cannot neglect the interaction term. This means that the entropy $S^{(1)}$ of system 1 depends on the energy of system 2, which is quite strange from a macroscopic point of view but rather natural when we reach the nanoscopic realm.

The SDE equations are valid provided that the CG variables X, P, E_1, E_2 are "sufficiently good". This means that, at the time scales of interest, there are no other CG variables that evolve on comparable time scales[3]. If this was the case, the memory kernels would capture the evolution of these slow variables (slow in the time scale of X, P, E_1, E_2) and could not be treated as a Dirac delta in time. The dynamics would not be Markovian.

For this level of description to be Markovian, it is obvious that there should be no spatial inhomogeneities in the system. This implies that the external forcing should be not very fast. If we hit with a hammer the piston (a rather fast process), a shock wave will be developed and the energy E_1 of system 1 is not sufficient to characterize these states of affairs. We expect that the correlation function entering the Green-Kubo expressions will signal the presence of this bouncing back and forth of the shock wave accross the piston. For example, the force-force correlation entering the definition of the friction coefficient will have correlation bumps every time the shock waves returns to the piston. This is quite a long memory effect that invalidates the Markovian assumption.

Also, we expect that a description with just the total energies can be adequate only if the system is relatively close to equilibrium. Imagine for a moment that system 2 as an extremely large temperature and system 2 a very small temperature. We expect that if we put in contact these two systems strong inhomogeneities will develop in the system that would spoil again this description.

In summary, for this level of description to be appropriate, the external forcing should be slowly varying in the time scale in which any possible inhomogeneity in the gas develops and the system should be relatively close to equilibrium. All these imprecise statements can only be made quantitative by studying concrete realizations of the model through numerical and simulation analysis. Some times this is referred to as the system

---

[3]In words of Green, the CG variables forms a *complete set of gross variables*[**?**].

tracing a "reversible path" that travels always through equilibrium states. We find this phrasing very misleading because a system described with the Thermo*dynamic* level of description is not reversible (as attested by the presence of transport coefficients) and it is not at equilibrium (as it clearly evolves). Some times these paths are referred to as "adiabatic". This seems more appropriate, but may be confused with "no heat", which is not correct. We prefer to refer to this issue by saying that the Thermodynamics level of description is valid for **slow external forcing near global equilibrium**.

The macroscopic piston

In order clarify the discussion, let us consider the SDE (9.2.38) in the limit when fluctuations can be neglected. This is, the two systems are very large and the piston massive. In the FPE corresponding to (9.2.38) this means to neglect the second derivative terms transforming the FPE into a deterministic Liouville equation. The SDE (9.2.38) becomes in the macroscopic limit a deterministic dynamics, obtained from (9.2.38) by neglecting the noise and the spurious drift terms. This results into the ODE

$$\frac{dX}{dt} = V$$

$$\frac{dP}{dt} = AP_1 + F^{\text{ext}}(t) - \gamma V$$

$$\frac{dE_1}{dt} = -VAP_1 + \kappa \left(\frac{1}{T_1} - \frac{1}{T_2} \right) + \gamma V^2$$

$$\frac{dE_2}{dt} = -\kappa \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \tag{9.2.39}$$

Let us discuss the physical meaning of all the terms in these equations. Focus first on the motion of the piston. In the equation for the momentum, we identify the force AP_1 that the gas exerts on the piston, which is just the pressure of the gas times the area of the piston. The second term is just the external force that acts on the piston. Finally, the third tie is a friction force proportional to the velocity. This means that the interaction of the gas induces a friction on the piston that opposes its motion and, in the case of a constant external force, would eventually stop its motion.

We can look at the equation of the energy E_1 of system 1. The first term is the reversible work that the piston does on the gas. This term is reversible because it will not contribute to the increase of the total entropy of the system (see below). The second term is a heat conduction term that is positive if $T_2 > T_1$, that is, there will be a transfer of energy from system 2 to system 1 if system 2 is hotter. For this reason, we may refer to κ as the thermal transfer coefficient. The last term proportional to the square of the velocity of the piston is a viscous heating term. The physical origin of this term is that the kinetic energy of the piston that is lost while the friction forces try to stop the motion of the piston has to go into the internal energy of system 1. This is captured by the viscous heating term. Finally, the equation for the energy E_2 of system 2 simply says that the energy of system 1 gained by heat conduction is lost by system 2.

Recall that the transport coefficients κ, γ scale with the area of the piston. From a macroscopic point of view, this makes quite a lot of sense: the transfer of energy between

systems will be larger the larger is the area of contact, the rest of conditions kept equal. And similarly with the transfer of momentum.

Finally, imagine that the container of system 2 is separated a distance larger than the range of interaction of the forces $\hat{\mathbf{F}}_{i_1 i_2}$ between the atoms of system 1 and system 2. In this case, the microscopic heat $\hat{Q}$ defined in (9.2.10) vanishes. This implies that the thermal transfer coefficient κ defined in (9.2.24) vanishes. In this case, the CG dynamics (9.2.39) of the two systems become decoupled.

Heat, work, Clausius

It is interesting to recover the First and Second Laws of Thermodynamics from the ODEs (9.2.39). As we know, these laws are already embodied in the structure of the equations. However, a definition of work and heat, two crucial concepts in thermodynamics is still missing. And we also miss the Clausius identity that relates "entropy" with "heat". Let us see how these results arise from the set of ODEs (9.2.39).

Let us start with the total energy of the system (9.2.6). If we compute the time derivative of this energy by using the ODEs (9.2.39), we obtain

$$\frac{d}{dt} E(a, t) = X \frac{dF^{\text{ext}}}{dt}(t) \tag{9.2.40}$$

When the external force is time-independent, then energy is conserved. If we compute the time derivative of the total entropy (9.2.12) by using the derivatives (9.2.16) we obtain

$$\frac{dS_B}{dt} = \kappa \left(\frac{1}{T_1} - \frac{1}{T_2} \right)^2 + \gamma \frac{V^2}{T_1} \geq 0 \tag{9.2.41}$$

The first term in the entropy production is due to the irreversible transport of energy between systems 1 and 2. The second one is due to the irreversible friction on the momentum of the piston. This is the Second Law for the total entropy of the "universe".

In order to find the First Law in the usual form as it appears in Thermodynamics, consider the system formed by system 1 plus the piston. We denote this system as system 1'. The energy of this composite system is given by

$$E_{1'} = E_1 + \frac{P^2}{2M} \tag{9.2.42}$$

If we compute the time derivative of this energy and use the dynamics (9.2.39) we obtain

$$\frac{d}{dt} E_{1'} = \dot{W} + \dot{Q} \tag{9.2.43}$$

where the work per unit time $\dot{W}$ done by the external force and the heat per unit time $\dot{Q}$ given by the system 2 are defined as

$$\dot{W} = V F^{\text{ext}}(t)$$
$$\dot{Q} = \kappa \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \tag{9.2.44}$$

Eq (9.2.43) is a neat example of the First Law expressing conservation of energy.

Now, consider the evolution of the entropy of system 1. By using the chain rule, equations of state (9.2.16), and the dynamic equations (9.2.39)

$$\frac{dS^{(1)}}{dt} = \frac{\partial S^{(1)}}{\partial E_1}\frac{dE_1}{dt} + \frac{\partial S^{(1)}}{\partial E_2}\frac{dE_2}{dt} + \frac{\partial S^{(1)}}{\partial X}\frac{dX}{dt} \quad = \frac{\dot{Q}}{T_1} - \frac{\partial S^{(1)}}{\partial E_2}\dot{Q} + \frac{\gamma V^2}{T_1} \quad (9.2.45)$$

If the systems are very large, we may neglect the dependence of the entropy of system 1 on the energy E_2 of the reservoir, leading to

$$\frac{dS^{(1)}}{dt} = \frac{\dot{Q} + \gamma V^2}{T_1} \quad (9.2.46)$$

This is almost Clausius' connection between the entropy of the system with the heat, although in the case of a moving piston, there is some "heat" that is dissipated through the friction force on the piston. In the case of a very massive piston for which $V \sim 0$ we recover Clausius equality

$$\frac{dS^{(1)}}{dt} = \frac{\dot{Q}}{T_1} \quad (9.2.47)$$

In some textbooks the Clausius identity is restricted by saying "valid for reversible isothermal processes". In view of the above discussion, such an expression makes not much sense because any process described with the ODEs (9.2.39) describes an *irreversible non-isothermal* process and, yet, the Clausius equality is valid for this system. Obviously, what should be understood by such qualifications is that the Clausius equality is valid whenever the CG level of Thermodynamics (based on total energies) is valid. As we have discussed, this is possible when the external forcing is *slow* and the system is sufficiently *close to equilibrium*.

The equilibrium solution

If the external force does not depend on time the above set of equations have a unique equilibrium state, characterized by the values $E_1^{\text{eq}}, E_2^{\text{eq}}, X^{\text{eq}}, P^{\text{eq}}$. In this case, the total energy (9.2.6) is conserved, while total entropy always increases until it reaches the maximum when the time derivative (9.2.41) takes the zero value. This happens when the temperatures are equal and the velocity of the piston vanishes. In the equilibrium state we have

$$T_1(X^{\text{eq}}, E_1^{\text{eq}}, E_1^{\text{eq}}) = T_2(X^{\text{eq}}, E_1^{\text{eq}}, E_1^{\text{eq}})$$
$$V^{\text{eq}} = \frac{P^{\text{eq}}}{M} = 0 \quad (9.2.48)$$

Together with energy conservation and the momentum equation

$$E_1^{\text{eq}} + E_2^{\text{eq}} - F^{\text{ext}}X^{\text{eq}} = E_0$$
$$AP_1(X^{\text{eq}}, E_1^{\text{eq}}, E_1^{\text{eq}}) - F^{\text{ext}} = 0 \quad (9.2.49)$$

we have a system of three equations for the three unknowns $E_1^{\text{eq}}, E_2^{\text{eq}}, X^{\text{eq}}$. At equilibrium the temperatures are equal, the pressure force balances the external force and the piston does not move.

When the external force depends on time, there is obviously no stationary solution.

The equilibrium distribution

The entropy of this microcanonical description is (6.3.13). When evaluated at the solution $P(E_1, E_2, t)$ of the FPE, this entropy increases until it reaches its maximum at the equilibrium probability distribution of the energies in each subsystem $P^{\text{eq}}(E_1, E_2)$. Therefore, the FPE describes the tendency towards equilibrium.

As discussed in Sec 3.8, the equilibrium probability $P^{\text{eq}}(E_1, E_2)$ of the system is given by Einstein formula (3.8.8), this is

$$P^{\text{eq}}(E_1, E_2) = \phi(E_1 + E_2) \exp\{S_B(E_1, E_2)\} \qquad (9.2.50)$$

where $\phi(E)$ is a function that is fixed by the initial distribution of the energy of the total system. The most probable state E_1^*, E_2^* is the one that maximizes the probability $P^{\text{eq}}(E_1, E_2)$. By setting the derivatives to zero we obtain from (**??**)

$$\frac{\partial S}{\partial E_1}(E_1, E_2) = \frac{1}{T_1(E_1^*)} - \phi'(E) = 0$$

$$\frac{\partial S}{\partial E_2}(E_1, E_2) = \frac{1}{T_2(E_2^*)} - \phi'(E) = 0 \qquad (9.2.51)$$

Therefore, the most probable values E_1^*, E_2^* are the solution of the equations

$$T_1(E_1^*) = T_2(E_2^*) \qquad (9.2.52)$$

subject to $E_1^* + E_2^* = E$. The most probable state is, therefore, the one that equates the temperatures of both systems.

A simple motor

The equations for the thermalized piston are new. Recently SDE for this system have been formulated in Ref. [**?**] in the linearized regime. We could linearize the equations but then all the thermodynamic structure is lost. On the other hand, linear equations lead to exponential and oscillatory decay and they may provide insight into the consideration of time scales.
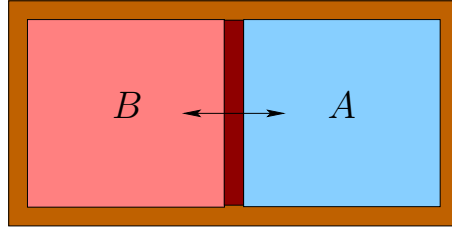
Figure 9.2: Two containers filled with a gas are separated by a moving frictionless piston.

9.3 The adiabatic piston

Description of the system

In the present section, we consider a different piston system, shown in Fig. (9.2). In this case, two gases with N_1 and N_2 particles each are separated by a piston of mass M. The whole system is isolated and the walls of the containers are assumed to be made of hard interactions in such a way that we can model them as external potentials (this is, we disregard the atomic nature of the container walls). The piston does not suffer from any friction with the container walls. Initially, the piston has its position fixed with a mechanical device (like a brake) an each gas is an isolated system with a given energy E_1^0, E_2^0. At $t = 0$ the brake releases the piston and the system evolves towards a new equilibrium position. This system is known in the literature as **the adiabatic piston**, and has received some attention recently. In this section, we seek to obtain the dynamic equations for the evolution of the adiabatic piston within the Theory of Coarse-Graining.

The microscopic level

This idealized system is described by the following Hamiltonian

$$\hat{H}(z_1, z_2, X, P) = \hat{H}_1(z_1, X) + \hat{H}_2(z_2, X) + \frac{P^2}{2M} \qquad (9.3.1)$$

Here, z_1, z_2 denote the degrees of freedom of systems $1, 2$, respectively and X, P is the position and momentum of the piston. We will denote the microstate of the system with $z = z_1, z_2, X, P$. The piston, as well as the rest of container walls, is assumed to have no internal degrees of freedom (or in other words, their internal degrees of freedom decay much faster towards its equilibrium state determined by their overall center of mass position and momenta). $\hat{H}_1(z_1, X)$ is the Hamiltonian of system 1, $\hat{H}_2(z_2, X)$ is the Hamiltonian of system 2. We assume that each of these Hamiltonians contain a potential term that accounts for the interaction with the container walls and forbids the escape of any particle from each container. Because one of this walls is the piston, we

assume that inside $\hat{H}_1(z_1, X), \hat{H}_2(z_2, X)$ we have a potential term of the form

$$\phi_1(z) = \sum_{i_1}^{N_1} \phi_1(X - x_{i_1})$$

$$\phi_2(z) = \sum_{i_2}^{N_2} \phi_2(X - x_{i_2}) \tag{9.3.2}$$

where $\phi(X - x)$ is the potential energy of interaction between a particle at x with the piston. This potential term is also very steep and forbids the direct interaction between particles of system 1 and 2. All the interaction between the gases is through the piston subsystem. The last term in the Hamiltonian (9.3.1) is the kinetic energy of the piston.

The CG variables

We choose as CG variables the total energy $\hat{H}(z)$ in (9.3.1) of the system, the energies $\hat{H}_1(z_1, X), \hat{H}_2(z_2, X)$ of each sub system (including the interaction energy with the piston) and the position $\hat{X}$ and momentum $\hat{P}$ of the piston. The CG variables take numerical values $a = \{X, P, E_1, E_2\}$. Note that the total energy can be entirely describable in terms of the CG variables chosen. Therefore, total energy will not be included in the list of CG variables. The CG energy function (6.3.7) is

$$E(a) = E_1 + E_2 + \frac{P^2}{2M} \tag{9.3.3}$$

The time derivatives of the CG variables

Now we compute the time derivatives of the CG variables

$$i\mathcal{L}\hat{A}_\mu(z) \rightarrow \begin{cases} i\mathcal{L}\hat{X} = \frac{P}{M} \\[2mm] i\mathcal{L}\hat{P} = -\frac{\partial \hat{H}}{\partial X} = \hat{F}^1(z_1, X) + \hat{F}^2(z_2, X) = F^0(z, X) \\[2mm] i\mathcal{L}\hat{H}_1(z_1, X) = -V\hat{F}^1(z_1, X) \\[2mm] i\mathcal{L}\hat{H}_2(z_2, X) = -V\hat{F}^2(z_2, X) \end{cases}$$

$$\tag{9.3.4}$$

here $\hat{F}^1(z_1, X)$ and $\hat{F}^2(z_2, X)$ are the total microscopic forces that the fluid 1 and 2 exert on the piston. $\hat{F}^0(z, X)$ is the total net force felt by the piston due to both fluids.

In what follows we will particularize the SDE (6.3.15) to this particular selection of CG variables. We focus on the SDE rather than on the FPE because the latter is rather cumbersome to write down with all their partial derivatives. However, both are mathematically equivalent. We will consider separately each term that appears in the SDE (6.3.15), summarized in Sec. 3.11. These terms are the equilibrium probability

$P^{\text{eq}}(a)$, the drift terms $V_\mu(a)$, for which we need to compute the time derivatives $i\mathcal{L}\hat{A}_\mu$ of the CG variables. Finally, we will consider the dissipative matrix $D_{\mu\nu}(a)$.

The entropy

The entropy (6.3.5) is given by

$$S(X, P, E_1, E_2) = \ln \int dz \rho_0 \Psi_a(z) \tag{9.3.5}$$

where we use the compact notation

$$\Psi_a(z) = \delta(E_1 - \hat{H}_1(z_1, X))\delta(E_2 - \hat{H}_2(z_2, X))\delta(\hat{X} - X)\delta(\hat{P} - P) \tag{9.3.6}$$

Here the constant ρ^0 accounting for the correct Boltzmann counting is, instead of (1.8.19)

$$\rho^0 = \frac{1}{N_1! N_2! h^{3(N+1)}} \tag{9.3.7}$$

where we assume that the particles of each system are distinguishable from each other.

We will compute approximately the entropy by assuming that the potential of interaction with the piston is very steep and, therefore, its effect is just to limit the regions V_1, V_2 on which the particles of type $1, 2$, respectively, may move, this is

$$\begin{aligned}
S(X, P, E_1, E_2) &= \ln \int_{V_1} dz_1 \int_{V_2} dz_2 \rho^0 \delta(E_1 - \hat{H}_1(z_1, X))\delta(E_2 - \hat{H}_2(z_2, X)) \\
&= S_1(N_1, V_1, E_1) + S_2(N_2, V_2, E_2)
\end{aligned} \tag{9.3.8}$$

where we have introduced the entropy functions of each subsystem

$$\begin{aligned}
S_1(N_1, V_1, E_1) &= \ln \Omega(N_1, V_1, E_1) \\
S_2(N_2, V_2, E_2) &= \ln \Omega(N_2, V_2, E_2)
\end{aligned} \tag{9.3.9}$$

and the volumes of each container are given by

$$\begin{aligned}
V_1 &= XA \\
V_2 &= V - XA
\end{aligned} \tag{9.3.10}$$

with V being the total volume of the system and A the area of the section of the container. Observe that the entropy does not depend on the momentum of the piston.

The derivatives of the entropy are given by

$$\frac{\partial S_B}{\partial a_\mu}(a) \rightarrow \begin{cases} \frac{\partial S}{\partial X} &= A\left[\frac{P_1(N_1,V_1,E_1)}{T_1(N_1,V_1,E_1)} - \frac{P_2(N_2,V_2,E_2)}{T_2(N_2,V_2,E_2)}\right] \\[2mm] \frac{\partial S_B}{\partial P} &= 0 \\[2mm] \frac{\partial S_B}{\partial E_1} &= \frac{1}{T_1(N_1,V_1,E_1)} \\[2mm] \frac{\partial S_B}{\partial E_2} &= \frac{1}{T_2(N_2,V_2,E_2)} \end{cases} \tag{9.3.11}$$

Here, we have introduced the usual definitions of temperature $T(N,V,E)$ and pressure $P(N,V,E)$ for each subsystem

$$\frac{\partial S_B}{\partial E} = \frac{1}{T(N,V,E)}$$

$$\frac{\partial S_B}{\partial V} = \frac{P(N,V,E)}{T(N,V,E)} \tag{9.3.12}$$

The drift term

The drift term (6.3.2) now has the following form

$$V_\mu(a) \rightarrow \begin{cases} V_X(X,P,E_1,E_2) = \frac{P}{M} \\[2mm] V_P(X,P,E_1,E_2) = \langle i\mathcal{L}P\rangle^{X,P,E_1,E_2} = F_1 + F_2 \\[2mm] V_{E_1}(X,P,E_1,E_2) = -VF_1 \\[2mm] V_{E_2}(X,P,E_1,E_2) = -VF_2 \end{cases} \tag{9.3.13}$$

Here, the mean force that the gas 1 exerts on the piston is defined as

$$F_1(X,E_1,E_2) \equiv \langle \hat{F}_1\rangle^a \tag{9.3.14}$$

and similarly for F_2. Note that the average forces F_1, F_2 that the gas 1,2 exert on the piston are, in fact, independent of the momentum of the piston. The third component of the drift term has this form because $i\mathcal{L}X$ is proportional to the momentum that is, by itself, a CG variable. We still need to find the particular functional form of the forces $F_1(X,P,E_1,E_2), F_2(X,P,E_1,E_2)$. We will use the exact identity (6.3.9) that, by using (9.3.11) and (9.3.13) in the present situation reduces to

$$\frac{F_1}{T_1} + \frac{F_2}{T_2} - \frac{\partial S_B}{\partial X} = \frac{\partial F_1}{\partial E_1} + \frac{\partial F_1}{\partial E_2} \tag{9.3.15}$$

Note that the last term scales as $1/N_1, 1/N_2$ with respect to the rest of terms in the equation. If we neglect them, we obtain

$$\frac{F_1}{T_1} + \frac{F_2}{T_2} = \frac{\partial S}{\partial X} \tag{9.3.16}$$

By comparing with the explicit form for the derivative of the entropy with respect to the position of the piston in (9.3.13) we conclude that

$$F_1 = AP_1(N_1, V_1, E_1)$$
$$F_2 = AP_2(N_1, V_1, E_1) \tag{9.3.17}$$

this is, the average force is given by the pressure times the area of the piston.

To this end, we resort to the fact that the present level of description is of the GENERIC type. Therefore, the drift term is of the form of Eq. (3.6.4)

$$v_\mu(a) = \langle i\mathcal{L}\hat{A}_\mu \rangle^a = L_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu}(a) \tag{9.3.18}$$

The reversible matrix (3.6.5) is given by

$$L_{\mu\nu}(a) \equiv \left\langle \frac{\partial \hat{A}_\mu}{\partial z} J \frac{\partial \hat{A}_\nu}{\partial z} \right\rangle^a = \left\langle \{\hat{A}_\mu, \hat{A}_\nu\} \right\rangle^a = \left\langle \frac{\partial \hat{A}_\mu}{\partial q_i}\frac{\partial \hat{A}_\nu}{\partial p_i} - \frac{\partial \hat{A}_\nu}{\partial q_i}\frac{\partial \hat{A}_\mu}{\partial p_i} \right\rangle^a \tag{9.3.19}$$

which, in the present example becomes

$$\begin{pmatrix} 0 & L_{12} & L_{1X} & L_{1P} \\ -L_{12} & 0 & L_{2X} & L_{2P} \\ -L_{1X} & -L_{1P} & 0 & L_{XP} \\ -L_{1P} & -L_{2P} & -L_{XP} & 0 \end{pmatrix} \tag{9.3.20}$$

We see that we only need to compute six terms. The first term is

$$L_{12} = \left\langle \frac{\partial \hat{H}_1}{\partial q_i}\frac{\partial \hat{H}_2}{\partial p_i} - \frac{\partial \hat{H}_2}{\partial q_i}\frac{\partial \hat{H}_1}{\partial p_i} \right\rangle^a = 0 \tag{9.3.21}$$

Because H^1, H^2 depend only on a common coordinate X, the only surviving term would be the ones for X. But because H^1, H^2 do not depend on the conjugate P, the above element vanish. The next term is

$$L_{X2} = \left\langle \frac{\partial X}{\partial q_i}\frac{\partial \hat{H}_2}{\partial p_i} - \frac{\partial \hat{H}_2}{\partial q_i}\frac{\partial X}{\partial p_i} \right\rangle^a = 0 \tag{9.3.22}$$

This term vanishes because for the same reason as the previous element. For identical

reasons $L_{X1} = 0$. Next consider

$$L_{XP} = \left\langle \frac{\partial X}{\partial q_i} \frac{\partial P}{\partial p_i} - \frac{\partial P}{\partial q_i} \frac{\partial X}{\partial p_i} \right\rangle^a = 1 \qquad (9.3.23)$$

The remaining terms are

$$L_{1P} = \left\langle \frac{\partial \hat{H}_1}{\partial q_i} \frac{\partial P}{\partial p_i} - \frac{\partial P}{\partial q_i} \frac{\partial \hat{H}_1}{\partial p_i} \right\rangle^a = \left\langle \frac{\partial \hat{H}_1}{\partial X} \right\rangle^a = -F_1(X, E_1, E_2) \qquad (9.3.24)$$

and with a similar expression for L_{2P}. The final form of the reversible matrix is

$$\begin{pmatrix} 0 & 0 & 0 & -F_1 \\ 0 & 0 & 0 & -F_2 \\ 0 & 0 & 0 & 1 \\ F_1 & F_2 & -1 & 0 \end{pmatrix} \qquad (9.3.25)$$

The drift term will be

$$\begin{pmatrix} 0 & 0 & 0 & -F_1 \\ 0 & 0 & 0 & -F_2 \\ 0 & 0 & 0 & 1 \\ F_1 & F_2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ V \end{pmatrix} = \begin{pmatrix} -VF_1 \\ -VF_1 \\ 0 \\ F_1 + F_2 \end{pmatrix} \qquad (9.3.26)$$

Let us make use of the exact identity (3.6.23) by computing explicitly each of the terms that appear there. On one hand

$$L_{\mu\nu}(a) \frac{\partial S}{\partial a_\mu} = \begin{pmatrix} 0 & 0 & 0 & -F_1 \\ 0 & 0 & 0 & -F_2 \\ 0 & 0 & 0 & 1 \\ F_1 & F_2 & -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{T_1} \\ \frac{1}{T_2} \\ \frac{\partial S}{\partial X} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{F_1}{T_1} + \frac{F_2}{T_2} - \frac{\partial S}{\partial X} \end{pmatrix} \qquad (9.3.27)$$

On the other hand

$$\frac{\partial}{\partial a_\mu} L_{\mu\nu}(a) = \frac{\partial}{\partial E_1} L_{1\nu} + \frac{\partial}{\partial E_2} L_{2\nu} + \frac{\partial}{\partial X} L_{X\nu} + \frac{\partial}{\partial P} L_{P\nu} = \begin{pmatrix} -\frac{\partial F_1}{\partial P} \\ -\frac{\partial F_2}{\partial P} \\ 0 \\ \frac{\partial F_1}{\partial E_1} + \frac{\partial F_1}{\partial E_2} \end{pmatrix} \qquad (9.3.28)$$

The dissipative matrix

The only remaining object to compute is the dissipative matrix (6.3.11). We note that because the time derivative of the position of the piston is the momentum, which is itself a relevant variable, the projected current concerning the position will vanish. Therefore the dissipative matrix will have the form

$$
D_{\mu\nu}(a) \to \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & D_{PP} & D_{P1} & D_{P2} \\
0 & D_{1P} & D_{11} & D_{12} \\
0 & D_{2P} & D_{21} & D_{22}
\end{pmatrix}
\tag{9.3.29}
$$

The elements of this matrix are

$$
D_{PP}(a) = \int_0^{\Delta t} dt \, \langle \delta i \mathcal{L} P \exp\{i\mathcal{L}t\} \delta i \mathcal{L} P \rangle^a
$$

$$
D_{\mu P}(a) = \int_0^{\Delta t} dt \, \left\langle \delta i \mathcal{L} \hat{H}_\mu \exp\{i\mathcal{L}t\} \delta i \mathcal{L} P \right\rangle^a \qquad \mu = 1, 2
$$

$$
D_{\mu\nu}(a) = \int_0^{\Delta t} dt \, \left\langle \delta i \mathcal{L} \hat{H}_\mu \exp\{i\mathcal{L}t\} \delta i \mathcal{L} \hat{H}_\nu \right\rangle^a \qquad \mu, \nu = 1, 2 \tag{9.3.30}
$$

By using the (9.3.4), we may write the different elements of the dissipative matrix as

$$
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & D_{PP} & D_{P1} & D_{P2} \\
0 & D_{1P} & D_{11} & D_{12} \\
0 & D_{2P} & D_{21} & D_{22}
\end{pmatrix}
= \frac{1}{\beta}
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & \gamma_{00} & -V\gamma_{01} & -V\gamma_{02} \\
0 & -V\gamma_{01} & V^2\gamma_{11} & V^2\gamma_{12} \\
0 & -V\gamma_{02} & V^2\gamma_{12} & V^2\gamma_{22}
\end{pmatrix}
\tag{9.3.31}
$$

where we have introduced the friction coefficients as the force-force correlations

$$
\gamma_{11} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^1 \exp\{i\mathcal{L}t\} \delta F^1 \rangle^a
$$

$$
\gamma_{12} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^2 \exp\{i\mathcal{L}t\} \delta F^1 \rangle^a
$$

$$
\gamma_{22} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^2 \exp\{i\mathcal{L}t\} \delta F^2 \rangle^a
$$

$$
\gamma_{01} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^1 \exp\{i\mathcal{L}t\} \delta F^0 \rangle^a = \gamma_{11} + \gamma_{12}
$$

$$
\gamma_{02} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^2 \exp\{i\mathcal{L}t\} \delta F^0 \rangle^a = \gamma_{12} + \gamma_{22}
$$

$$
\gamma_{00} \equiv \beta \int_0^{\Delta t} dt \, \langle \delta F^0 \exp\{i\mathcal{L}t\} \delta F^0 \rangle^a = \gamma_{11} + 2\gamma_{12} + \gamma_{22} \tag{9.3.32}
$$

Note that we have defined the friction coefficients with a prefactor β as it is conventionally done in the theory of Brownian motion. Here β is the inverse of a thermal energy. In

principle these friction coefficients depend on the state $a = (X, P, E_1, E_2)$ due to the conditional expectations. Note that the dissipative matrix $D(a)$ is symmetric, positive definite, and satisfy the GENERIC degeneracy

$$
D_{\mu\nu}(a)\frac{\partial E}{\partial a_\nu} = 0 \quad \longrightarrow \quad \frac{1}{\beta} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \gamma_{00} & -V\gamma_{01} & -V\gamma_{02} \\ 0 & -V\gamma_{01} & V^2\gamma_{11} & V^2\gamma_{12} \\ 0 & -V\gamma_{02} & V^2\gamma_{12} & V^2\gamma_{22} \end{pmatrix} \begin{pmatrix} 0 \\ V \\ 1 \\ 1 \end{pmatrix} = 0 \quad (9.3.33)
$$

Therefore, the irreversible part of the dynamics will automatically conserve total energy. As not all friction coefficients are independent, we choose as the independent ones $\gamma_{12}, \gamma_{01}, \gamma_{02}$. In terms of these independent coefficients, the dissipative matrix becomes

$$
\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{PP} & D_{P1} & D_{P2} \\ 0 & D_{1P} & D_{11} & D_{12} \\ 0 & D_{2P} & D_{21} & D_{22} \end{pmatrix} = \frac{1}{\beta} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \gamma_{00} & -V\gamma_{01} & -V\gamma_{02} \\ 0 & -V\gamma_{01} & V^2(\gamma_{01} - \gamma_{12}) & V^2\gamma_{12} \\ 0 & -V\gamma_{02} & V^2\gamma_{12} & V^2(\gamma_{02} - \gamma_{12}) \end{pmatrix}
$$
$$(9.3.34)$$

Therefore, the irreversible term is

$$
D_{\mu\nu}(a)\frac{\partial S}{\partial a_\nu}(a) \to \frac{1}{\beta}
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & (\gamma_{01}+\gamma_{02}) & -V\gamma_{01} & -V\gamma_{02} \\
0 & -V\gamma_{01} & V^2(\gamma_{01}-\gamma_{12}) & V^2\gamma_{12} \\
0 & -V\gamma_{02} & V^2\gamma_{12} & V^2(\gamma_{02}-\gamma_{12})
\end{pmatrix}
\begin{pmatrix}
\frac{\partial S}{\partial X} \\
-\beta V \\
\frac{1}{T_1} \\
\frac{1}{T_2}
\end{pmatrix}
$$

$$
= \frac{1}{\beta}
\begin{pmatrix}
0 \\
-V\gamma_{01}\frac{1}{T_1} - V\gamma_{02}\frac{1}{T_2} - C'V(\gamma_{01}+\gamma_{02}) \\
V^2(\gamma_{01}-\gamma_{12})\frac{1}{T_1} + V^2\gamma_{12}\frac{1}{T_2} + C'V^2\gamma_{01} \\
V^2\gamma_{12}\frac{1}{T_1} + V^2(\gamma_{02}-\gamma_{12})\frac{1}{T_2} + C'V^2\gamma_{02}
\end{pmatrix}
$$

$$
= \frac{1}{\beta}
\begin{pmatrix}
0 \\
-V\left[\gamma_{01}\frac{1}{T_1} + \gamma_{02}\frac{1}{T_2}\right] \\
V^2\gamma_{12}\left(\frac{1}{T_1}-\frac{1}{T_2}\right) + V^2\gamma_{01}\frac{1}{T_1} \\
V^2\gamma_{12}\left(\frac{1}{T_2}-\frac{1}{T_1}\right) + V^2\gamma_{02}\frac{1}{T_2}
\end{pmatrix}
\tag{9.3.35}
$$

The friction coefficients appearing in these expressions depend in general on the state of the system. We will approximate the above conditional expectations by simple equilibrium averages, in such a way that the dissipative matrix no longer depends on the state of the system. In this way, we will have equilibrium Green-Kubo expressions for these friction coefficients. The term depending on the derivatives of the dissipative matrix in the SDE (3.4.41) will be zero in the present approximation.

The random terms

We still have to specify the random forces $d\tilde{E}_1, d\tilde{E}_2, d\tilde{P}$. We know that these forces need to satisfy the fluctuation-dissipation theorem. The main physical guidance to construct the random forces is the assimilation of the random forces with the projected currents entering into the Green-Kubo expressions, Eq. (3.5.42). In this case, it seems natural to

introduce the following random terms

$$dP̃ = dF̃_1 + dF̃_2$$
$$dẼ_1 = -V dF̃_1$$
$$dẼ_2 = -V dF̃_2 \tag{9.3.36}$$

where $dF̃_\mu$ are themselves random terms to be specified further. The Fluctuation-Dissipation theorem (3.4.44) is

$$\begin{pmatrix} dP̃ \\ dẼ_1 \\ dẼ_2 \end{pmatrix} dP̃, \left(dẼ_1, dẼ_2\right) = \frac{2dt}{\beta} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \gamma_{00} & -V\gamma_{01} & -V\gamma_{02} \\ 0 & -V\gamma_{01} & V^2\gamma_{11} & V^2\gamma_{12} \\ 0 & -V\gamma_{02} & V^2\gamma_{12} & V^2\gamma_{22} \end{pmatrix} \tag{9.3.37}$$

By using the ansatz (9.3.36) we have

$$\begin{pmatrix} dF̃_1 + dF̃_2 \\ -V dF̃_1 \\ -V dF̃_2 \end{pmatrix} \left(dF̃_1 + dF̃_2, -V dF̃_1, -V dF̃_2\right)$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (dF̃_1 + dF̃_2)(dF̃_1 + dF̃_2) & -V(dF̃_1 + dF̃_2)dF̃_1 & -V(dF̃_1 + dF̃_2)dF̃_2 \\ 0 & -V(dF̃_1 + dF̃_2)dF̃_1 & V^2 dF̃_1 dF̃_1 & V^2 dF̃_1 dF̃_2 \\ 0 & -V(dF̃_1 + dF̃_2)dF̃_2 & V^2 dF̃_2 dF̃_1 & V^2 dF̃_2 dF̃_2 \end{pmatrix} \tag{9.3.38}$$

Therefore, we can always satisfy the Fluctuation-Dissipation theorem (9.3.37) by requiring

$$\frac{2dt}{\beta} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{pmatrix} dt = \begin{pmatrix} dF̃_1 \\ dF̃_2 \end{pmatrix} \left(dF̃_1, dF̃_2\right) \tag{9.3.39}$$

Here the stochastic forces are a linear combination of Wiener process

$$\begin{pmatrix} dF̃_1 \\ dF̃_2 \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix} \tag{9.3.40}$$

We fix the matrix coefficients a, b, c by inserting (9.3.40) into (9.3.39), this is

$$
\begin{aligned}
\frac{2dt}{\beta} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{pmatrix} dt &= \begin{pmatrix} d\tilde{F}_1 \\ d\tilde{F}_2 \end{pmatrix} \left(d\tilde{F}_1, d\tilde{F}_2 \right) \\
&= \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix} (dW_1, dW_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \\
&= \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} dt \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & (a+c)b \\ (a+c)b & b^2 + c^2 \end{pmatrix} dt
\end{aligned}
$$

$$(9.3.41)$$

Therefore, we have to satisfy

$$
\frac{2}{\beta} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \tag{9.3.42}
$$

this, is we need to take the square root of the first matrix. This is problem is easily solved by diagonalizing the matrix of friction coefficients.

The SDE

By collecting the drift term (9.3.13) with (9.3.17), and the dissipative term (9.3.35) we obtain the stochastic differential equation (3.4.41) of this level of description

$$
dX = V dt
$$

$$
dP = A \left[P_1 - P_2 \right] dt - V \left[\gamma_{01} \frac{1}{\beta T_1} + \gamma_{02} \frac{1}{\beta T_2} \right] dt + d\tilde{P}
$$

$$
dE_1 = -VAP_1 + V^2 \frac{\gamma_{12}}{\beta} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) dt + V^2 \frac{\gamma_{01}}{\beta} \frac{1}{T_1} dt + d\tilde{E}_1
$$

$$
dE_2 = -VAP_1 + V^2 \frac{\gamma_{12}}{\beta} \left(\frac{1}{T_2} - \frac{1}{T_1} \right) dt + V^2 \frac{\gamma_{02}}{\beta} \frac{1}{T_2} dt + d\tilde{E}_2 \tag{9.3.43}
$$

These are the stochastic evolution equations for the relevant variables. The physical significance of each term in these equations is very neat. For the sake of the discussion, let us for a while assume that the random terms $d\tilde{E}_1, d\tilde{E}_2, d\tilde{P}$ can be neglected. We will discuss latter when this is a reasonable approximation. In this way, the above equation become deterministic equations. The most salient aspect of the equations is that they conserve the total energy $E = E_1 + E_2 + \frac{P^2}{2M}$. The value of E is, therefore, fixed by the initial conditions for our variables $E_1(0), E_2(0), P(0)$. The deterministic equations also ensure that the total entropy (9.3.3) is a non-decreasing function of time. This can be proved by a direct calculation. Alternatively, we may check that the dissipative matrix (9.3.37) is positive definite. The deterministic equations therefore, dictate an evolution

in which the energy is conserved and the entropy always increases. This implies that the above system of equations has an equilibrium final state X^*, P^*, E_1^*, E_2^* which is determined by the maximum of the entropy subject to the constraint that the energy is conserved. This equilibrium state, therefore, maximizes in an unconstrained way the function

$$S(X, P, E_1, E_2) + \beta_0 \left(E_1 + E_2 + \frac{P^2}{2M} \right) \tag{9.3.44}$$

where β_0 is a Lagrange multiplier. By setting to zero the partial derivatives at the maximum

$$0 = \frac{\partial S}{\partial E_1} + \beta_0 = \frac{1}{T_1(N_1, V_1^*, E_1^*)} + \beta_0$$

$$0 = \frac{\partial S}{\partial E_2} + \beta_0 = \frac{1}{T_2(N_2, V_2^*, E_2^*)} + \beta_0$$

$$0 = \frac{\partial S}{\partial X} = A \left[\frac{P_1(N_1, V_1^*, E_1^*)}{T_1(N_1, V_1^*, E_1^*)} - \frac{P_2(N_2, V_2^*, E_2^*)}{T_2(N_2, V_2^*, E_2^*)} \right]$$

$$0 = \frac{\partial S}{\partial P} + \beta_0 \frac{P^*}{M} \tag{9.3.45}$$

Therefore, the equilibrium state is such that the temperatures and the pressures of the two systems are identical and the piston is at rest.

$$T_1(N_1, V_1^*, E_1^*) = T_1(N_2, V_2^*, E_2^*)$$
$$P_1(N_1, V_1^*, E_1^*) = P_1(N_2, V_2^*, E_2^*) \tag{9.3.46}$$

The above system of equations has as unknowns $V_1^*, V_2^*, E_1^*, E_2^*$. Together with the two equations $V_1^* + V_2^2 = V$ and $E_1^* + E_2^* = E$, we have sufficient information to obtain, in principle, the final values of the macroscopic state of the system.

We discuss now the physical meaning of all the terms in the SDE (9.3.43). Let us start with the reversible part of the dynamics. The reversible part of the dynamics in the momentum equation tells us that there is a net force depending on the difference of pressures in each system. In the energy equations, the terms $-VF_1 = \dot{V}_1 P_1$, $-VF_2 = \dot{V}_2 P_2$ are the energy per unit time lost by each gas in doing such a reversible work on the piston. This reversible work done on the piston produces a reversible change of internal energy.

The irreversible term is a friction force proportional to the velocity acting on the piston. The effective friction coefficient is given by

$$\left[\gamma_{01} \frac{1}{\beta T_1} + \gamma_{02} \frac{1}{\beta T_2} \right] \tag{9.3.47}$$

This friction force will stop the motion of the piston (in the absence of fluctuations). Therefore, any kinetic energy of the piston at the initial time will dissipate producing an increment of the energies E_1, E_2 through the terms

$$V^2 \frac{\gamma_{01}}{\beta} \frac{1}{T_1} \qquad\qquad V^2 \frac{\gamma_{02}}{\beta} \frac{1}{T_2} \qquad\qquad (9.3.48)$$

which are proportional to the square of the velocity. These terms may be understood as a viscous heating of the two gasses due to the friction with the piston.

In the deterministic case, the term in the first equation

$$V^2 \frac{\gamma_{12}}{\beta} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) = V^2 \frac{\gamma_{12}}{\beta} \left(\frac{T_2 - T_1}{T_1 T_2} \right) = \kappa(T_2 - T_1) \qquad (9.3.49)$$

produces an increment of E_1 whenever $T_1 < T_2$ and can be interpreted as a heat conduction term. When 2 is hotter, the energy of the system 1 increases. Note that the origin of this energy transfer from system 2 to system 1 is due entirely to the random motion of the piston. Recall that the piston is modeled as a structureless object and there is no heat conduction through the material of the piston. The coefficient κ can be interpreted as a heat conductivity coefficient. Quite remarkably, this conductivity is proportional to the velocity squared of the piston. This inusual dependence should be attributed to the special way in which energy is transferred from one system to the other, through the fluctuations of the velocity of the piston.

The equations (9.3.43) have been obtained by Gruber [?] by using a phenomenological approach. The advantage of a formulation based on statistical mechanics is the explicit expressions for the transport coefficients in terms of Green-Kubo formulae. Öttinger has given also dynamic equations in the context of the GENERIC framework. The proposal by Öttinger applies, however to a different type of underlying model in which thermal conduction is taking place throughout the material of the piston itself.

9.4   Energy and volume exchange when the piston is heat conducting

Description of the system

10

Stretching a single macromolecule
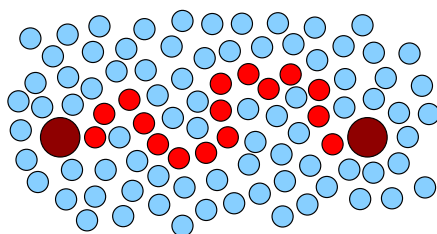
Figure 10.1: Cartoon of the system considered in this chapter. A linear macromolecule is pulled from to beads which are fixed in space and exert forces on the end monomers of the macromolecule. The macromolecule is floating in a solvent.

10.1 Description of the system

A linear macromolecule in solution has two ends subject to external forces. For simplicity, we assume that the forces at each end have the same direction and have opposite sign. The forces may be time dependent and, in particular, can be set instantaneously to zero. We seek for a CG description of how the system evolves in time. A cartoon of the problem is in Fig **??**.

Atomic Force Microscopy experiments have been conducted in which macromolecules like DNA or proteins have been subjected to external forces in order to study its elastic properties and relate these properties with relevant biological functions [**?**].

10.1.1   The microscopic level

At the microscopic level the state of the system $z = (z_1, z_2)$ is given by the positons $\mathbf{q}_{i_1}$ and momenta $\mathbf{p}_{i_1}$ of the N_1 monomers of the macromolecule and the the positons $\mathbf{q}_{i_2}$ and momenta $\mathbf{p}_{i_2}$ of the N_2 atoms of the surrounding fluid. The fluid is contained in a flask. The two particles at the ends of the macromolecule feel a force $\mathbf{F}^{\text{ext}}(t)$ equal and opposite. **[Warning! With this forcing, the center of mass of the molecule may diffuse away!]**

We need to anchor the beads at $\mathbf{Q}_1(t)$ and the force on monomer 1 is given by $k(\mathbf{Q}_1(t) - \mathbf{q}_{1_1})$. The two spring potentials are

$$\frac{k}{2}(\mathbf{Q}_1(t) - \mathbf{q}_{1_1})^2 + \frac{k}{2}(\mathbf{Q}_2(t) - \mathbf{q}_{N_1})^2 \tag{10.1.1}$$

Unfortunately, this external forcing is not linear in the external forcing!

The Hamiltonian of the system is

$$\hat{H}_t(z) = \hat{H}^{(1)}(z_1) + \hat{H}^{(2)}(z_1, z_2) - \mathbf{F}^{\text{ext}}(t)\cdot(\mathbf{q}_{N_1} - \mathbf{q}_{1_1}) \tag{10.1.2}$$

with

$$\hat{H}^{(1)}(z_1) = \sum_{i_1} \frac{\mathbf{p}_{i_1}^2}{2m_{i_1}} + \phi^{(1)}(q_1)$$

$$\hat{H}^{(2)}(z_1, z_2) = \sum_{i_2} \frac{\mathbf{p}_{i_2}^2}{2m_{i_2}} + \phi^{(2)}(q_2) + \sum_{i_1, i_2} \phi^I(\mathbf{q}_{i_1} - \mathbf{q}_{i_2}) \tag{10.1.3}$$

Here $\hat{H}^{(1)}(z_1)$ is the Hamiltonian of the macromolecule in vacuo and depends only on the degrees of freedom z_1 of the macromolecule. The interaction potential energy of the monomers of the macromolecule is $\phi^{(1)}(q_1)$. The Hamiltonian of the solvent is given by $\hat{H}^{(2)}(z_1, z_2)$ contains, in addition to the potential of interaction $\phi^{(2)}(q_2)$ between solvent particles, the interaction potential energy $\phi^I(q_1, q_2)$ between the solvent particles and the monomers.

The Hamiltonian (10.4.6) leads to the equations of motion

$$\dot{\mathbf{q}}_{i_1} = \frac{\mathbf{p}_{i_1}}{m_{i_1}}$$

$$\dot{\mathbf{p}}_{i_1} = \sum_{j_1} \hat{\mathbf{F}}_{i_1 j_1} + \sum_{j_2} \hat{\mathbf{F}}_{i_1 j_2} - \delta_{i_1 1_1}\mathbf{F}^{\text{ext}}(t) + \delta_{i_1 N_1}\mathbf{F}^{\text{ext}}(t)$$

$$\dot{\mathbf{q}}_{i_2} = \frac{\mathbf{p}_{i_2}}{m_{i_2}}$$

$$\dot{\mathbf{p}}_{i_2} = \sum_{j_1} \hat{\mathbf{F}}_{i_2 j_1} + \sum_{j_2} \hat{\mathbf{F}}_{i_2 j_2} \tag{10.1.4}$$

10.2   Some possible CG levels

As we have repeatedly emphasized throughout the book, the Theory of Coarse-Graining is like a black-box, in that given the list of CG variables, the theory provides its equations of motion. However there is no guarantee that the selected list is "correct". In fact, many other possible selections for the list of CG variables can be thought of in the system of a stretched macromolecule. The following are several possible levels of description for this system, attending to different possibilities for the selection of CG variables, in increasing order of resolution:

1. *Isothermal Thermodynamics level*, given by the total energy $\hat{H}^{(0)} = \hat{H}^{(1)} + \hat{H}^{(2)}$ and the end-to-end vector $\hat{\mathbf{R}}$.

2. *Non-isothermal Thermodynamics level*, given by $\hat{H}^{(1)}$, $\hat{H}^{(2)}$ and the end-to-end vector $\hat{\mathbf{R}}$.

3. *Non-isothermal Inertial Thermodynamics level*, given by $\hat{H}^{(1)}$, $\hat{H}^{(2)}$, the end-to-end vector $\hat{\mathbf{R}}$, and the relative velocity $\hat{\mathbf{V}}$.

4. *Hydrodynamic level*, where the solvent is described at the hydrodynamic level with the conserved mass, momentum, and energy fields $\hat{\rho}_{\mathbf{r}}, \hat{\mathbf{g}}_{\mathbf{r}}, \hat{e}_{\mathbf{r}}$ and the macromolecule is described with $\hat{\mathbf{R}}, \hat{\mathbf{V}}$.

Each level of description will result in a FPE governing the probability $P(a, t)$ of the variables. The first level of description will result in a FPE for the probability distribution of $E, \mathbf{R}$. Because the effect of "heating" the macromolecule as a consequence of the external forcing is expected to hardly affect the total ammount of energy E in the system, the dynamics of $\mathbf{R}$ will be just a diffusion process in isothermal conditions. This level of description is probably adequate to discuss situations where the external forcing is very "slow".

   If the external forcing is "rapid", in such a way that the macromolecule is "heated" then, including the energy of the macromolecule as one of the CG variables will allow to study the transfer of energy between the macromolecule and the system. The second level above, called Non-isothermal Thermodynamics level keeps track of the two energies $\hat{H}^{(1)}$ of the macromolecule and $\hat{H}^{(2)}$ of the solvent. On heuristic grounds, however it seems that the time scale of variation of the energy $\hat{H}^{(1)}$ of the macromolecule cannot be much different from the time scale of the relative velocity. This suggest that, *perhaps*, the third level of description that includes the velocity is more adequate to describe the energy transfer between the macromolecule and the solvent.

   Finally, it could very well happen that the "rapid" external forcing of the molecule would excite appreciable coherent motions of the solvent. If this is the case, we would need to include in the description the variables that describe these coherent motions. These variables are the mass, momentum, and energy densities of the solvent. If we think of an stretched rubber band inside a glass of water[1] that is released and snaps violently, we would clearly require a hydrodynamic description to keep track of what the fluid is doing in reaction to the rubber band. For nanoscopic macromolecules, however,

---

[1]This would be described by a Hamiltonian identical to (10.4.6) but with huge numbers N_1, N_2 of atoms.

the hydrodynamic description (that will be considered in detail in Chapter **??**) itself may break down. In this case, *perhaps*, the third level of description is the adequate one.

In the following subsections we consider the two levels of description 1 and 3 above.

10.3 Isothermal Thermodynamics level

The CG variables

The selected CG variables are

$$\hat{A}(z) \rightarrow \begin{cases} \hat{\mathbf{R}}(z_1) & = \mathbf{q}_N - \mathbf{q}_1 \\ \\ \hat{H}^{(0)}(z) & = \hat{H}^{(1)}(z_1) + \hat{H}^{(2)}(z_1, z_2) \end{cases} \tag{10.3.1}$$

that take numerical values

$$a \rightarrow \mathbf{R}, E \tag{10.3.2}$$

The CG variables are both even under time reversal, this is $\hat{A}_\mu(\epsilon \cdot z) = \varepsilon \cdot \hat{A}(z)$ with the time reversal signature $\varepsilon = (1, 1)$.

The CG energy

The CG energy (6.3.7) is

$$E(a, t) \rightarrow E(\mathbf{R}, E) = E - \mathbf{F}^{\mathrm{ext}}(t) \cdot \hat{\mathbf{R}} \tag{10.3.3}$$

with derivatives

$$\frac{\partial E}{\partial a_\mu}(a, t) \rightarrow (-\mathbf{F}^{\mathrm{ext}}(t), 1) \tag{10.3.4}$$

The time derivatives of the CG variables

The action of the Liouville operator $i\mathcal{L}_t$ (1.4.3) on the total energy of the unperturbed system is

$$i\mathcal{L}_t \hat{\mathbf{R}} = \hat{\mathbf{V}}(z_1)$$

$$i\mathcal{L}_t \hat{H}^{(0)} = -\frac{\partial \hat{H}_t}{\partial z} \cdot J \cdot \frac{\partial \hat{H}^{(0)}}{\partial z} = -\frac{\partial (\hat{H}^{(0)} - \mathbf{F}^{\mathrm{ext}} \cdot \hat{\mathbf{R}})}{\partial z} \cdot J \cdot \frac{\partial \hat{H}^{(0)}}{\partial z} = -\mathbf{F}^{\mathrm{ext}}(t) \cdot \hat{\mathbf{V}}(z_1) \tag{10.3.5}$$

The entropy

The entropy (6.3.5) is given by

$$S_B(a) \rightarrow S(\mathbf{R}, E) = \ln \int dz \delta(\hat{R} - \mathbf{R}) \delta(\hat{H}^{(0)}(z) - E) \tag{10.3.6}$$

Note that the entropy depends parametrically on the time-independent external force. We may express this entropy in terms of the microcanonical entropy $\ln \Omega(E)$ by introducing the equilibrium distribution $P^{\mathrm{eq}}(\mathbf{R})$ of $\mathbf{R}$ in the microcanonical ensemble which is

$$P^{\mathrm{eq}}(\mathbf{R}) = \frac{\Omega(\mathbf{R}, E)}{\Omega(E)}$$

$$\Omega(\mathbf{R}, E) = \int dz \delta(\hat{H}^0(z) - E) \delta(\hat{\mathbf{R}} - \mathbf{R})$$

$$\Omega(E) = \int dz \delta(\hat{H}^0(z) - E) \tag{10.3.7}$$

In this way, the entropy is

$$S_B(\mathbf{R}, E) = S_B^{(0)}(E) + \ln P^{\mathrm{eq}}(\mathbf{R}) \tag{10.3.8}$$

where the usual Boltzmann entropy of the system is

$$S_B^{(0)}(E) = \ln \Omega(E) = \int dz \delta(\hat{H}^0(z) - E) \tag{10.3.9}$$

We should bear in mind that all the quantitites entering (10.3.8) are defined in terms of the Hamiltonian $\hat{H}^{(0)}(z)$ and, therefore, they depend parametrically on the time-independent external forcing $\mathbf{F}^{\mathrm{ext}}$. Note the the first term should scale as the number of particles in the solvent while the second term does not scale as the number of particles. We cannot, however, neglect the second term, to the risk of throwing the baby with the bath water. In fact, the derivatives of the entropy are

$$\frac{\partial S_B}{\partial a_\mu}(a) \rightarrow \begin{cases} \frac{\partial S_B}{\partial \mathbf{R}} = \frac{\partial \ln P^{\mathrm{eq}}(\mathbf{R})}{\partial \mathbf{R}} \equiv \frac{\overline{\mathbf{F}}(\mathbf{R}, E)}{T_0(E)} \\[2mm] \frac{\partial S_B}{\partial E} = \frac{1}{T_0(E)} + \frac{\partial}{\partial E} \ln P^{\mathrm{eq}}(\mathbf{R}) \end{cases} \tag{10.3.10}$$

where T_0 is the microcanonical temperature and in the last line defines the mean force $\overline{\mathbf{F}}(E, \mathbf{R})$. Should we neglect the second term in the entropy (10.3.8) we would not have any mean force, which is crucial for the dynamics of $\mathbf{R}$. We argue, nevertheless that the second contribution in the temperature is negligible if the solvent system is sufficient large. Heuristically, we have a derivative with respect to an extensive variable E of an object that is size independent[2]. Therefore, we will neglect this term.

---

[2] We can use equivalence of ensembles, substitute the microcanonical ensemble with a canonical one, and perform the derivative. It will throw out a factor inversely proportional to the heat capacity, the latter scaling with the size of the system.

The reversible drift term

The reversible drift term (6.3.2) is given by

$$V_\mu(a) \to \begin{cases} V_{\mathbf{R}}(\mathbf{R}, E) = & \left\langle i\mathcal{L}_t\hat{\mathbf{R}} \right\rangle^{\mathbf{R},E} \\ V_E(\mathbf{R}, E) = & \left\langle i\mathcal{L}_t\hat{H}^{(0)} \right\rangle^{\mathbf{R},E} \end{cases} \tag{10.3.11}$$

where the conditional expectation is defined as

$$\langle \cdots \rangle^a \to \frac{1}{\Omega(\mathbf{R}, E)} \int dz \delta(\hat{\mathbf{R}} - \mathbf{R})\delta(\hat{H}^{(0)} - E) \cdots \tag{10.3.12}$$

Because the time derivatives (10.3.5) are linear in the velocities and the Hamiltonian is even, the conditional expectations vanish. Therefore, the reversible drift vanishes in this level of description.

The dissipative matrix

The dissipative matrix (6.3.11) is

$$D_{\mu\nu}(a, t) \to T_0 \begin{pmatrix} D_{\mathbf{RR}} & D_{\mathbf{R}E} \\ D_{E\mathbf{R}} & D_{EE} \end{pmatrix} \tag{10.3.13}$$

where the different elements are given in terms of the Green-Kubo expressions

$$D_{\mathbf{RR}}(a, t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \delta i\mathcal{L}_t\hat{\mathbf{R}} \exp\{i\mathcal{L}_t t'\}\delta i\mathcal{L}_t\hat{\mathbf{R}} \right\rangle^a$$

$$D_{E\mathbf{R}}(a, t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \delta i\mathcal{L}_t\hat{H}^{(0)} \exp\{i\mathcal{L}_t t'\}\delta i\mathcal{L}_t\hat{\mathbf{R}} \right\rangle^a$$

$$D_{E\mathbf{R}}(a, t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \delta i\mathcal{L}_t\hat{\mathbf{R}} \exp\{i\mathcal{L}_t t'\}\delta i\mathcal{L}_t\hat{H}^{(0)} \right\rangle^a$$

$$D_{EE}(a, t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \delta i\mathcal{L}_t\hat{H}^{(0)} \exp\{i\mathcal{L}_t t'\}\delta i\mathcal{L}_t\hat{H}^{(0)} \right\rangle^a \tag{10.3.14}$$

where we have introduce the temperature factor T_0 in order to have contact with usual definitions of diffusion coefficients. In particular, we introduce the following diffusion tensor

$$\mathbf{D}(a, t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \hat{\mathbf{V}} \exp\{i\mathcal{L}_t t'\}\hat{\mathbf{V}} \right\rangle^a \tag{10.3.15}$$

that involves the autocorrelation function of the relative velocity. This diffusion coefficient has the usual dimensions of length squared over time.

By using this diffusion coefficient and the form of the time derivatives (10.3.5) in the

Green-Kubo expressions (10.3.14) the dissipative matrix (10.3.13)

$$D_{\mu\nu}(a,t) \to T_0 \begin{pmatrix} \mathbf{D}(a,t) & -\mathbf{D}(a,t)\cdot\mathbf{F}^{\mathrm{ext}}(t) \\ -\mathbf{F}^{\mathrm{ext}}(t)\cdot\mathbf{D}(a,t) & \mathbf{F}^{\mathrm{ext}}(t)\cdot\mathbf{D}(a,t)\cdot\mathbf{F}^{\mathrm{ext}}(t) \end{pmatrix} \qquad (10.3.16)$$

Because the time signature of the CG variables is $\varepsilon = (1,1)$, Onsager's reciprocity (6.3.12) implies that this matrix is symmetric which implies, in turn, that the diffusion tensor (10.3.15) is symmetric. It is manifestly semi-definite positive, satisfying the degeneracy (4.4.25).

In principle, this diffusion coefficient depends on the state $a = \mathbf{R}, E$ and it depends on time because the dynamics is happening under the external forcing that depends on time. Recall that implicitly, it also depends on the time-independent external force. We expect that the unidirectionality imposed by the external force translates into a particular axisymmetric structure for the diffusion tensor

$$\mathbf{D}(\mathbf{R}, E, t) = D_{||}(E, R, t)\mathbf{nn} + D_{\perp}(E, R, t)(\mathbf{1} - \mathbf{nn}) \qquad (10.3.17)$$

Here, $\mathbf{n} = \mathbf{R}/R$ is the unit vector and $R = |\mathbf{R}|$ is the end to end distance. With this assumption, from the six elements of the diffusion tensor, only two become important. By contracting the tensor with the unit vectors $\mathbf{n}$ and $\mathbf{t}$ (any unit vector normal to $\mathbf{n}$) we obtain the Green-Kubo expressions for the two diffusion coefficients

$$D_{||}(a,t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \hat{V}_{||} \exp\{i\mathcal{L}_t t'\}\hat{V}_{||} \right\rangle^a$$

$$D_{\perp}(a,t) = \frac{1}{T_0} \int_0^{\Delta t} dt' \left\langle \hat{V}_{\perp} \exp\{i\mathcal{L}_t t'\}\hat{V}_{\perp} \right\rangle^a \qquad (10.3.18)$$

where $\hat{V}_{||} = \mathbf{r}\cdot\mathbf{V}$ and $\hat{V}_{\perp} = \mathbf{t}\cdot\mathbf{V}$.

The spurious drift

Because the dissipative matrix depends on the state, the spurious drift that enters the SDE (6.3.15) is given in the present problem

$$\frac{\partial D_{\mu\nu}^S}{\partial a_\nu} \to T_0 \begin{pmatrix} \frac{\partial \mathbf{D}(a,t)}{\partial E}\cdot\mathbf{F}^{\mathrm{ext}}(t) - \frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t) \\ \mathbf{F}^{\mathrm{ext}}(t)\cdot\frac{\partial \mathbf{D}(a,t)}{\partial E}\cdot\mathbf{F}^{\mathrm{ext}}(t) - \mathbf{F}^{\mathrm{ext}}(t)\cdot\frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t) \end{pmatrix} \qquad (10.3.19)$$

We will neglect the terms containing derivatives with respect to the extensive variable E as they scale as the inverse of the system size. Therefore, the spurious drift becomes

$$\frac{\partial D_{\mu\nu}^S}{\partial a_\nu} \to T_0 \begin{pmatrix} -\frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t) \\ -\mathbf{F}^{\mathrm{ext}}(t)\cdot\frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t) \end{pmatrix} \qquad (10.3.20)$$

The irreversible drift

By multiplying the dissipative matrix (10.3.13) with the derivatives of the entropy we obtain the reversible drift

$$D_{\mu\nu}(a,t)\frac{\partial S_B}{\partial a_\nu} \to T_0 \begin{pmatrix} \mathbf{D}(a,t) & -\mathbf{D}(a,t)\cdot\mathbf{F}^{\text{ext}}(t) \\ -\mathbf{F}^{\text{ext}}(t)\cdot\mathbf{D}(a,t) & \mathbf{F}^{\text{ext}}(t)\cdot\mathbf{D}(a,t)\cdot\mathbf{F}^{\text{ext}}(t) \end{pmatrix} \begin{pmatrix} \frac{\overline{\mathbf{F}}}{T_0} \\ \frac{1}{T_0} \end{pmatrix} \quad (10.3.21)$$

The noise

The noise is readily inferred from the structure of the time derivatives (10.3.5), this is, we guess that

$$d\mathbf{R}(t) = (2\mathbf{D}(a,t))^{1/2}d\mathbf{W}(t)$$
$$d\tilde{E}_0(t) = -\mathbf{F}^{\text{ext}}(t)\cdot d\mathbf{R}(t) \quad (10.3.22)$$

where $d\mathbf{W}(t)$ is a vector of independent Wiener processes. The square root of the diffusion tensor exists because the matrix is semi-definite positive. With the model (10.3.17) for the diffusion tensor a much simpler form that does not require to compute the square root is given by the following noise term

$$d\mathbf{R}(t) = (2D_{||})^{1/2}d\mathbf{V}(t) + (2D_\perp)^{1/2}\mathbf{n}\times d\mathbf{V}(t) \quad (10.3.23)$$

The covariance of this noise is given by

$$d\mathbf{R}(t)d\mathbf{R}^T(t) = 2D_{||}d\mathbf{V}(t)d\mathbf{V}(t) + 2D_\perp^{1/2}(\mathbf{r}\times d\mathbf{V}(t))(\mathbf{n}\times d\mathbf{V}(t))$$
$$= 2(D_{||}\mathbf{nn} + D_\perp(\mathbf{1} - \mathbf{nn}))dt \quad (10.3.24)$$

where we have used $(\mathbf{n}\times d\mathbf{V})_\mu = \epsilon_{\mu ab}n_a dV_b$, where $\epsilon_{\mu ab}$ is the Levi-Civitta symbol, the Ito rule $dV_a dV_b = \delta_{ab}dt$, and the identity $\epsilon_{\mu ab}\epsilon_{\nu a'b} = \delta_{aa'}\delta_{\mu\nu} - \delta_{a\nu}\delta_{a'\mu}$.

The final SDE

The SDE (6.3.15) is obtained now by collecting the results for the irreversible drift (10.3.21) for the spurious drift (10.3.20) and for the noise terms (10.3.22)

$$d\mathbf{R}(t) = -\mathbf{D}(a,t)\cdot(\mathbf{F}^{\text{ext}}(t) - \overline{\mathbf{F}})dt - \frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t)dt + d\mathbf{R}(t)$$

$$dE(t) = \mathbf{F}^{\text{ext}}(t)\cdot\mathbf{D}(a,t)\cdot(\mathbf{F}^{\text{ext}}(t) - \overline{\mathbf{F}}) - \mathbf{F}^{\text{ext}}(t)\cdot\frac{\partial \mathbf{D}}{\partial \mathbf{R}}(a,t) + d\tilde{E}_0(t) \quad (10.3.25)$$

Discussion

Note that, quite naturally, the equation for the energy of the system can be expressed as

$$dE(t) = \mathbf{F}^{\text{ext}}(t)\cdot d\mathbf{R}(t) \quad (10.3.26)$$

this is, the change in the energy of the system is due entirely to the work done by the external force. Note also, that the energy E scales with the size of the system, whereas the right hand side of (10.3.26) does not depend on the system. This means that for very large systems, the energy E may be taken as practically constant. The value of E fixes the temperature $T_0(E)$ of the system which will remain constant.

There are three ingredients in the SDE (10.3.25), the mean force, and the two diffusion coefficients. The mean force, defined in (10.3.10), is given by

$$\overline{\mathbf{F}}(\mathbf{R}) = T_0 \frac{\partial \ln P^{\mathrm{eq}}(\mathbf{R})}{\partial \mathbf{R}} \tag{10.3.27}$$

The probability $P^{\mathrm{eq}}(\mathbf{R})$ is a bell-shaped function peaked at the average value $\overline{\mathbf{R}}(\mathbf{F}^{\mathrm{ext}})$ that is, obviously a function of the constant external force. Imagine now the following experiment. The system is initially under the constant force $\mathbf{F}^{\mathrm{ext}}$ and at $t = 0$ we apply to the ends of the stretched macromolecule another constant force $\Delta\mathbf{F}^{\mathrm{ext}} = -\mathbf{F}^{\mathrm{ext}}$, that is, we release the macromolecule. The

The sampling problem

Consider the release experiment in which we start from an end-to-end vector t is far away from the equilibrium value. One problem that we encounter is that we need to know, for updating purposes with the SDE, the value of the mean force at that highly non-equilibrium value. This means that we need to know the value of $P^{\mathrm{eq}}(\mathbf{R})$ at the tails, where the sampling of the probability is very poor.

$$\begin{aligned} P^{\mathrm{eq}}(\mathbf{R}) &= \int dz \rho_E(z) \delta(\hat{\mathbf{R}} - \mathbf{R}) \simeq \int dz \frac{1}{Z(\beta_0)} \exp\{-\beta_0 \hat{H}^{(0)}(z)\} \delta(\hat{\mathbf{R}} - \mathbf{R}) \\ &= \exp\{\beta_0 \mathbf{F}^{\mathrm{ext}}\cdot\mathbf{R}\} \int dz \frac{1}{Z(\beta_0)} \exp\{-\beta_0(\hat{H}^{(0)}(z) - \mathbf{F}^{\mathrm{ext}}\cdot\hat{\mathbf{R}})\} \delta(\hat{\mathbf{R}} - \mathbf{R}) \\ &= \exp\{\beta_0 \mathbf{F}^{\mathrm{ext}}\cdot\mathbf{R}\} \frac{Z_F(\beta_0)}{Z(\beta_0)} \int dz \frac{1}{Z_F(\beta_0)} \exp\{-\beta_0(\hat{H}^{(0)}(z) - \mathbf{F}^{\mathrm{ext}}\cdot\hat{\mathbf{R}})\} \delta(\hat{\mathbf{R}} - \mathbf{R}) \\ &= \exp\{\beta_0 \mathbf{F}^{\mathrm{ext}}\cdot\mathbf{R}\} \frac{Z_F(\beta_0)}{Z(\beta_0)} P_F(\mathbf{R}) \end{aligned} \tag{10.3.28}$$

where $P_F(\mathbf{R})$ is the probability under the external forcing. We may run different simulations for different values of $\mathbf{F}^{\mathrm{ext}}$ and easily get $P_F(\mathbf{R})$. Finally, we would estimate the mean force from

$$\overline{\mathbf{F}}(\mathbf{R}) = \mathbf{F}^{\mathrm{ext}} + T_0 \frac{\partial \ln P_F}{\partial \mathbf{R}}(\mathbf{R}) \tag{10.3.29}$$

Note that the left hand side is independent on $\mathbf{F}^{\mathrm{ext}}$. For a fixed value of $\mathbf{F}^{\mathrm{ext}}$, around the average position $\overline{\mathbf{R}}(\mathbf{F}^{\mathrm{ext}})$ corresponding to this value, we will have a good sampling of $\overline{\mathbf{F}}(\mathbf{R})$, but a bad representation for values of $\mathbf{R}$ away from this average value. By "moving" the value of the external force. should allow one to recover with good precission the mean force for all values of $\mathbf{R}$.

Consider the correlations that appear in the diffusion coefficients (10.3.18). They

have the structure

$$\left\langle \hat{V} \exp\{i\mathcal{L}_t t'\}\hat{V} \right\rangle^{\mathbf{R},E} = \int dz \frac{1}{\Omega(\mathbf{R},E)} \delta(\hat{H}^{(0)}(z) - E)\delta(\hat{\mathbf{R}}(z) - \mathbf{R})\hat{V}(z)\exp\{i\mathcal{L}_t t'\}\hat{V}(z)$$

$$= \frac{1}{P^{\mathrm{eq}}(\mathbf{R})} \int dz \rho^{\mathrm{eq}}(z)\delta(\hat{\mathbf{R}}(z) - \mathbf{R})\hat{V}(z)\exp\{i\mathcal{L}_t t'\}\hat{V}(z) \quad (10.3.30)$$

We encounter here also the same sampling problem. The microstates that give high weight to the ensemble are only those corresponding to values of $\mathbf{R}$ near equilibrium. If we need to evaluate the diffusion coefficient at values far from this equilibrium value, we run into a sampling problem that would require prohibitely large simulations. The trick now is the same, use the canonical ensemble and multiply numerator and denominator with the factor $\exp\{-\beta\mathbf{F}^{\mathrm{ext}}\}$. This gives

$$\left\langle \hat{V} \exp\{i\mathcal{L}_t t'\}\hat{V} \right\rangle^{\mathbf{R},E} = \int dz \frac{1}{\Omega(\mathbf{R},E)} \delta(\hat{H}^{(0)}(z) - E)\delta(\hat{\mathbf{R}}(z) - \mathbf{R})\hat{V}(z)\exp\{i\mathcal{L}_t t'\}\hat{V}(z)$$

$$= \frac{1}{P_F^{\mathrm{eq}}(\mathbf{R})} \int dz \frac{1}{Z_F} \exp\{-\beta(\hat{H}^{(0)}(z) - \mathbf{F}^{\mathrm{ext}}\cdot\mathbf{R}\}\delta(\hat{\mathbf{R}}(z) - \mathbf{R})\hat{V}(z)\exp\{i\mathcal{L}_t t'\}\hat{V}(z)$$

$$(10.3.31)$$

For symmetry reasons, we expect that the dependence on $\mathbf{R}$ will be only through $R = \mathbf{R}\cdot\mathbf{n}$. Therefore the strategy is as follows. Instead of the Dirac delta function, use an approximation by bins, this is

$$\delta(\hat{\mathbf{R}}(z) - \mathbf{R}) \simeq \chi_\mu(\mathbf{n}\cdot\hat{\mathbf{R}}(z)) \quad (10.3.32)$$

where $\chi_\mu(R)$ is a discrete approximation of the form

$$\chi_\mu(R) = \frac{1}{\Delta R}\theta(R - R_\mu)\theta(R_{\mu+1} - R) \quad (10.3.33)$$

that takes the value 1 where R happens to be in the bin μ between $R_{\mu+1}$ and R_μ. In this way, we estimate the conditional correlation (10.3.31) as a set of ordinary correlation function of the function $\chi_\mu\hat{V}$ with $\hat{V}$ for each μ, this is

$$\left\langle \hat{V} \exp\{i\mathcal{L}_t t'\}\hat{V} \right\rangle^{\mathbf{R},E} \simeq \frac{1}{P^{\mathrm{eq}}(R)} \left\langle \chi_\mu \hat{V} \exp\{i\mathcal{L}_t t'\}\hat{V} \right\rangle^\beta \quad (10.3.34)$$

These correlations need to be performed for a range of different external forcings in order to obtain a good sampling for all required values of R.

10.4   Non-isothermal Inertial Thermodynamics level

The CG variables

We choose as CG variables

$$\hat{A}_\mu(z) \to \{\hat{\mathbf{R}}(z_1), \hat{\mathbf{V}}(z_1), \hat{H}^{(1)}(z_1), \hat{H}^{(2)}(z_1, z_2)\} \tag{10.4.1}$$

where the end-to-end vector $\hat{\mathbf{R}}(z_1)$ and the relative velocity $\hat{\mathbf{V}}(z_1)$ are defined as

$$\begin{aligned}
\hat{\mathbf{R}}(z_1) &= \mathbf{q}_{N_1} - \mathbf{q}_{1_1} \\
\hat{\mathbf{V}}(z_1) &= \frac{\mathbf{p}_{N_1}}{m_{N_1}} - \frac{\mathbf{p}_{1_1}}{m_{1_1}}
\end{aligned} \tag{10.4.2}$$

Why do we choose these variables? We include the end to end distance because it appears directly as the coupling function with the external force. As it is obvious that the force has a direct and observable effect on the end to end distance, we should include this variable in the description. However, what is the reason for including the other set of variables?

We choose the relative velocity because it is the time derivative of the coupling function $\hat{\mathbf{R}}(z_1)$. We could choose as an alternative the relative momentum $\hat{\mathbf{P}} = \hat{\mathbf{p}}_{N_1} - \mathbf{p}_{1_1}$ but this is not the time derivative, except for the case that the masses are equal $m_{N_1} = m_{1_1}$.

The time signature of this set of variables is $\varepsilon = (1, -1, 1, 1)$

The CG energy

The Hamiltonian (**??**) can be expressed in terms of the CG variables in the form

$$\hat{H}_t(z) = \hat{E}_1(z) + \hat{E}_2(z) - \mathbf{F}^{\text{ext}}(t) \cdot \hat{\mathbf{R}} \tag{10.4.3}$$

The CG energy is

$$E(a) = E_1 + E_2 - \mathbf{F}^{\text{ext}} \cdot \mathbf{R} \tag{10.4.4}$$

The gradient of the energy is

$$\frac{\partial E}{\partial a} = (-\mathbf{F}^{\text{ext}}, 0, 1, 1) \tag{10.4.5}$$

The time derivatives of the CG variables

The Hamiltonian

$$\hat{H}_t(z) = \hat{H}^{(1)}(z_1) + \hat{H}^{(2)}(z_1, z_2) - \mathbf{F}^{\text{ext}}(t)\cdot(\mathbf{q}_{N_1} - \mathbf{q}_{1_1}) \tag{10.4.6}$$

with

$$\hat{H}^{(1)}(z_1) = \sum_{i_1} \frac{\mathbf{p}_{i_1}^2}{2m_{i_1}} + \phi^{(1)}(q_1)$$

$$\hat{H}^{(2)}(z_1, z_2) = \sum_{i_2} \frac{\mathbf{p}_{i_2}^2}{2m_{i_2}} + \phi^{(2)}(q_2) + \sum_{i_1,i_2} \phi^I(\mathbf{q}_{i_1} - \mathbf{q}_{i_2}) \tag{10.4.7}$$

leads to the equations of motion

$$\dot{\mathbf{q}}_{i_1} = \frac{\mathbf{p}_{i_1}}{m_{i_1}}$$

$$\dot{\mathbf{p}}_{i_1} = \sum_{j_1} \hat{\mathbf{F}}_{i_1 j_1} + \sum_{j_2} \hat{\mathbf{F}}_{i_1 j_2} - \delta_{i_1 1_1} \mathbf{F}^{\text{ext}}(t) + \delta_{i_1 N_1} \mathbf{F}^{\text{ext}}(t)$$

$$\dot{\mathbf{q}}_{i_2} = \frac{\mathbf{p}_{i_2}}{m_{i_2}}$$

$$\dot{\mathbf{p}}_{i_2} = \sum_{j_1} \hat{\mathbf{F}}_{i_2 j_1} + \sum_{j_2} \hat{\mathbf{F}}_{i_2 j_2} \tag{10.4.8}$$

$$i\mathcal{L}_t\hat{\mathbf{R}} = \hat{\mathbf{V}}$$

$$i\mathcal{L}_t\hat{\mathbf{V}} = \frac{1}{m_{N_1}}\hat{\mathbf{F}}_{N_1} - \frac{1}{m_{1_1}}\hat{\mathbf{F}}_{1_1} + \frac{1}{\mu}\mathbf{F}^{\text{ext}}(t) = \frac{1}{\mu}\hat{\mathbf{F}}^{\text{ends}} + \frac{1}{\mu}\mathbf{F}^{\text{ext}}(t)$$

$$i\mathcal{L}_t\hat{E}_1 = \hat{Q} + \mathbf{F}^{\text{ext}}(t)\cdot\hat{\mathbf{V}}$$

$$i\mathcal{L}_t\hat{E}_2 = -\hat{Q} \tag{10.4.9}$$

where we have introduced the reduced mass **[Hay un signo incorrecto en $\mathbf{F}^{\text{ext}}(t)$]**

$$\frac{1}{\mu} \equiv \left[\frac{1}{m_{N_1}} + \frac{1}{m_{1_1}}\right] \tag{10.4.10}$$

and $\hat{\mathbf{F}}_{1_1}$ $(\hat{\mathbf{F}}_{N_1})$ is the force due to other atoms on the first (last) atom of the molecule. Finally, we have introduced the microscopic heat as

$$\hat{Q} \equiv \frac{1}{2}\sum_{i_1 j_2} \mathbf{v}_{i_1 j_2}\cdot\mathbf{F}_{i_1 j_2} \tag{10.4.11}$$

Finally we have introduced the force on the ends of the molecule as

$$\hat{\mathbf{F}}^{\text{ends}} = \mu \left[\frac{1}{m_{N_1}} \hat{\mathbf{F}}_{N_1} - \frac{1}{m_{1_1}} \hat{\mathbf{F}}_{1_1} \right] \qquad (10.4.12)$$

The entropy

The entropy is given by

$$S(\mathbf{R}, \mathbf{V}, E_1, E_2) = k_B \ln \int dz_1 dz_2 \delta \left(\mathbf{R} - \hat{\mathbf{R}}(z_1) \right) \delta \left(\mathbf{V} - \hat{\mathbf{V}}(z_1) \right)$$
$$\times \delta \left(E_1 - \hat{H}^{(1)}(z_1) \right) \delta \left(E_2 - \hat{H}^{(2)}(z_1, z_2) \right) \qquad (10.4.13)$$

Introduce the following structure factors

$$\Omega(E_1, E_2) = \int dz_1 dz_2 \delta(\hat{H}^{(1)}(z_1) - E_1) \delta(\hat{H}^{(2)}(z_1, z_2) - E_2)$$

$$\Omega^{(1)}(E_1) = \int dz_1 \delta(\hat{H}^{(1)}(z_1) - E_1)$$

$$\Omega^{(2)}(E_2) = \int dz_2 \delta \left(\sum_{i_2} \frac{\mathbf{P}_{i_2}^2}{m_{i_2}} + \phi^{(2)}(z_2) - E_1 \right) \qquad (10.4.14)$$

Note that $\Omega^{(2)}(E_2)$ does not include the potential energy of interaction between system 1 and 2. Therefore, $\Omega^{(1)}(E_1), \Omega^{(2)}(E_2)$ are the structure functions of the independent systems.

By multiplying and dividing with respect to $\Omega(E_1, E_2)$ and $\Omega^{(1)}(E_1)\Omega^{(2)}(E_2)$ we may obtain

$$S(\mathbf{R}, \mathbf{V}, E_1, E_2) = S^{(1)}(E_1) + S^{(2)}(E_2) + S^{(I)}(E_1, E_2) + \ln P(\mathbf{R}, \mathbf{V}, E_1, E_2) \quad (10.4.15)$$

where we have introduced the following conditional probability distribution

$$P(\mathbf{R}, \mathbf{V}; E_1, E_2) = \int dz_1 dz_2 \frac{\delta \left(\hat{H}^{(1)}(z_1) - E_1 \right) \delta \left(\hat{H}^{(2)}(z_1, z_2) - E_2 \right)}{\Omega(E_1, E_2)}$$
$$\times \delta \left(\mathbf{R} - \hat{\mathbf{R}}(z_1) \right) \delta \left(\mathbf{V} - \hat{\mathbf{V}}(z_1) \right) \qquad (10.4.16)$$

which is normalized according to

$$\int d\mathbf{R} d\mathbf{V} P(\mathbf{R}, \mathbf{V}; E_1, E_2) = 1 \qquad (10.4.17)$$

and we have introduced the interaction bit of the entropy as

$$S^{(I)}(E_1, E_2) = \ln \frac{\Omega(E_1, E_2)}{\Omega^{(1)}(E_1)\Omega^{(2)}(E_2)} \qquad (10.4.18)$$

Note that if we may neglect the interaction potential, then $S^{(I)}(E_1, E_2) = 0$.

$$\frac{\partial S_B}{\partial a_\mu}(a) \rightarrow \begin{cases} \frac{\partial S_B}{\partial \mathbf{R}} & \equiv \frac{\overline{\mathbf{F}}(\mathbf{R})}{T_1} \\[2mm] \frac{\partial S_B}{\partial \mathbf{V}} & = \mu \frac{\mathbf{V}}{T_1} \\[2mm] \frac{\partial S_B}{\partial E_1} & \equiv \frac{1}{T_1} \\[2mm] \frac{\partial S_B}{\partial E_2} & \equiv \frac{1}{T_2} \end{cases} \tag{10.4.19}$$

In blue are guesses. While $\overline{\mathbf{F}}(\mathbf{R})$ may be taken as a definition, the presence of the temperature T_1 (and not any other one) seems to be dictated by the reversibility condition below. The linear term in the velocity is because we assume a Gaussian in $\mathbf{V}$ for the probability $P(\mathbf{R}, \mathbf{V}; E_1, E_2)$.

The reversible drift term

$$\begin{aligned} \left\langle i\mathcal{L}_t \hat{\mathbf{R}} \right\rangle^a &= \mathbf{V} \\ \left\langle i\mathcal{L}_t \hat{\mathbf{V}} \right\rangle &= \frac{1}{\mu} \left\langle \hat{\mathbf{F}}^{\text{ends}} \right\rangle^a + \frac{1}{\mu} \mathbf{F}^{\text{ext}}(t) \\ \left\langle i\mathcal{L}_t \hat{E}_1 \right\rangle^a &= \mathbf{F}^{\text{ext}}(t) \cdot \mathbf{V} \\ \left\langle i\mathcal{L}_t \hat{E}_2 \right\rangle^a &= 0 \end{aligned} \tag{10.4.20}$$

where we have that $\left\langle \hat{Q} \right\rangle^a = 0$ because $\hat{Q}$ is a linear function of the velocities, which vanish by symmetry, as the Hamiltonian is quadratic in the velocities.

The reversibility condition

Note that the drift term (10.4.20) satisfies first identity in (**??**) expressing energy conservation. The second identity in (**??**) is the reversibility condition

$$V_\mu(a, \phi_t) \frac{\partial S}{\partial a_\mu}(a) = 0 \tag{10.4.21}$$

$$\mathbf{V} \cdot \frac{\overline{\mathbf{F}}(\mathbf{R})}{T_1} + \frac{1}{\mu} \left\langle \hat{\mathbf{F}}^{\text{ends}} \right\rangle^a \cdot \mu \frac{\mathbf{V}}{T_1} + \frac{1}{\mu} \mathbf{F}^{\text{ext}}(t) \cdot \mu \frac{\mathbf{V}}{T_1} + \mathbf{F}^{\text{ext}}(t) \cdot \mathbf{V} \frac{1}{T_1} = 0 \tag{10.4.22}$$

[**Hay un signo incorrecto en $\mathbf{F}^{\text{ext}}(t)$.**] Therefore, we arrive at the following expression

$$\overline{\mathbf{F}}(\mathbf{R}) = \left\langle \delta \hat{\mathbf{F}} \right\rangle^a \tag{10.4.23}$$

This allows to interprete $\overline{\mathbf{F}}(\mathbf{R})$ as the conditional average of the force acting on the extreme beads due to the surrounding fluid.

The dissipative matrix

The dissipative matrix (6.3.11) is very similar in structure to the forced piston. We note that because the time derivative of the position of the piston is the momentum, which is itself a relevant variable, the projected current concerning the position will vanish. Therefore the dissipative matrix will have the form

$$
D_{\mu\nu}(a) \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & D_{\mathbf{VV}} & D_{\mathbf{V}1} & D_{\mathbf{V}2} \\ 0 & D_{1\mathbf{V}} & D_{11} & D_{12} \\ 0 & D_{2\mathbf{V}} & D_{21} & D_{22} \end{pmatrix} \tag{10.4.24}
$$

where the vanishing elements are due to the fact that the time derivative of the end to end distancy is the velocity, which is itself a relevant variable. The different elements are given in terms of Green-Kubo expressions

$$
D_{\mathbf{VV}}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{\mathbf{V}} \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L} \hat{\mathbf{V}} \right\rangle^a
$$

$$
D_{\mu\mathbf{V}}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{H}_\mu \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L} \hat{\mathbf{V}} \right\rangle^a \qquad \mu = 1,2
$$

$$
D_{\mu\nu}(a,t) = \int_0^{\Delta t} dt' \left\langle \delta i \mathcal{L}_t \hat{H}_\mu \exp\{i\mathcal{L}_t t'\} \delta i \mathcal{L}_t \hat{H}_\nu \right\rangle^a \qquad \mu,\nu = 1,2 \tag{10.4.25}
$$

The fluctuations appearing in these correlations are $\delta\hat{G}(z) = \hat{G}(z) - \langle\hat{G}\rangle^a$. By using (9.2.10), these projected currents become

$$
\mathcal{Q}i\mathcal{L}_t \hat{\mathbf{R}} = 0
$$

$$
\mathcal{Q}i\mathcal{L}_t \hat{\mathbf{V}} = \frac{1}{\mu}\mathcal{Q}\hat{\mathbf{F}}^{\text{ends}} = \frac{1}{\mu}\left[\hat{\mathbf{F}}^{\text{ends}} - \overline{\mathbf{F}}(\mathbf{R})\right]
$$

$$
\mathcal{Q}i\mathcal{L}_t \hat{E}_1 = \hat{Q}
$$

$$
\mathcal{Q}i\mathcal{L}_t \hat{E}_2 = -\hat{Q} \tag{10.4.26}
$$

where the fluctuation of the force that the piston exerts on the gas with respect to the average force given by (9.2.20) is $\delta\hat{F}_X = \hat{F}_X - \langle\hat{F}_X\rangle^a$. Note that the external force does not enter into the projected current and, therefore, the dissipative matrix will not depend on the external force explicitly. The whole dependence on the external force is through the change in the microscopic dynamics $\exp\{i\mathcal{L}_t t'\}$, which is obviously affected by the

external forcing.

$$D_{\mu\nu}(a) \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & T_1\gamma & \zeta & -\zeta \\ 0 & \zeta & \kappa & -\kappa \\ 0 & -\zeta & -\kappa & \kappa \end{pmatrix} \tag{10.4.27}$$

Where we introduced the following transport coefficients, through the Green-Kubo correlations of these projected currents as

$$\kappa(a) = \int_0^{\Delta t} dt \left\langle \hat{Q}(t)\hat{Q} \right\rangle^a$$

$$\zeta(a) = \int_0^{\Delta t} dt \left\langle \hat{Q}(t)\delta\hat{\mathbf{F}}^{\mathrm{ends}} \right\rangle^a$$

$$\zeta'(a) = \int_0^{\Delta t} dt \left\langle \delta\hat{\mathbf{F}}^{\mathrm{ends}}(t)\hat{Q} \right\rangle^a$$

$$\gamma(a) = \frac{1}{T_1} \int_0^{\Delta t} dt \left\langle \delta\hat{\mathbf{F}}^{\mathrm{ends}}(t)\delta\hat{\mathbf{F}}^{\mathrm{ends}} \right\rangle^a \tag{10.4.28}$$

The introduction of the prefactor T_1 is just convenient and accords with a usual definition of the friction coefficient in Brownian motion. In principle the transport coefficients (9.2.24) depend on the state $a = (X, P, E_1, E_2)$, as they are defined in terms of conditional expectations.

Note that the dissipative matrix is symmetric, positive semi-definite (given that $\kappa, \gamma > 0$), where the semi comes from the degeneracy $D_{\mu\nu}\frac{\partial E}{\partial a_\nu} = 0$. This degeneracy is easily seen to be satisfyed when we multiply the dissipative matrix (10.4.27) with the gradient of the energy (10.4.5).

Onsager's symmetry (5.3.31) states that **[Check external force sign!]**

$$D_{\mu\nu}(a) = \varepsilon_\mu\varepsilon_\nu D_{\nu\mu}(\varepsilon\cdot a) \tag{10.4.29}$$

or more explicitly

$$\begin{pmatrix} D_{11}(a) & D_{12}(a) & 0 & D_{1P}(a) \\ D_{21}(a) & D_{22}(a) & 0 & D_{2P}(a) \\ 0 & 0 & 0 & 0 \\ D_{P1}(a) & D_{P2}(a) & 0 & D_{PP}(a) \end{pmatrix} = \begin{pmatrix} D_{11}(\varepsilon\cdot a) & D_{21}(\varepsilon\cdot a) & 0 & -D_{P1}(\varepsilon\cdot a) \\ D_{12}(\varepsilon\cdot a) & D_{22}(\varepsilon\cdot a) & 0 & -D_{P2}(\varepsilon\cdot a) \\ 0 & 0 & 0 & 0 \\ -D_{1P}(\varepsilon\cdot a) & -D_{2P}(\varepsilon\cdot a) & 0 & D_{PP}(\varepsilon\cdot a) \end{pmatrix}$$
$$\tag{10.4.30}$$

For example, the identity $D_{1P}(a) = -D_{P1}(\varepsilon \cdot a)$ leads to

$$
\begin{aligned}
D_{1P}(a) &= -\left(\zeta + \gamma T_1 V\right) \\
-D_{P1}(\varepsilon \cdot a) &= \left(\zeta' + \gamma T_1 V\right) \\
\zeta + \zeta' &= 0
\end{aligned}
\tag{10.4.31}
$$

Note that, from the definition through the Green-Kubo formula (9.2.24), the coefficient ζ involves correlations of the energy exchange between systems 1 and 2 and the fluctuation of the forces on the piston. If the system 1 is sufficiently large, these correlations are expected to be negligible. Therefore, we will neglect $\zeta \simeq 0$.

The Onsager's reciprocity

The spurious drift

The noise

The irreversible drift

$$
D_{\mu\nu}(a)\frac{\partial S_B}{\partial a_\nu} \rightarrow
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & T_1\gamma & \zeta & -\zeta \\
0 & \zeta & \kappa & -\kappa \\
0 & -\zeta & -\kappa & \kappa
\end{pmatrix}
\begin{pmatrix}
\frac{\overline{\mathbf{F}}(\mathbf{R})}{T_1} \\
\mu\frac{\mathbf{V}}{T_1} \\
\frac{1}{T_1} \\
\frac{1}{T_2}
\end{pmatrix}
=
\begin{pmatrix}
0 \\
\boldsymbol{\gamma}\cdot\mathbf{V} + \zeta\left[\frac{1}{T_1} - \frac{1}{T_2}\right] \\
\boldsymbol{\zeta}\cdot\mathbf{V} + \kappa\left[\frac{1}{T_1} - \frac{1}{T_2}\right] \\
-\boldsymbol{\zeta}\cdot\mathbf{V} - \kappa\left[\frac{1}{T_1} - \frac{1}{T_2}\right]
\end{pmatrix}
\tag{10.4.32}
$$

The final SDE

10.5   The forced molecule in vacuo

From the above equations we may obtain the dynamics of an externally forced macro-molecule that is not in contact with any other system. This may be achieved by setting to zero the interaction term $\phi^I(\mathbf{q}_{i_1} - \mathbf{q}_{j_2}) = 0$. This imples that the bath and molecule are decoupled. In particular, the transport coefficients $\boldsymbol{\zeta}, \kappa$ vanish, while the friction coefficient $\boldsymbol{\gamma}$ involves only the correlation of the force due to the atoms of the molecule.$<$

11

Fluctuation Theorems for Thermodynamics

In Sec 4.6.3 we have considered the Detailed Balance Theorem that reflects the microscopic time reversibility of Hamilton's equations at the macroscopic level of the CG variables. In that section we considered the general case for arbitrary CG variables. In this section we will consider the form of the Detailed Balance Theorem when we choose the usual CG variables in the Thermodynamic level of description.

11.0.1   Work and Heat

When the Hamiltonian is time dependent, total energy is no longer a conserved quantity. Two new concepts enter then into the picture: Work and Heat. Both can be defined in terms of functions in phase space and the microscopic dynamics of the system. In this subsection, we seek for a definition of work and heat in terms of functions of phase space.

    We consider a system composed of two subsystems A, B. System A is the one we are interested in. System B will act as a **thermal bath**, a concept still to be defined in microscopic terms. The microscopic state is specified by $z = (z_1, z_2)$ where z_1 (z_2) is the collection of positions and momenta of the particles making up A (B). The Hamiltonian $\hat{H}_t(z)$ of the whole system is given by (11.0.1) where

$$\hat{H}_t(z) = \hat{H}^A(z) + \hat{H}^B(z) + \hat{C}(z)\cdot\phi_t \tag{11.0.1}$$

Here $\hat{H}^A(z)$ is the unperturbed Hamiltonian of system A, $\hat{H}^B(z)$ is the Hamiltonian of system B including the interaction potential energy between systems A and B, and $\hat{C}_\mu(z)\cdot\phi_t$ is a time-dependent potential describing the effect of external field. This last term is the only one that depends explicitly on time. For simplicity we assume that the external potential only acts on the degrees of freedom of system A, this is $\hat{C}(z) = \hat{C}(z_1)$. The evolution of the microstate with initial condition z under the Hamiltonian (11.0.1) is given by the flow map $z_t = \hat{\Phi}_{[\phi,0,t]}(z)$.

Let us compute the time derivative of the total energy defined in Eq. (11.0.1)

$$\frac{d}{dt}\hat{H}_t(z_t) = \frac{\partial \hat{H}_t}{\partial z}(z_t)J\frac{\partial \hat{H}_t}{\partial z}(z_t) + \partial_t \hat{H}_t(z_t) = \hat{C}_\mu(z_t)\cdot\dot{\phi}_t \tag{11.0.2}$$

This tells us that the change in total energy of the composite system is due solely to the time-dependence of the external potential. If the external potential is time-independent, then total energy is conserved. We introduce the variation in time of total energy as

$$\Delta \hat{H}_{[\phi,0,t]}(z) = \hat{H}_t(z_t) - \hat{H}_0(z) \tag{11.0.3}$$

which is given by

$$\Delta \hat{H}_{[\phi,0,t]}(z) = \int_0^t dt'\frac{d}{dt'}\hat{H}_{t'}(z_{t'}) = \int_0^t dt'\hat{C}_\mu(z_{t'})\cdot\dot{\phi}_{t'} \tag{11.0.4}$$

Alternatively, this total variation is due to the total variation of each of the contributions, this is

$$\Delta \hat{H}_{[\phi,0,t]}(z) = \Delta \hat{H}^A_{[\phi,0,t]}(z) + \Delta \hat{H}^B_{[\phi,0,t]}(z) + \Delta V_t^{\text{ext}}(z) \tag{11.0.5}$$

where we have introduced the following variations

$$\Delta \hat{H}^{(1)}_{[\phi,0,t]}(z) = \hat{H}^{(1)}(z_t) - \hat{H}^{(1)}(z)$$
$$\Delta \hat{H}^B_{[\phi,0,t]}(z) = \hat{H}^B(z_t) - \hat{H}^B(z)$$
$$\Delta V_t^{\text{ext}}(z) = \hat{C}(z_t)\cdot\phi_t - \hat{C}(z)\cdot\phi_0 \tag{11.0.6}$$

By using (11.0.4) and (11.0.5) we may express the variation of the energy of system A in the form of the **microscopic First Law of Thermodynamics**

$$\Delta \hat{H}^A_{[\phi,0,t]}(z) = \hat{W}_{[\phi,0,t]}(z) + \hat{Q}_{[\phi,0,t]}(z) \tag{11.0.7}$$

where we have introduced the definitions for **work** $\hat{W}_{[\phi,0,t]}(z)$ and **heat** $\hat{Q}_{[\phi,0,t]}(z)$ as

$$\hat{W}_{[\phi,0,t]}(z) = \int_0^t dt'\hat{C}(z_{t'})\cdot\dot{\phi}_{t'} - \hat{C}(z_t)\cdot\phi_t + \hat{C}(z)\cdot\phi_0$$
$$\hat{Q}_{[\phi,0,t]}(z) = -\Delta \hat{H}^B_{[\phi,0,t]}(z) \tag{11.0.8}$$

The work is defined entirely in terms of the external forcing, while the heat involves the variations of the energy of system B and the coupling energy between both systems. By noting that

$$\hat{C}(z_t)\cdot\phi_t - \hat{C}(z)\cdot\phi_t = \int_0^t dt'\frac{d}{dt'}\left(\hat{C}(z_{t'})\cdot\phi_{t'}\right)$$
$$= \int_0^t dt'i\mathcal{L}(\phi_{t'})\hat{C}(z_{t'})\cdot\phi_{t'} + \int_0^t dt'\hat{C}(z_{t'})\cdot\dot{\phi}_{t'} \tag{11.0.9}$$

the work becomes

$$\hat{W}_{[\phi,0,t]}(z) = -\int_0^t dt' i\mathcal{L}(\phi_{t'})\hat{C}(z_{t'})\cdot\phi_{t'} \tag{11.0.10}$$

Note that the work and heat are *functionals* of the protocol ϕ_t as they depend on all the values attained by the protocol. This is why we use the detailed notation. However, when confusion does not arise, we will shorten the notation as

$$\hat{W}_{[\phi,0,t]}(z) = \hat{W}_t(z) \tag{11.0.11}$$

11.0.2   Detailed Balance for an isolated system under external forcing

When our system A does not interact with any other system B, the Hamiltonian is

$$\hat{H}_t(z) = \hat{H}^A(z) + \hat{A}(z) \cdot \phi_t \tag{11.0.12}$$

where $\hat{H}^A(z)$ is the Hamiltonian of the unperturbed system and $\hat{A}(z) \cdot \phi_t$ is the external forcing, with a given protocol ϕ_t. With respect to the Hamiltonian (11.0.1) we see that $\hat{H}^B(z) = 0$ and there is no heat in the problem $\hat{Q}_{[\phi,0,t]} = 0$ as seen in (11.0.8). In this case, the level of description of Thermodynamics is defined by the selection of the energy $\hat{H}^A(z)$ as the CG variable. In more specific examples later on we will discuss under what conditions it is appropriate to choose this sole variable for the CG description of the system.

Note that the Hamiltonian is even under time reversal $\varepsilon = 1$, because $\hat{H}^A(\epsilon \cdot z) = \hat{H}^A(z)$. We make no assumptions on the parity of the coupling functions $\hat{C}(z)$ that will transform as $\hat{C}(\epsilon \cdot z) = \varepsilon \cdot \hat{C}(z)$.

The joint probability distributions $P_{[\phi]}(E_0, t_0, E_1, t_1)$ that the unperturbed Hamiltonian $\hat{H}^A(z)$ of the system takes the value $\hat{H}^A(z) = E_0$ at time t_0 and the value $\hat{H}^A(z) = E_1$ at time t_1 is, by definition

$$P_{[\phi]}(E_0, t_0, E_1, t_1) = \int dz \rho_0(z) \delta(\hat{H}^A(z) - E_0) \delta(\hat{H}^A(\hat{z}^F(z)) - E_1) \tag{11.0.13}$$

where $\hat{z}^F(z) = \hat{\Phi}_{[\phi, t_0, t_1]}(z)$. Note that the Detailed Balance Theorem assumes from a start that the initial ensemble is given by an ensemble of the relevant form. In the present case, this relevant ensemble is

$$\rho_0(z) = \frac{P_0(\hat{H}^A(z))}{\Omega(\hat{H}^A(z))} \tag{11.0.14}$$

where $P_0(E)$ is the arbitrary distribution of the energy at the initial time. With (11.0.14), the joint probability becomes

$$P_{[\phi]}(E_0, t_0, E_1, t_1) = \frac{P_0(E_0)}{\Omega(E_0)} \int dz \rho_0^N \delta(\hat{H}^A(z) - E_0) \delta(\hat{H}^A(\hat{z}^F(z)) - E_1) \tag{11.0.15}$$

The conditional probability of having E_1 at time t_1 given that the energy was E_0 at t_0 is

$$P_{[\phi]}(E_0, t_0 | E_1 t_1) = \frac{P(E_0, t_0, E_1, t_1)}{P_0(E_0)}$$

$$= \frac{1}{\Omega(E_0)} \int dz \rho_0^N \delta(\hat{H}^A(z) - E_0) \delta(\hat{H}^A(\hat{z}^F(z)) - E_1) \tag{11.0.16}$$

It is precisely for this kind of conditional probability that we have proved the Detailed Balance Theorem (4.6.78). Therefore, the Theorem of Detailed Balance for the level of

Thermodynamics is

$$\frac{P_{[\phi]}(E_0, t_0 | E_1, t_1)}{P_{[\varepsilon \cdot \phi^R]}(E_1, t_0 | E_0, t_1)} = e^{S_B(E_1) - S_B(E_0)} \tag{11.0.17}$$

where the bare entropy is the one introduced originally by Boltzmann, this is

$$S_B(E) = \ln \int dz \rho_0^N \delta(\hat{H}^A(z) - E) \tag{11.0.18}$$

As we have seen in the previous chapter, the entropy scales with the number of degrees of freedom of the system. This implies that for large systems the probability of performing a jump from an energy E_0 of a low entropy $S_B(E_0)$ to an energy E_1 with a higher entropy $S_B(E_1) > S_B(E_0)$ under a protocol ϕ_t is exponentially more probable than the reverse jump under the protocol $\varepsilon \cdot \phi_t^R$.

Note that we may express the transition probabilities in terms of the work, according to the microscopic expression (11.0.7) of the First Law that we write as

$$\hat{H}^A\left(\hat{z}^F(z)\right) = \hat{H}^A(z) + \hat{W}^F(z) \tag{11.0.19}$$

this is

$$
\begin{aligned}
P_{[\phi]}(E_0, t_0 | E_1 t_1) &= \frac{1}{\Omega(E_0)} \int dz \rho_0^N \delta(\hat{H}^A(z) - E_0) \delta(\hat{H}^A(\hat{z}^R(z) - E_1) \\
&= \frac{1}{\Omega(E_0)} \int dz \rho_0^N \delta(\hat{H}^A(z) - E_0) \delta(\hat{W}^F(z) + \hat{H}^A(z) - E_1) \\
&= \frac{1}{\Omega(E_0)} \int dz \rho_0^N \delta(\hat{H}^A(z) - E_0) \delta(\hat{W}^F(z) - (E_1 - E_0)) \\
&= \int dz \rho_0^N \rho_0^{\text{mic}}(z) \delta(\hat{W}^F(z) - (E_1 - E_0)) \\
&\equiv P_{E_0}(\hat{W}^F = \Delta E)
\end{aligned}
\tag{11.0.20}
$$

where the last equality defines the probability that the work takes a particular value, where the probability is computed with the microcanonical ensemble

$$\rho_0^{\text{mic}}(z) = \frac{1}{\Omega(E_0)} \delta(\hat{H}^A(z) - E_0) \tag{11.0.21}$$

Observe that a *transition* probability of the energy is expressed as *probability*. In a similar way, we may express the reverse conditional probability as

$$P_{[\varepsilon \cdot \phi^R]}(E_1, t_0 | E_0, t_1) = P_{E_1}(\hat{W}^R(z) = -\Delta E) \tag{11.0.22}$$

By using (11.0.20), (11.0.22) the Detailed Balance Theorem (4.6.78) now reads

$$\frac{P_{E_0}(\hat{W}_{[\phi, t_0, t_1]} = \Delta E)}{P_{E_1}(\hat{W}_{[\varepsilon \cdot \phi^R, t_0, t_1]} = -\Delta E)} = \exp\{S_B(E_1) - S_B(E_0)\} \tag{11.0.23}$$

which is reminiscent of Crook's theorem, although in the present case it allows to estimate directly the entropy function.

Run two MD simulations with no external forcing at two different energies E_0, E_1 to generate two swarms of initial conditions. Start from the microstates of E_0 a simulation with a prescribed protocol. Measure at time t the distribution of energy $P^R(E)$ under this protocol. Now, start from the microstates E_1 forced with the reverse protocol and measure again the distribution of energy $P^B(E)$. The Detailed Balance Theorem states that $S_B(E + E_0) - S(E_0) = \ln P^R(E)/P^B(-E)$. This is valid for any protocol!

For large systems, the exponential is extremely large (small) if $\Delta E > 0$ ($\Delta E < 0$). This implies that it is exponentially more probable that the work done during the protocol ϕ_t increases the energy than the work done during the protocol ϕ_t^R decreases the energy. For example, if we design a cyclic protocol in which in the first part we force the system with the protocol ϕ_t and in the second part with the reversed protocol ϕ_t^R, it is much more probable that the system gains energy than rather looses it. These protocols cannot "cool" the system.

11.0.3   Detailed Balance for energy of the system and heat bath

We now consider the case Detailed Balance Theorem when the system A is interacting with a system B. The Hamiltonian is then given by (11.0.1). For the present level of description we choose the energy of each system $\hat{H}^A(z), \hat{H}^B(z)$ as the CG variables in the description. These functions take numerical values E^A, E^B. The bare entropy of this level of description is defined as

$$S_B(E^A, E^B) = \ln \Omega(E^A, E^B) \tag{11.0.24}$$

where the measure is defined as

$$\Omega(E^A, E^B) = \int dz \delta(\hat{H}^A(z) - E^A)\delta(\hat{H}^B(z) - E^B) \tag{11.0.25}$$

We may now consider the Detailed Balance Theorem (4.6.78). First, note that the transition probability entering the theorem is, in the present level of description

$$P_{[\phi]}(E_0^A, E_0^B, t_0 | E_1^A, E_1^B, t_1)$$
$$= \frac{1}{\Omega(E_0^A, E_0^B)} \int dz \delta(\hat{H}^A(z) - E_0^A)\delta(\hat{H}^A(z_t) - E_1^A)\delta(\hat{H}^B(z) - E_0^B)\delta(\hat{H}^B(z_t) - E_1^B) \tag{11.0.26}$$

We may express the transition probability in terms of the work and heat, by using the First Law (11.0.7) and the definitions (11.0.8) in the form

$$\hat{H}^A(z_t) = \hat{H}^A(z) + \hat{W}_{[\phi,0,t]}(z) + \hat{Q}_{[\phi,0,t]}(z)$$
$$\hat{H}^B(z_t) = \hat{H}^B(z) - \hat{Q}_{[\phi,0,t]}(z) - \Delta\hat{H}^I_{[\phi,0,t]}(z) \tag{11.0.27}$$

where $\Delta \hat{H}^I_{[\phi,0,t]}(z) = \hat{H}^I(z_t) - \hat{H}^I(z)$ is the variation in the coupling interaction. By substitution of (11.0.27) into (11.0.26) we obtain

$$P_{[\phi]}(E_0^A, E_0^B, t_0 | E_1^A, E_1^B, t_1)$$

$$= \frac{1}{\Omega(E_0^A, E_0^B)} \int dz$$

$$\times \delta(\hat{H}^A(z) - E_0^A) \delta\left(\hat{H}^A(z) + \hat{W}_{[\phi,0,t]}(z) + \hat{Q}_{[\phi,0,t]}(z) - E_1^A\right)$$

$$\times \delta(\hat{H}^B(z) - E_0^B) \delta\left(\hat{H}^B(z) - \hat{Q}_{[\phi,0,t]}(z) - \Delta \hat{H}^I_{[\phi,0,t]}(z) - E_1^B\right)$$

$$= \frac{1}{\Omega(E_0^A, E_0^B)} \int dz$$

$$\times \delta(\hat{H}^A(z) - E_0^A) \delta\left(\hat{W}_{[\phi,0,t]}(z) + \hat{Q}_{[\phi,0,t]}(z) - (E_1^A - E_0^A)\right)$$

$$\times \delta(\hat{H}^B(z) - E_0^B) \delta\left(-\hat{Q}_{[\phi,0,t]}(z) - \Delta \hat{H}^I_{[\phi,0,t]}(z) - (E_1^B - E_0^B)\right)$$

$$= \frac{1}{\Omega(E_0^A, E_0^B)} \int dz$$

$$\times \delta(\hat{H}^A(z) - E_0^A) \delta\left(\hat{W}_{[\phi,0,t]}(z) - \Delta \hat{H}^I_{[\phi,0,t]}(z) - \Delta E^B - \Delta E^A\right)$$

$$\times \delta(\hat{H}^B(z) - E_0^B) \delta\left(\hat{Q}_{[\phi,0,t]}(z) + \Delta \hat{H}^I_{[\phi,0,t]}(z) + \Delta E^B\right) \qquad (11.0.28)$$

For the sake of the argument, we will assume that $\Delta \hat{H}^I_{[\phi,0,t]}(z)$ is negligible in front of the other terms in these expressions

$$P_{[\phi]}(E_0^A, E_0^B, t_0 | E_1^A, E_1^B, t_1) = \int dz \rho_0^{\mathrm{mic}}(z) \delta\left(\hat{W}_{[\phi,0,t]}(z) - \Delta E^{AB}\right) \delta\left(\hat{Q}_{[\phi,0,t]}(z) + \Delta E^B\right)$$

$$(11.0.29)$$

where we have introduced the microcanonical ensemble

$$\rho_0^{\mathrm{mic}}(z) = \frac{1}{\Omega(E_0^A, E_0^B)} \delta(\hat{H}^A(z) - E_0^A) \delta(\hat{H}^B(z) - E_0^B) \qquad (11.0.30)$$

Observe that the *transition* probability of the energy is expressed as the *probability* that the work and heat take a particular value, where the probability is computed with the microcanonical ensemble (11.0.30).

By using (11.0.29), the Detailed Balance Theorem (4.6.78) now reads

$$\frac{P^{\mathrm{mic}}(\hat{W} = \Delta E^{AB}, \hat{Q} = -\Delta E^B)}{P^{\mathrm{mic}}(\hat{W} = -\Delta E^{AB}, \hat{Q} = +\Delta E^B)} = \exp\{S_B(E_1^A, E_1^B) - S_B(E_0^A, E_0^B)\} \qquad (11.0.31)$$

How this theorem is related to the Fluctuation Theorem

A crucial assumption made in the Detailed Balance theorem is that the initial ensemble is of the relevant type (4.6.69). This gives a dissipation function given by

$$\hat{\Sigma}_{[\phi,0,t]} = \ln \frac{\rho_0(z)}{\rho_0(z_t)} = \ln \frac{P_0(\hat{H}_0(z))}{P_0(\hat{H}_0(z_t))} + \ln \frac{\Omega_0(\hat{H}_0(z_t))}{\Omega(\hat{H}_0(z))} \qquad (11.0.32)$$

If the initial distribution of energy corresponds to that of a thermalized system

$$P_0(E) = \frac{1}{Z(\beta_0)}\Omega(E)\exp\{-\beta_0 E\} \qquad (11.0.33)$$

and the dissipation function takes the simple form

$$\hat{\Sigma}_{[\phi,0,t]}(z) = \beta_0(H_0(z_t) - H_0(z)) = \beta_0 \hat{W}_t(z) \qquad (11.0.34)$$

[Faltan factores $\Omega(\hat{H}_0(z_t))$ y $Z(\beta_0)$, completar...]

11.0.4   Crooks Theorem

Assume that the Hamiltonian is even under time reversal (the couplings also). Consider the variation of total energy of the system

$$\Delta\hat{H}_{[\phi,t_0,t_1]}(z) = \hat{H}(\hat{\Phi}_{[\phi,t_0,t_1]}(z),\phi_{t_1}) - \hat{H}(z,\phi_{t_0}) \qquad (11.0.35)$$

The probability that $\Delta\hat{H}$ takes a value W sampled from an initial canonical ensemble is

$$
\begin{aligned}
P_{[\phi]}(W) &= \int dz \rho_0^N \frac{\exp\{-\beta\hat{H}(z,t_0)\}}{Z(\phi_{t_0})} \delta\left(\Delta\hat{H}_{[\phi,t_0,t_1]}(z) - W\right) \\
&= \int dz \rho_0^N \frac{\exp\{-\beta\hat{H}(z,t_0)\}}{Z(\phi_{t_0})} \delta\left(\hat{H}(\hat{\Phi}_{[\phi,t_0,t_1]}(z),\phi_{t_1}) - \hat{H}(z,\phi_{t_0}) - W\right) \\
&= e^W \int dz \rho_0^N \frac{\exp\{-\beta\hat{H}(\hat{\Phi}_{[\phi,t_0,t_1]}(z),\phi_{t_1})\}}{Z(\phi_{t_0})} \delta\left(\hat{H}(\hat{\Phi}_{[\phi,t_0,t_1]}(z),\phi_{t_1}) - \hat{H}(z,\phi_{t_0}) - W\right) \\
&= e^W \frac{Z(\phi_{t_1})}{Z(\phi_{t_0})} \int dz \rho_0^N \frac{\exp\{-\beta\hat{H}(z',\phi_{t_1})\}}{Z(\phi_{t_1})} \delta\left(\hat{H}(z',\phi_{t_1}) - \hat{H}(\hat{\Phi}_{[\phi,t_0,t_1]}^{-1}(z'),\phi_{t_0}) - W\right)
\end{aligned}
$$
$$(11.0.36)$$

Now use microscopic reversibility

$$\hat{\Phi}_{[\phi,t_0,t_1]}^{-1}(z') = \epsilon \cdot \hat{\Phi}_{[\phi^R,t_0,t_1]}(\epsilon \cdot z') \qquad (11.0.37)$$

and the fact that $\phi_{t_1}^R = \phi_{t_0}$, $\phi_{t_0}^R = \phi_{t_1}$

$$P_{[\phi]}(W) = e^{\beta W} \frac{Z(\phi_{t_1})}{Z(\phi_{t_0})} \int dz' \rho_0^N \frac{\exp\{-\beta \hat{H}(z', \phi_{t_0}^R)\}}{Z(\phi_{t_0}^R)} \delta \left(\hat{H}(z', \phi_{t_1}) - \hat{H}(\epsilon \cdot \hat{\Phi}_{[\phi^R, t_0, t_1]}(\epsilon \cdot z'), \phi_{t_1}^R) - W \right)$$

$$= e^{\beta W} \frac{Z(\phi_{t_1})}{Z(\phi_{t_0})} \int dz \rho_0^N \frac{\exp\{-\beta \hat{H}(z, \phi_{t_0}^R)\}}{Z(\phi_{t_0}^R)} \delta \left(\hat{H}(z, \phi_{t_0}^R) - \hat{H}(\hat{\Phi}_{[\phi^R, t_0, t_1]}(z'), \phi_{t_1}^R) - W \right)$$

$$= e^{\beta W} \frac{Z(\phi_{t_1})}{Z(\phi_{t_0})} P_{[\phi^R]}(-W) \tag{11.0.38}$$

We arrive, therefore, to Crook's theorem

$$\frac{P_{[\phi]}(W)}{P_{[\phi^R]}(W)} = \exp\{\beta(W - F(\phi_t) + F(\phi_0))\} \tag{11.0.39}$$

where the dimensional free energy is $F(\phi_t) = -k_B T \ln Z(\phi_t)$.

11.0.5 Jarzynski identity for an isolated system under external forcing

The CG variable of the system is the Hamiltonian *at the initial time*, this is $\hat{A}(z) = \hat{H}_{t_0}(z)$. Consider an initial ensemble given by a canonical distribution

$$\overline{\rho}_{t_0}(z) = \frac{1}{Z(\beta_0)} e^{-\beta_0 \hat{H}_{t_0}(z)} \tag{11.0.40}$$

We now consider the average over initial conditions of the exponential function of the total work, this is

$$\left\langle e^{-\beta_0 \hat{W}} \right\rangle_0 = \int dz \overline{\rho}_{t_0}(z) e^{-\beta_0 \left(\hat{H}_{t_1}(\hat{\Phi}_{[\phi, t_0, t_1]}(z)) - \hat{H}_{t_0}(z) \right)} \tag{11.0.41}$$

Substitution of (11.0.40) in this expression readily gives **the Jarzynski identity**

$$\left\langle e^{-\beta_0 W} \right\rangle_0 = \frac{\int dz e^{-\beta_0 \hat{H}_{t_1}(z)}}{\int dz e^{-\beta_0 \hat{H}_{t_0}(z)}} = e^{\Phi_{t_1}(\beta_0) - \Phi_{t_0}(\beta_0)} \tag{11.0.42}$$

where we have introduced the thermodynamic potential (2.7.10) corresponding to the level of description of thermodynamics, this is

$$\Phi_t(\beta_0) \equiv -\ln \int dz e^{-\beta_0 \hat{H}_t(z)} \tag{11.0.43}$$

The dimensionless thermodynamic potential $\Phi_t(\beta)$ of the level of description of thermodynamics is related to the standard free energy as

$$\Phi_t(\beta) = \beta_0 F(\beta) \tag{11.0.44}$$

[The interest of Jarzynski identity]

We may compare this result with the result that we have obtained in (11.0.46).

In this case the "generalized work" is

$$-\lambda_t \hat{H}^A(z_t) + \lambda_0 \hat{A} \to -\beta(t)\hat{H}^A(z_t) + \beta(0)\hat{H}^A \qquad (11.0.45)$$

$$\left\langle \exp\{-\lambda_t \hat{A}(t) + \lambda_0 \hat{A}\} \right\rangle_0 = \exp\{-(\Phi(a_t) - \Phi(a_0))\} \qquad (11.0.46)$$

11.0.6   Transient Fluctuation theorem for a forced isolated system

12

Smoluchowski Equation

13

The Langevin Thermostate and Brownian motion

13.1 Introduction

Brownian motion is perhaps the paradigmatic problem for coarse-graining and has been at the origin of all the advancements in the theory, since Einstein and then Mori.

We may also understand the problem in slightly more general terms. Imagine that we have a system of interacting particles that are interacting with a thermal bath. The coarse-graining of the thermal bath degrees of freedom will lead to thermostated microscopic equations.

A thermal bath is a system that, when coupled to a given system of interest, allows one to control the temperature of such a system. Experimentally, a thermal bath is usually a large system with a moving liquid. A thermal bath may be used to function as a thermostate.

The two main characteristics of a thermal bath are 1) an extremely large heat capacity and 2) the time scales of the processes taking place within the thermal bath are much faster than those of the system of interest. A thermal bath made of a solid, for example, would have a very large heat capacity and a very high thermal conductivity, in such a way that any energy transfer from or to the system of interest and the bath does not alter neither the local (due to high thermal conductivity) nor the global (due to the large heat capacity) temperature of the bath. One way of having a large heat capacity is simply having a very large system, because the heat capacity is an extensive quantity.

In this work, we describe the dynamics of the reference system coupled with a thermal bath through the theory of coarse graining. In fact, the two defining treats of a thermal bath imply that the only "slow" variable to be kept regarding the bath variables is its total energy.

[Provocative claim: I don't like NEMD, and the use of artificial thermostates of the kind of Nosé-Hoover]

13.2   The theory of coarse-graining

In this section, we review the theory of coarse-graining as was put forward by Zwanzig [?, ?]. At a coarse-grained level of description, the system is described by a set of functions $X_k(z)$ which depend on the set of position and momenta z of the molecules of the system. The selection of the CG variables $X_k(z)$ is a crucial step in the description of a non-equilibrium system. One basic requirement is that they are *slow* variables[?]. When this is the case, Green [?] and later Zwanzig [?] derived from the microscopic dynamics of the system a general Fokker-Planck equation for the set of CG variables x

$$\partial_t P(x,t) = -\frac{\partial}{\partial x} \cdot \left[A(x) + M(x) \cdot \frac{\partial S}{\partial x} \right] P(x,t)$$
$$+ e k_B \frac{\partial}{\partial x} \cdot M(x) \cdot \frac{\partial}{\partial x} P(x,t) \tag{13.2.1}$$

The different objects in this equation have a well-defined microscopic definition. For example, the reversible drift is

$$A(x) = \langle LX \rangle^x \tag{13.2.2}$$

where L is the Liouville operator and the conditional average is defined by

$$\langle \ldots \rangle^x = \frac{1}{\Omega(x)} \int dz \rho^{\text{eq}}(z) \delta(\hat{X}(z) - x) \cdots \tag{13.2.3}$$

where $\rho^{\text{eq}}(z)$ stands for the equilibrium measure and $\delta(\hat{X}(z) - x)$ is actually a product of Dirac delta functions, one for every function $X_k(z)$. The "volume" of phase space compatible with a prescribed value x of the CG variables is

$$\Omega(x) = \int dz \rho^{\text{eq}}(z) \delta(\hat{X}(z) - x) \tag{13.2.4}$$

and is closely related to the entropy of the mesolevel x which is defined through

$$S(x) \equiv k_B \ln \Omega(x) \tag{13.2.5}$$

where k_B is Boltzmann's constant.

Finally, the irreversible symmetric operator M is the matrix of transport coefficients expressed in the form of Green-Kubo formulas,

$$M(x) = \frac{1}{k_B} \int_0^\infty \langle \delta LX \exp\{iQLt'\} \delta LX \rangle^x dt' \tag{13.2.6}$$

where $\delta LX = LX - \langle LX \rangle^x$ is the so called projected current. The dynamic operator $\exp\{iQLt'\}$ is usually named the projected dynamics, which is, strictly speaking different from the real dynamics $\exp\{Lt'\}$. The projection operator Q is defined from its action on any phase function $\hat{F}(z)$ [?]

$$Q\hat{F}(z) = \hat{F}(z) - \langle F \rangle^{\hat{X}(z)} \tag{13.2.7}$$

The projected dynamics can be usually approximated by the real dynamics, but then the upper infinite limit of integration in Eq. (17.5.1) has to be replaced by τ, a time which is long in front of the correlation time of the integrand, but short in front of the time scale of evolution of the macroscopic variables. This is the well-known plateau problem [?],[?],[?]. Very recently, we have shown in a systematic way that a better strategy is to approximate the real dynamics with a constrained dynamics that does not suffer from the plateau problem [?]. The operator M is symmetric and positive definite [?].

The Ito stochastic differential equation that is mathematically equivalent to the Fokker-Planck equation is given by

$$dx = \left[A(x) + M(x)\cdot\frac{\partial S}{\partial x} + k_B\frac{\partial}{\partial x}\cdot M(x) \right] dt + d\tilde{x} \tag{13.2.8}$$

where $d\tilde{x}$ is a linear combination of independent increments of the Wiener process. Their covariance is given by the Fluctuation-Dissipation theorem

$$d\tilde{x}d\tilde{x} \quad = \quad 2k_B M(x)dt \tag{13.2.9}$$

In situations where the thermal fluctuations can be neglected (formally in the limit $k_B \to 0$ [?]), one obtains a set of deterministic equations of the form

$$\dot{x} = A(x) + M(x)\cdot\frac{\partial S}{\partial x} \tag{13.2.10}$$

In summary, the three basic objects to compute in Zwanzig theory are the entropy $S(x)$, the reversible drift $A(x)$ and the friction matrix $M(x)$. When the energy $\hat{H}(z)$ of the system may be written in terms of the CG variables, i.e. $\hat{H}(z) = E(\hat{X}(z))$, then the drift term $A(x)$ may be rewritten in the GENERIC form [?],[?].

13.3 CG variables for the dynamics coupled to a bath

At a microscopic level the system composed of the reference system A and the bath system B is supposed to be isolated and described by classical mechanics. The Hamiltonian of the system $A + B$ is given by

$$\hat{H}(z, z') = \hat{H}_A(z) + \hat{H}_B(z') + \hat{H}_I(z, z') \tag{13.3.1}$$

where $z = \{\mathbf{r}_i, \mathbf{p}_i, i = 1, \cdots N\}$ are the collection of positions and momenta of the system A z' are the corresponding ones of system B, $\hat{H}_A(z), \hat{H}_B(z')$ are the Hamiltonians of the respective isolated systems and $\hat{H}_I(z, z')$ is the interaction energy between the two systems. Because the total system is isolated, the equilibrium ensemble of the whole system is given by the microcanonical ensemble

$$\rho^{\text{eq}}(z, z') = \frac{1}{\Omega_0(E_0)}\delta(\hat{H}(z, z') - E_0) \tag{13.3.2}$$

where E_0 is the value of the total energy and $\Omega_0(E_0)$ is the usual normalization.

As CG variables $\hat{X}(z)$ we choose the following functions of phase space: $X(z, z') =$

$\{\hat{\mathbf{r}}_i, \hat{\mathbf{p}}_i, i = 1, \cdots N, \mathscr{E}(z, z')\}$, where

$$\mathscr{E}(z, z') = \hat{H}_B(z') + \hat{H}_I(z, z') \tag{13.3.3}$$

is the part of the total energy that depends on bath variables. These phase functions take numerical values $x = \{r, p, E\}$, where r, p denotes the full microstate of system A

Let us consider first the equilibrium probability $\Omega(r, p, E)$ at the CG level, which is given by

$$\Omega(r, p, E) = \int dz dz' \rho^{\text{eq}}(z, z') \prod_i^N \delta(\hat{\mathbf{q}}_i - \mathbf{q}_i) \delta(\hat{\mathbf{p}}_i - \mathbf{p}_i) \delta(\hat{H}_B(z') + \hat{H}_I(z, z') - E)$$

$$= \frac{1}{\Omega_0(E_0)} \delta(\hat{H}_A(r, p) + E - E_0) \int dz' \delta(\hat{H}_B(z') + \hat{H}_I(r, p, z') - E) \tag{13.3.4}$$

We expect that for the immense majority of microstates z', the value of the energy of the bath will be much larger than the energy of interaction. In this way, we will take the approximation

$$\int dz' \delta(\hat{H}_B(z') + \hat{H}_I(r, p, z') - E) \approx \int dz' \delta(\hat{H}_B(z') - E)$$

$$\equiv \Omega_B(E) \equiv \exp\{S_B(E)/k_B\} \tag{13.3.5}$$

where we have introduced the usual definition of the Boltzmann entropy $S_B(E)$ for the bath with k_B the Boltzmann constant. The resulting CG equilibrium distribution $\Omega(x)$ defined in Eq. (13.2.4) is now

$$\Omega(r, p, E) = \delta(\hat{H}_A(r, p) + E - E_0) \frac{\Omega_B(E)}{\Omega_0(E_0)} \tag{13.3.6}$$

where now $\Omega_0(E_0)$ is the new normalization factor, given by

$$\Omega_0(E_0) \approx \int dr dp dE \delta(\hat{H}_A(r, p) + E - E_0) \exp\{S_B(E)/k_B\} \tag{13.3.7}$$

It is interesting to compute the marginal $\Omega(r, p)$ which is given by

$$\Omega(r, p) = \int dE \Omega(r, p, E)$$

$$= \frac{\Omega_B(E_0 - \hat{H}_A(r, p))}{\Omega_0(E_0)}$$

$$= \frac{\exp\{k_B^{-1} S_B(E_0 - \hat{H}_A(r, p))\}}{\Omega_0(E_0)} \tag{13.3.8}$$

As we expect that, for the typical values r, p encountered, the value of $\hat{H}_A(r, p) \ll E_0$, we may Taylor expand to first order the argument of the entropy and obtain the well

known canonical ensemble result

$$\Omega(r, p) \approx \frac{\exp\{-\beta_B(E_0)\hat{H}_A(r, p)\}}{Z(E_0)} \tag{13.3.9}$$

where

$$\beta_B(E_0) \equiv \frac{1}{k_B}\frac{\partial S_B}{\partial E}(E_0)$$

$$Z(E_0) = \int dr dp \exp\{-\beta_B(E_0)\hat{H}_A(r, p)\} \tag{13.3.10}$$

We need to define the entropy function (13.2.5) and we should be careful in dealing with the logarithm of a Dirac delta function. Such an object does not exists, and we have to regularize the Dirac delta function appropriately. One possibility is to approximate $\delta(x)$ with a regular function

$$\overline{\delta}(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left\{-\frac{x^2}{2\sigma^2}\right\} \tag{13.3.11}$$

where the approximation will be good for $\sigma \to 0$. In this way, the entropy $S(r, p, E)$ of the CG level is given by

$$S(r, p, E) = k_B \ln\left[\overline{\delta}(\hat{H}_A(r, p) + E - E_0)\right] + S_B(E) - S_0(E_0) \tag{13.3.12}$$

The derivatives of the entropy are

$$\frac{\partial S}{\partial \mathbf{q}_i}(r, p, E) = G(r, p, E)\mathbf{F}_i^A(r)$$

$$\frac{\partial S}{\partial \mathbf{p}_i}(r, p, E) = -G(r, p, E)\frac{\mathbf{p}_i}{m_i}$$

$$\frac{\partial S}{\partial E}(r, p, E) = \frac{\partial S_B}{\partial E}(E) - G(r, p, E) \tag{13.3.13}$$

where we have introduced the function

$$G(r, p, E) \equiv -k_B\frac{\overline{\delta}'(\hat{H}_A(r, p) + E - E_0)}{\overline{\delta}(\hat{H}_A(r, p) + E - E_0)} \tag{13.3.14}$$

where the prime denotes the derivative.

Let us move on to the conditional averages (13.2.3) for an arbitrary function $F(z, z')$ of the full phase space of the composite system. It will take the form, after integrating some Dirac delta functions

$$\langle F \rangle^{rpE} = \frac{1}{\Omega(r, p, E)}\int dz dz' \rho^{\mathrm{eq}}(z, z')\delta(\hat{r} - r)\delta(\hat{p} - p)\delta(\hat{H}_B(z') + \hat{H}_I(z, z') - E)F(z, z')$$

$$\approx \frac{1}{\Omega_B(E)}\int dz' \delta(\hat{H}_B(z') - E)F(r, p, z') \tag{13.3.15}$$

where, again, we have assumed that the interaction energy may be neglected in the calculation of the averages over z'. Note that if the function over which the conditional average is take depends only on z, the conditional average does nothing, this is

$$\langle F\rangle^{rpE} = F(r,p) \tag{13.3.16}$$

Next, we need the time derivatives LX of the CG variables that in the present case are given by

$$
\begin{aligned}
L\mathbf{q}_i &= \frac{\mathbf{p}_i}{m_i}\\
L\mathbf{p}_i &= \mathbf{F}_i^A(r) + \mathbf{F}_i^I(r,r')\\
L\mathscr{E} &= L\hat{H}_B(z') + L\hat{H}_I(z,z')
\end{aligned}
\tag{13.3.17}
$$

The total energy can be expressed in terms of the CG variables, because

$$\hat{H}(z,z') = \sum_i^N \frac{\mathbf{p}_i^2}{2m} + V^A(r) + \mathscr{E}(z,z') = E(r,p,\mathscr{E}) \tag{13.3.18}$$

Because the total energy is conserved $L\hat{H}(z,z') = 0$ which takes the form

$$\frac{\partial E}{\partial \mathbf{q}_i}(r,p,\mathscr{E})\mathbf{v}_i + \frac{\partial E}{\partial \mathbf{p}_i}(r,p,\mathscr{E})\mathbf{F}_i + L\mathscr{E} = 0 \tag{13.3.19}$$

which becomes simply

$$L\mathscr{E} = -\mathbf{F}_i^I\mathbf{v}_i \tag{13.3.20}$$

The reversible drift $A(x)$ (13.2.2) now takes the form

$$
\begin{aligned}
\langle L\mathbf{q}_i\rangle^{rpE} &= \frac{\mathbf{p}_i}{m_i}\\
\langle L\mathbf{p}_i\rangle^{rpE} &= \mathbf{F}_i^A(r) + \langle \mathbf{F}_i^I\rangle^{rpE}\\
\langle L\mathscr{E}\rangle^{rpE} &= -\langle \mathbf{F}_i^I\rangle^{rpE}\mathbf{v}_i
\end{aligned}
\tag{13.3.21}
$$

The dissipative matrix $M(x)$ will have the structure

$$M(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & M_{ij} & M_{iE} \\ 0 & M_{Ei} & M_{EE} \end{pmatrix} \tag{13.3.22}$$

where the different transport coefficients are given by the Green-Kubo expressions

$$M_{ij}(r,p,E) = \frac{1}{k_B} \int_0^\infty dt' \left\langle \delta\mathbf{F}_i^I \delta\mathbf{F}_j^I(t') \right\rangle^{rpE}$$

$$M_{iE}(r,p,E) = \frac{1}{k_B} \int_0^\infty dt' \left\langle \delta\mathbf{F}_i^I \delta L\mathscr{E}(t') \right\rangle^{rpE}$$

$$= \frac{1}{k_B} \int_0^\infty dt' \left\langle \delta\mathbf{F}_i^I \delta\mathbf{F}_j^I(t') \right\rangle^{rpE} \mathbf{v}_j = M_{ij}\mathbf{v}_j$$

$$M_{Ei}(r,p,E) = -\mathbf{v}_i M_{ij}$$

$$M_{EE}(r,p,E) = \mathbf{v}_i M_{ij}\mathbf{v}_j \tag{13.3.23}$$

The irreversible term will be

$$M(x)\frac{\partial S}{\partial x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & M_{ij} & -\mathbf{v}_i M_{ij} \\ 0 & -\mathbf{v}_i M_{ij} & \mathbf{v}_i M_{ij}\mathbf{v}_j \end{pmatrix} \begin{pmatrix} G\mathbf{F}_j^A \\ -G\frac{\mathbf{p}_j}{m_j} \\ \frac{\partial S_B}{\partial E} - G \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{\partial S_B}{\partial E}M_{ij}\mathbf{v}_j \\ \frac{\partial S_B}{\partial E}M_{ij}\mathbf{v}_i\mathbf{v}_j \end{pmatrix}$$

$$\tag{13.3.24}$$

The final SDE (13.2.8) governing the dynamics of the CG variables are simply (we forget for the time being the term $k_B \nabla M$)

$$d\mathbf{q}_i = \mathbf{v}_i dt$$

$$d\mathbf{p}_i = \mathbf{F}_i^A(r)dt + \langle \mathbf{F}_i^I \rangle^{rpE} - M_{ik}\mathbf{v}_k \frac{\partial S_B}{\partial E}dt + d\tilde{\mathbf{p}}_i$$

$$dE = -\langle \mathbf{F}_i^I \rangle^{rpE}\mathbf{v}_i + M_{ik}\mathbf{v}_k\mathbf{v}_i \frac{\partial S_B}{\partial E}dt + d\tilde{E} \tag{13.3.25}$$

We have to include the term $k_B \nabla M$ which is of the form

$$k_B \nabla M(x) = \left(\frac{\partial}{\partial\mathbf{q}_j}, \frac{\partial}{\partial\mathbf{p}_j}, \frac{\partial}{\partial E}\right) \begin{pmatrix} 0 & 0 & 0 \\ 0 & M_{ij} & M_{iE} \\ 0 & M_{Ei} & M_{EE} \end{pmatrix} = k_B \begin{pmatrix} 0 \\ \frac{\partial M_{ij}}{\partial\mathbf{p}_j} - \frac{\partial M_{ij}}{\partial E}\mathbf{v}_j \\ \sum_j M_{ij}\frac{1}{m_j} - \frac{\partial M_{ij}}{\partial\mathbf{p}_j}\mathbf{v}_i + \mathbf{v}_i \frac{\partial M_{ij}}{\partial E}\mathbf{v}_j \end{pmatrix}$$

$$\tag{13.3.26}$$

In the exponential approximation

$$M_{ij} \approx \tau \langle \delta\mathbf{F}_i^I \delta\mathbf{F}_j^I \rangle^{rpE} \tag{13.3.27}$$

This matrix element is momentum independent and, therefore,

$$k_B \nabla M(x) = k_B \begin{pmatrix} 0 \\ -\frac{\partial M_{ij}}{\partial E}\mathbf{v}_j \\ \sum_j M_{ij}\frac{1}{m_j} + \mathbf{v}_i\frac{\partial M_{ij}}{\partial E}\mathbf{v}_j \end{pmatrix} \tag{13.3.28}$$

Concerning the noise terms, Eq. (**??**) suggest that $d\tilde{E} = -\mathbf{v}_i d\tilde{\mathbf{p}}_i$ and the only concern is to find a random force $d\tilde{\mathbf{p}}_i$ such that $d\tilde{\mathbf{p}}_i d\tilde{\mathbf{p}}_j = 2k_B M_{ij} dt$. The situation simplifies a lot if the matrix M_{ij} is diagonal. In this case, $d\tilde{\mathbf{p}}_i$ is just proportional to a vector Wiener process.

13.4   Conclusion

The resulting dynamic equations are rather simple, the bath is simply providing a friction, and the kinetic energy lost by the friction is invested in heating the bath. Of course, if the bath is very large, the last equation decouples because the temperature $T_B^{-1}(E) = \frac{\partial S_B}{\partial E}$ is practically a constant. In that case, we just have a Stokesian thermostate.

I think this approach makes sense if, when considering specific examples, we show that having an actual implementation of a thermal bath is, in fact, a rather elusive thing. This is not a positive result, I'm afraid.

13.5   GENERIC structure

We know that we should have a GENERIC structure because we can express the total energy of the composite system in terms of the CG variables. I believe that this structure should give us information about the reversible term $\langle \mathbf{F}_i^I \rangle^{rpE}$ and enforce conditions on the possible approximations for $M(x)$.

The reversible part of the dynamics should produce no increase in the entropy (13.3.12) of the CG level of description. The time derivative of the entropy due to the reversible motion is

$$\frac{dS}{dt}(r,p,E) = \frac{\partial S}{\partial \mathbf{q}_i}\frac{d\mathbf{q}_i}{dt}\Big|_{\text{rev}} + \frac{\partial S}{\partial \mathbf{p}_i}\frac{d\mathbf{p}_i}{dt}\Big|_{\text{rev}} + \frac{\partial S}{\partial E}\frac{dE}{dt}\Big|_{\text{rev}}$$
$$= G(r,p,E)\mathbf{F}_i^A(r)\mathbf{v}_i - \mathbf{F}_i^A(r)G(r,p,E)\mathbf{v}_i = 0 \tag{13.5.1}$$

Therefore, the reversible part automatically conserves the entropy.

Therefore the degeneracy condition $M(x)\nabla E(x) = 0$ now takes the form

$$M(x)\frac{\partial E}{\partial x} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & M_{ij} & M_{iE} \\ 0 & M_{Ei} & M_{EE} \end{pmatrix}\begin{pmatrix} -\mathbf{F}_j^A \\ \mathbf{v}_j \\ 1 \end{pmatrix} \tag{13.5.2}$$

This gives the following identities

$$M_{ij}\mathbf{v}_j + M_{iE} = 0$$
$$M_{jE}\mathbf{v}_j + M_{EE} = 0 \qquad\qquad (13.5.3)$$

which imply

$$M_{iE} = -M_{ik}\mathbf{v}_k$$
$$M_{EE} = -M_{jE}\mathbf{v}_j = M_{ik}\mathbf{v}_k\mathbf{v}_i \qquad\qquad (13.5.4)$$

These identities could be obtained also from the Green-Kubo expressions (17.5.6). By using these identities in the SDE (**??**)

The CG variables are $\Psi_{\mathbf{R},\mathbf{P}}$, $\Psi_{\mathbf{R}}$

14

A porous membrane

14.1   Description of the system

A gas flows slowly through a porous membrane separating two chambers.

15

Macroscopic Chemical Reactions

15.1 Keizer's theory

Keizer's theory encodes the law of mass action which seems to be the correct way to treat chemical reactions. The macroscopic variables are collected into a vector $\mathbf{n}$ that contains in each component the number of molecules of each specie. The evolution equation proposed by Keizer is given by the "canonical form"

$$\frac{d\mathbf{n}}{dt} = \boldsymbol{\omega}\left[\exp\{-\boldsymbol{\lambda}\cdot\mathbf{n}^+\} - \exp\{-\boldsymbol{\lambda}\cdot\mathbf{n}^-\}\right] \tag{15.1.1}$$

Here $\boldsymbol{\lambda} = \frac{\partial S}{\partial \mathbf{n}}(\mathbf{n})$ are the conjugate variables of the relevant variable $\mathbf{n}$, and $\mathbf{n}^\pm$ are "the molecular amounts involved in each direction of the elementary process of the reaction". The stochiometric coefficients of the reaction are $\boldsymbol{\omega} = \mathbf{n}^- - \mathbf{n}^+$.

Note that this equation satisfies a Second Law

$$\frac{dS}{dt}(\mathbf{n}) = \Omega\exp\{-\boldsymbol{\lambda}\cdot\mathbf{n}^-\}\mathscr{A}(\exp\{\mathscr{A} - 1) \geq 0 \tag{15.1.2}$$

where we have introduced the affinity

$$\mathscr{A} = \boldsymbol{\lambda}\cdot\boldsymbol{\omega} \tag{15.1.3}$$

The inequality is true because $\mathscr{A}(e^{\mathscr{A}} - 1) \geq 0$ and $\Omega > 0$. Note that for any vector $\mathbf{c}$ which is perpendicular to $\boldsymbol{\omega}$, we have $\mathbf{c}\cdot\boldsymbol{\omega} = 0$ and, therefore, $\dot{\mathbf{n}}\cdot\mathbf{c} = 0$, which means that $\mathbf{n}\cdot\mathbf{c}$ is a conserved quantity.

From a statistical mechanics point of view, Keizer's theory is incomplete, as it does not provide explicit first principles molecular expressions for the entropy $S(\mathbf{n})$ or the rate Ω.

15.1.1   The reaction $\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$

We consider now a chemical reaction of the form

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB} \tag{15.1.4}$$

For the chemical reaction $\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{AB}$ we have

$$
\begin{aligned}
\mathbf{n}^+ &= (1, 1, 0) \\
\mathbf{n}^- &= (0, 0, 1) \\
\boldsymbol{\omega} &= (-1, -1, 1) \\
\mathbf{c}_A &= (1, 0, 1) \\
\mathbf{c}_B &= (0, 1, 1)
\end{aligned}
\tag{15.1.5}
$$

Here $\mathbf{c}_A \cdot \mathbf{n} = N_A^0$ is the total number of atoms of type A (either free or bounded to a B atom) and $\mathbf{c}_B \cdot \mathbf{n} = N_B^0$ is the total number of atoms of type B.

Explicitly, the dynamics governing this reaction is given by

$$
\begin{aligned}
\dot{N}_A &= -\Omega(\exp\{-\lambda_A - \lambda_B\} - \exp\{-\lambda_{AB}\}) \\
\dot{N}_B &= -\Omega(\exp\{-\lambda_A - \lambda_B\} - \exp\{-\lambda_{AB}\}) \\
\dot{N}_{AB} &= \ \ \Omega(\exp\{-\lambda_A - \lambda_B\} - \exp\{-\lambda_{AB}\})
\end{aligned}
\tag{15.1.6}
$$

It is obvious that the three variables n_A, n_B, n_{AB} are not independent. In particular, conservation of the number of atoms N_A of type A and N_B of type B lead to the two conservation laws

$$
\begin{aligned}
N_A + N_{AB} &= N_A^0 \\
N_B + N_{AB} &= N_B^0
\end{aligned}
\tag{15.1.7}
$$

These two conservation laws imply that there is only one independent variable, the so called **progress variable** of the reaction, that we choose to be n_{AB}. The evolution of n_{AB} can be expressed as

$$\dot{N}_{AB} = \Omega \exp\{-\lambda_A - \lambda_B\}(1 - \exp\{\mathscr{A}\}) \tag{15.1.8}$$

where we have introduced the affinity (15.1.3). In this expression the conjugate variables $\lambda_A, \lambda_B, \lambda_{AB}$, which are the derivatives of entropy $S(N_A, N_B, N_{AB})$ should be understood as functions of N_A, N_B, N_{AB}. By expressing N_A, N_B as a function of N_{AB} from (15.1.7), we end up with a closed equation for N_{AB} in (15.1.8). **[In order to make close contact with the theory of CG, I would like to express λ_A, λ_B in terms of the affinity $\mathscr{A}$. In this way, we would obtain the correct flow.]**

The question now is what is the functional form of the entropy function $S(\mathbf{n})$. The entropy is related to the equilibrium probability through the Boltzmann-Planck-Einstein postulate

$$P^{\mathrm{eq}}(\mathbf{n}) = \exp\{S(\mathbf{n})\} \tag{15.1.9}$$

Because

$$P(\mathbf{n}) = \int dz \rho^{\mathrm{eq}}(z) \delta(\mathbf{n} - \hat{\mathbf{n}}(z)) \tag{15.1.10}$$

In the present case we have

$$P(N_A, N_B, N_{AB}) = \int dz \rho^{\mathrm{eq}}(z) \delta(N_A - \hat{N}_A(z)) \delta(N_B - \hat{N}_B(z)) \delta(N_{AB} - \hat{N}_{AB}(z))$$

$$= \delta(N_A + N_{AB} - N_A^0) \delta(N_B + N_{AB} - N_B^0) \int dz \rho^{\mathrm{eq}}(z) \delta(N_{AB} - \hat{N}_{AB}(z)) \tag{15.1.11}$$

Note that the integral depends implicitly on N_A^0, N_B^0 because we are integrating over the microscopic degrees of freedom z of all the atoms in the system. The two Dirac delta function reflect the conservation of atoms and state that the probability of finding a configuration N_A, N_B, N_{AB} that does not comply with the conservation relations (15.1.7) is zero. It is obvious that the only thing that can be called "entropy" in this expression is

$$\exp\left\{S(E, N_A^0, N_B^0, N_{AB})\right\} = \int dz \rho^{\mathrm{eq}}(z) \delta(N_{AB} - \hat{N}_{AB}(z)) \tag{15.1.12}$$

[**Conclusion: As long as we consider the CG mapping given by the phase functions (15.2.1), the Lagrange multipliers λ_A, λ_B loose their meaning! However, in reaction-diffusion systems, the number of particles N_A, N_B, N_{AB} are independent. The little cells are open systems.]**

The ideal solution and the usual law of mass action

To fix ideas, let us consider the ideal gas or ideal solution case. The functional form of the entropy is not know in general, but for an ideal gas the chemical potentials have the well-known form

$$\lambda_A = -\ln N_A + ctn$$
$$\lambda_B = -\ln N_B + ctn$$
$$\lambda_{AB} = -\ln N_{AB} + ctn \tag{15.1.13}$$

that leads to

$$\dot{N}_{AB} = C(N_A N_B - N_{AB}) \tag{15.1.14}$$

which is a common expression of the Law of Mass Action.

There are two different ways to solve this equation. Either one considers the extended

set

$$\dot{N}_A = -C(N_A N_B - N_{AB})$$
$$\dot{N}_B = -C(N_A N_B - N_{AB})$$
$$\dot{N}_{AB} = C(N_A N_B - N_{AB}) \tag{15.1.15}$$

and solves it for arbitrary initial conditions N_A, N_B, N_{AB}. The above equations obviously conserve the total numbers of atoms N_A^0, N_B^0 in (**??**) where the values N_A^0, N_B^0 are then fixed by the initial conditions. Or alternatively, one may use (**??**) to express $N_A = N_A^0 - N_{AB}$, $N_B = N_B^0 - N_{AB}$, and obtain a closed equation for N_{AB}.

$$\dot{n}_{AB} = C((N_A^0 - N_{AB})(N_B^0 - N_{AB}) - N_{AB}) \tag{15.1.16}$$

15.2   From the theory of coarse-graining

We have two types of atoms A, B that react to form a diatomic molecule AB. The CG variables are, in addition to the total energy of the system, the following phase functions

$$\hat{N}_{AB}(z) = \sum_{i_A i_B} \theta(\mathbf{r}_{i_A} - \mathbf{r}_{i_B})$$
$$\hat{N}_A(z) = N_A^0 - \hat{N}_{AB}(z)$$
$$\hat{N}_B(z) = N_B^0 - \hat{N}_{AB}(z) \tag{15.2.1}$$

Here, $\hat{N}_{AB}(z)$ counts the pairs of atoms A, B that are close together. If we understand that this grouping is a molecule AB, the variable $\hat{N}_{AB}(z)$ counts the number of molecules AB that correspond to the microstate z. The number $\hat{N}_A(z)$ of atoms of type A is obtained by subtracting from the total number of atoms N_A^0 those that are "bounded" with an atom of type B, and analogously for $\hat{N}_B(z)$. Note that the above three variables are not independent because N_A^0, N_B^0 are assumed to be known. We have a case of redundant information, treated in general in Sec **??**.

If we proceed formally as if the variables were independent, the relevant ensemble (**??**) corresponding to this level of description is

$$\bar{\rho}(z) = \frac{1}{Z[\beta, \lambda]} \exp\left\{-\beta \hat{H}(z) - \lambda_A \hat{N}_A(z) - \lambda_B \hat{N}_B(z) - \lambda_{AB} \hat{N}_{AB}(z)\right\}$$

$$Z[\beta, \lambda] = \int dz \exp\left\{-\beta \hat{H}(z) - \lambda_A \hat{N}_A(z) - \lambda_B \hat{N}_B(z) - \lambda_{AB} \hat{N}_{AB}(z)\right\} \tag{15.2.2}$$

The conjugate variables and the average of the CG variables are related by (**??**) that in

the present case becomes

$$\frac{\partial \Phi}{\partial \beta}[\beta, \lambda] = E$$

$$\frac{\partial \Phi}{\partial \lambda_A}[\beta, \lambda] = N_A$$

$$\frac{\partial \Phi}{\partial \lambda_B}[\beta, \lambda] = N_B$$

$$\frac{\partial \Phi}{\partial \lambda_{AB}}[\beta, \lambda] = N_{AB} \tag{15.2.3}$$

where the dimensionless thermodynamic potential is

$$\Phi[\beta, \lambda] = -\ln Z[\beta, \lambda] \tag{15.2.4}$$

The entropy (**??**) for this level of description is given by

$$S[E, N_A, N_B, N_{AB}] = -\Phi[\beta, \lambda] + \lambda_A N_A + \lambda_B N_B + \lambda_{AB} N_{AB} \tag{15.2.5}$$

where β, λ are understood as functions of the averages E, N_A, N_B, N_{AB}.

The conjugate variables are given in terms of the derivatives of the entropy functional

$$\frac{\partial S}{\partial E}[E, N_A, N_B, N_{AB}] = \beta$$

$$\frac{\partial S}{\partial N_A}[E, N_A, N_B, N_{AB}] = \lambda_A$$

$$\frac{\partial S}{\partial N_B}[E, N_A, N_B, N_{AB}] = \lambda_B$$

$$\frac{\partial S}{\partial N_{AB}}[E, N_A, N_B, N_{AB}] = \lambda_{AB} \tag{15.2.6}$$

The time derivatives of the CG variables now take the form

$$i\mathcal{L}\hat{N}_A(z) = -P$$

$$i\mathcal{L}\hat{N}_B(z) = -P$$

$$i\mathcal{L}\hat{N}_{AB}(z) = \quad P \tag{15.2.7}$$

where the production term is

$$P = \sum_{i_A j_B} \theta'\left(|\mathbf{r}_{i_A} - \mathbf{r}_{j_B}|\right) \mathbf{e}_{i_A j_B} \cdot \mathbf{v}_{i_A j_B} \tag{15.2.8}$$

Note that the time derivatives reflect the conservation of the total number of atoms of type A and B.

The dissipative matrix will have the following structure

$$
\mathscr{D}_{\mu\nu} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \mathscr{R} & \mathscr{R} & -\mathscr{R} \\ 0 & \mathscr{R} & \mathscr{R} & -\mathscr{R} \\ 0 & -\mathscr{R} & -\mathscr{R} & \mathscr{R} \end{pmatrix} \tag{15.2.9}
$$

The vanishing first row and column correspond to the total energy variable. The chemical reaction bit of the transport equations will be

$$
\begin{aligned}
\partial_t N_A(t) &= -\mathscr{R}\mathscr{A} \\
\partial_t N_B(t) &= -\mathscr{R}\mathscr{A} \\
\partial_t N_{AB}(t) &= \mathscr{R}\mathscr{A}
\end{aligned} \tag{15.2.10}
$$

where the affinity is

$$
\mathscr{A} \equiv -\lambda_A - \lambda_B + \lambda_{AB} \tag{15.2.11}
$$

Note that the structure of the affinity is

$$
\mathscr{A} = \boldsymbol{\omega}\cdot\boldsymbol{\lambda} \tag{15.2.12}
$$

The evolution of the averages (15.2.10) has been obtained under the assumption, in the entropy function, that the variables were independent. However, we used in the time derivatives (15.2.7) the explicit form of the CG variables (15.2.1) that obviously shows that they are not. As we already explained in the Sec. **??**, when the set of variables is redundant, it is not possible to invert the equations of state $a(\lambda)$.

The correct way to proceed is to use only one of the independent variables as, for example, $\hat{N}_{AB}(z)$. Alternatively, we may use the definitions (15.2.1) in the relevant ensemble (15.2.2), that then leads to

$$
\overline{\rho}(z) = \frac{1}{Z[\beta,\lambda]} \exp\left\{-\beta\hat{H}(z) - \lambda_A N_A^0 - \lambda_B N_B^0 - \mathscr{A}\hat{N}_{AB}(z)\right\}
$$

$$
Z[\beta,\lambda] = \int dz \exp\left\{-\beta\hat{H}(z) - \lambda_A N_A^0 - \lambda_B N_B^0 - \mathscr{A}\hat{N}_{AB}(z)\right\} \tag{15.2.13}
$$

Obviously, the factor involving λ_A, λ_B cancel in the numerator and denominator, and we are left with only one conjugate parameter, the affinity $\mathscr{A}$. The relevant ensemble is

$$
\overline{\rho}(z) = \frac{1}{Z(\beta,\mathscr{A},N_A^0,N_B^0)} \exp\left\{-\beta\hat{H}(z) - \mathscr{A}\hat{N}_{AB}(z)\right\}
$$

$$
Z(\beta,\mathscr{A},N_A^0,N_B^0) = \int dz \exp\left\{-\beta\hat{H}(z) - \mathscr{A}\hat{N}_{AB}(z)\right\} \tag{15.2.14}
$$

We have made explicit the dependence on the total number of atoms of each type of the partition function.

The thermodynamic potential is now a

$$\Phi(\beta, \mathscr{A}, N_A^0, N_B^0) = -\ln Z(\beta, \mathscr{A}, N_A^0, N_B^0) \tag{15.2.15}$$

whose derivative is

$$N_{AB} = \frac{\partial \Phi}{\partial \mathscr{A}}(\beta, \mathscr{A}) \tag{15.2.16}$$

This equation can be inverted, because the matrix of second derivatives of the thermodynamic potential has no zero eigenvalues now (as opposed to the thermodynamic potential (15.2.4)). In this way we may obtain the equation of state $\mathscr{A}(N_{AB})$ that gives the affinity in terms of the independent relevant variable N_{AB}. The entropy now becomes simply

$$S(E, N_{AB}, N_A^0, N_B^0) = -\Phi(\beta, \mathscr{A}, N_A^0, N_B^0) + \mathscr{A} N_{AB} \tag{15.2.17}$$

This entropy depends only on the independent relevant variable. The affinity is just the derivative of the entropy

$$\mathscr{A} = \frac{\partial S}{\partial N_{AB}}(E, N_{AB}) \tag{15.2.18}$$

The entropies (15.2.5) and (15.2.17) are different functions, as they depend on a different number of variables. Their relation is easily obtained to be

$$S[E, N_{AB}] = S[E, N_A^0 - N_{AB}, N_B^0 - N_{AB}, N_{AB}] \tag{15.2.19}$$

[Warning: The entropy depends implicitly on the total number N_A^0, N_B^0 of particles. They do not play the role of dynamic variables, but different values of N_A^0, N_B^0 give different values of the entropy. In that sense, we should include these variables, as well as the volume, in the description!]

[As far as I understand, in a local theory, when the cells can be considered as open systems, the total number of atoms N_A, N_B in each cell may vary, and we are entitled to consider them as independent variables. However, in well-mixed closed system, N_A, N_B are not independent, they are fixed. This is the truth. Therefore, we need to formulate Keizer's theory for a closed system. What is the prediction of Keizer for a closed system??]

15.3   The local approximation in non-local field theories (unsolved)

I want to show that the local form of the entropy may be obtained from the Gibbs-Jaynes entropy under the assumption that the conjugate variables vary little in molecular correlation distance.

The relevant ensemble for a field theory where the CG variables are $\hat{A}_{\mathbf{r}}(z)$ is

$$\overline{\rho}(z) = \rho_0 \frac{\exp\{-\int d\mathbf{r}\lambda(\mathbf{r})\hat{A}_{\mathbf{r}}(z)\}}{Z[\lambda]} \tag{15.3.1}$$

where the conjugate variables satisfy

$$A(\mathbf{r}) = -\frac{\delta \ln Z}{\delta \lambda(\mathbf{r})}[\lambda] \tag{15.3.2}$$

and the entropy is

$$S[\overline{\rho}] = -\ln Z[\lambda] + \int d\mathbf{r}\lambda(\mathbf{r})A(\mathbf{r}) \tag{15.3.3}$$

I want to find out what kind of relevant ensemble would lead to a local entropy of the form

$$S[\overline{\rho}] \simeq \int d\mathbf{r}s(A(\mathbf{r})) \tag{15.3.4}$$

If we take the functional derivative of this local form we have

$$\lambda(\mathbf{r}) = \frac{\delta S}{\delta A(\mathbf{r})} = s'(A(\mathbf{r})) \tag{15.3.5}$$

and the second derivative gives

$$\frac{\delta \lambda(\mathbf{r})}{\delta A(\mathbf{r}')} = s''(A(\mathbf{r}))\delta(\mathbf{r} - \mathbf{r}') \tag{15.3.6}$$

which is delta correlated. The inverse matrix is also diagonal, then the local approximation implies that

$$\frac{\delta A(\mathbf{r})}{\delta \lambda(\mathbf{r}')} = -\frac{\delta^2 \ln Z}{\delta \lambda(\mathbf{r})\delta \lambda(\mathbf{r}')}[\lambda] \simeq \frac{1}{s''(A(\mathbf{r}))}\delta(\mathbf{r} - \mathbf{r}') \tag{15.3.7}$$

What form of the relevant ensemble should we take in order to have a partition function with this property? Clearly, we need to have something like

$$\Phi(\lambda) = -\ln Z(\lambda) = \int d\mathbf{r}f(\lambda(\mathbf{r})) \tag{15.3.8}$$

Where

$$Z(\lambda) = \text{Tr}\left[\int d\mathbf{r}\lambda(\mathbf{r})\hat{A}_{\mathbf{r}}\right] \tag{15.3.9}$$

We can always write this as

$$\overline{\rho}(z) = \rho_0 \exp\left\{-\lambda(\mathbf{r}')\int d\mathbf{r}\hat{A}_{\mathbf{r}}(z)\right\}\frac{\exp\{-\int d\mathbf{r}(\lambda(\mathbf{r}) - \lambda(\mathbf{r}'))\hat{A}_{\mathbf{r}}(z)\}}{Z[\lambda]} \tag{15.3.10}$$

where the point $\mathbf{r}'$ is arbitrary. If we insert this into the Gibbs-Jaynes functional we have

$$S[\overline{\rho}] = -\text{Tr}\left[\overline{\rho}\ln\exp\left\{-\lambda(\mathbf{r}')\int d\mathbf{r}\hat{A}_{\mathbf{r}}(z)\right\}\frac{\exp\{-\int d\mathbf{r}(\lambda(\mathbf{r}) - \lambda(\mathbf{r}'))\hat{A}_{\mathbf{r}}(z)\}}{Z[\lambda]}\right]$$

$$= -\lambda(\mathbf{r}')A + \ln Z[\lambda] - \int d\mathbf{r}(\lambda(\mathbf{r}) - \lambda(\mathbf{r}'))A(\mathbf{r}) \tag{15.3.11}$$

where the extensive variable is

$$A = \int d\mathbf{r}A(\mathbf{r}) \tag{15.3.12}$$

The partition function is

$$Z[\lambda] = \text{Tr}\left[\rho_0 \exp\left\{-\lambda(\mathbf{r}')\hat{A}\right\}\exp\left\{-\int d\mathbf{r}(\lambda(\mathbf{r}) - \lambda(\mathbf{r}'))\hat{A}_{\mathbf{r}}(z)\right\}\right] \tag{15.3.13}$$

I cannot express this as a correlation function times conjugate variables (that would imply that neglecting variations of the conjugate variable in the length scale of correlations, we would obtain a local form).

16

Dielectric response

17

Blob models for complex molecules

We use the theory of coarse-graining in order to obtain the dynamic equations for coarse-grained representations, termed as thermal blobs, of complex molecules. Each molecule is described in a coarse way by the position and momentum of the center of mass and the internal energy of the molecule. The resulting dynamic equations have the structure of the energy conserving Dissipative Particle Dynamics model. Under suitable approximations, we provide explicit computable microscopic expressions for each object (entropy, mean force, friction and conductivity coefficients) appearing in the coarse-grained model. The model should be of use in situations in which energy transfer in non-isothermal situation needs to be captured while fully atomistic simulations are still too heavy to conduct.

17.1 Introduction

Molecular dynamics simulation of liquids made of complex molecules is limited by the large number of degrees of freedom involved. Only small samples and small time spans are accessible. In order to overcome this limitation there is much interest in constructing coarse-grained models in which each molecule is represented by a smaller number of degrees of freedom. Towards this aim much attention has been payed to the obtention of *effective or coarse-grained (CG) potentials* governing the interaction of the coarse degrees of freedom. Usually, certain functional forms for the CG potentials are assumed and they are parametrized in order to obtain structural properties like the radial distribution function. Several methods have been developed in order to obtain the appropriate parameters: iterative adjustment of potential parameters starting from an approximation based on the potentials of mean force [?],[?], through solving of the Ornstein-Zernicke equation [?], by using the inverse Monte Carlo technique [?],[?], by matching thermodynamic properties [?] or, more recently, by using directly the underlying all-atom interactions through a force matching procedure [?],[?]. The CG potentials produce correct results for the study of equilibrium and structural properties, as they have been designed just to recover such behaviour. However, dynamic properties like dif-

fusion or, more generally, time correlations, are not necessarily recovered from a simple molecular dynamics simulation using the CG potentials. The basic reason for this failure is that a coarse-graining procedure eliminates degrees of freedom that should appear in the coarse-grained dynamics in the form of *dissipation* and *thermal noise*, both connected through the Fluctuation-Dissipation theorem. Recently, the program of dynamic coarse-graining has produced dynamic equations for the positions and momenta of the centers of mass of the blobs [**?**],[**?**],[**?**],[**?**]. The resulting equations are generalizations of the Dissipative Particle Dynamics model, that include tensorial frictions. The model is isothermal by design and cannot be applied to non-isothermal situations in which transport of energy is important. In this Letter, we present a systematic CG procedure for the description of complex molecules when the transport of energy is an important aspect of the problem.

17.2   Coarse variables

Consider a fluid made of M molecules, identical for simplicity. Each molecule is composed of n atoms and the total number of atoms is $N = Mn$. Greek indices label molecules, while Latin indices label atoms. Monomers in a polymer chain could be also interpreted as "molecules" in what follows. At the *microscopic level* the system is described by the set of all positions $\mathbf{q}_i$ and momenta $\mathbf{p}_i = m_i \mathbf{v}_i$ $(i = 1, \cdots, N)$ of the atoms. The collection of all these degrees of freedom will be denoted by the microstate z. The microstate of the system evolves according to Hamilton's equations with Hamiltonian given by

$$\hat{H}(z) = \sum_i^N \frac{p_i^2}{2m_i} + V(q) \tag{17.2.1}$$

The potential energy is

$$V(q) = \frac{1}{2} \sum_{ij}^N \phi_{ij} \tag{17.2.2}$$

We have assumed a pair-wise potential energy, where ϕ_{ij} is the potential of interaction between atoms, which in general depends only on the distance between these atoms. Self interaction of the atoms is not considered, so $\phi_{ii} = 0$, etc.

Each complex molecule will be described at a coarse-level as a single *thermal blob*. We introduce the indicator $\delta_\mu(i)$ that takes the value 1 if the atom i belongs to the molecule μ and zero otherwise. This indicator satisfies

$$\sum_\mu^M \delta_\mu(i) = 1$$
$$\delta_\mu(i)\delta_\nu(i) = \delta_{\mu\nu}\delta_\mu(i) \tag{17.2.3}$$

These expressions reflect that the atom i belongs to one and only one blob. The variables used to describe the blob are the CoM variables of the blob, this is the position

$\hat{\mathbf{R}}_\mu$, momentum $\hat{\mathbf{P}}_\mu$ and internal energy $\hat{\mathscr{e}}_\mu$ of the blob. These coarse-variables are the following functions of the microstate

$$\hat{\mathbf{R}}_\mu(z) = \frac{1}{M_\mu} \sum_i^N m_i \mathbf{q}_i \delta_\mu(i)$$

$$\hat{\mathbf{P}}_\mu(z) = \sum_i^N \mathbf{p}_i \delta_\mu(i)$$

$$\hat{\mathscr{e}}_\mu(z) = \sum_i^N \frac{m_i}{2} (\mathbf{v}_i - \hat{\mathbf{v}}_\mu(z))^2 \delta_\mu(i) + \phi_\mu(z) \qquad (17.2.4)$$

where the mass, velocity, and potential energy of the μ-th blob are defined by

$$M_\mu = \sum_i^N m_i \delta_\mu(i)$$

$$\hat{\mathbf{v}}_\mu(z) = \frac{\hat{\mathbf{P}}_\mu(z)}{M_\mu}$$

$$\phi_\mu(z) = \frac{1}{2} \sum_{ij}^N \phi_{ij} \delta_\mu(i) \qquad (17.2.5)$$

The internal energy $\hat{\mathscr{e}}_\mu$ of the blob has a kinetic contribution due to the velocity relative to the CoM of each atom in the molecule plus the potential of interaction between atoms of different molecules.

The definition of the internal energy of the cells in Eq. (17.2.4) is crucial in order to write the Hamiltonian of the full system in terms of the CG variables. In fact, the Hamiltonian (4.2.1) can be exactly written as

$$\hat{H}(z) = \sum_\mu^M \left[\frac{\hat{\mathbf{P}}_\mu^2(z)}{2M_\mu} + \hat{\mathscr{e}}_\mu(z) \right] \qquad (17.2.6)$$

By using in Eq. (17.2.6) the Eqs. (17.2.3) and (17.2.4), we indeed recover (4.2.1). Because the microscopic Hamiltonian can be expressed in terms of the CG variables, the resulting dynamic equations will have the GENERIC structure.

Not only the total energy but also the total momentum can be expressed as a function of the microstate through the coarse-grained variables, this is

$$\hat{\mathbf{P}}(z) \equiv \sum_i^N \mathbf{p}_i = \sum_\mu^M \hat{\mathbf{P}}_\mu(z) \qquad (17.2.7)$$

The time derivatives of the coarse variables play a fundamental role in the form of the final form of the Fokker-Planck equation. The time derivative LX is the result of

applying the Liouville operator $L = \{\,\cdot\,, H\}$ to the CG variables. This gives

$$L\hat{\mathbf{R}}_\mu(z) = \mathbf{V}_\mu(z)$$

$$L\hat{\mathbf{P}}_\mu(z) = \sum_\nu^M \hat{\mathbf{F}}_{\mu\nu}(z)$$

$$L\hat{\mathscr{E}}_\mu(z) = \sum_\nu^M \left[\hat{Q}_{\mu\nu}(z) - \frac{1}{2}\hat{\mathbf{F}}_{\mu\nu}(z){\cdot}\mathbf{V}_{\mu\nu}(z) \right] \qquad (17.2.8)$$

where $\mathbf{V}_{\mu\nu} = \mathbf{V}_\mu - \mathbf{V}_\nu$ is the relative velocity and the total force $\hat{\mathbf{F}}_{\mu\nu}$ that the molecule ν exerts on molecule μ is defined as

$$\hat{\mathbf{F}}_{\mu\nu}(z) \equiv \sum_{ij}^N \delta_\mu(i)\delta_\nu(j)\mathbf{F}_{ij} \qquad (17.2.9)$$

Here, $\mathbf{F}_{ij}$ is the force that atom j exerts on atom i. Note that Newton's Third Law $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ implies that $\hat{\mathbf{F}}_{\mu\nu} = -\hat{\mathbf{F}}_{\nu\mu}$. Finally, we have introduced the energy transfer from blob ν to blob μ as

$$\hat{Q}_{\mu\nu}(z) = \sum_{ij}^N \mathbf{F}_{ij} \left(\frac{(\mathbf{v}_i - \mathbf{v}_\mu) + (\mathbf{v}_j - \mathbf{v}_\nu)}{2} \right) \delta_\mu(i)\delta_\nu(j)$$

$$(17.2.10)$$

Note that $\hat{Q}_{\mu\nu}(z) = -\hat{Q}_{\nu\mu}(z)$.

17.3   The equilibrium distribution

In the theory of coarse-graining, a crucial element is the equilibrium probability of the macroscopic variables. This probability, through the Einstein fluctuation formula leads to the explicit expression of the entropy of the selected coarse-grained level. The equilibrium probability of the macroscopic variables is given in Eq. (13.2.4) that in the present selection of CG variables becomes

$$P^{\mathrm{eq}}(\mathbf{R}, \mathbf{P}, \mathscr{E}) = \int dz \rho^{\mathrm{eq}}(z) \prod_\mu^M \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z))\delta(\mathbf{P}_\mu - \hat{\mathbf{P}}_\mu(z))\delta(\mathscr{E}_\mu - \hat{\mathscr{E}}_\mu(z)) \quad (17.3.1)$$

where $\rho^{\mathrm{eq}}(z)$ is the equilibrium ensemble at the microscopic level. This equilibrium ensemble is a function of the microstate z of the system through the dynamic invariants. For example, if we assume a molecular ensemble with total energy E_0 and total momentum $\mathbf{P}_0$, the above equilibrium distribution function will have the form

$$\rho^{\mathrm{eq}}(z) = \frac{1}{\Omega(E_0, \mathbf{P}_0)}\delta(\hat{H}(z) - E_0)\delta(\hat{\mathbf{P}}(z) - \mathbf{P}_0) \qquad (17.3.2)$$

where $\Omega(E_0, \mathbf{P}_0)$ is the normalization

$$\Omega(E_0, \mathbf{P}_0) = \int dz \delta(\hat{H}(z) - E_0)\delta(\hat{\mathbf{P}}(z) - \mathbf{P}_0) \qquad (17.3.3)$$

By inserting (17.3.2) into (17.3.1) and using (17.2.6),(17.2.7), we obtain

$$P^{\mathrm{eq}}(\mathbf{R}, \mathbf{P}, \mathscr{E}) = \frac{1}{\Omega(E_0, \mathbf{P}_0)} \delta\left(\sum_\mu^M \frac{\mathbf{P}_\mu^2}{2M_\mu} + \mathscr{E}_\mu - E_0\right) \delta\left(\sum_\mu^M \mathbf{P}_\mu - \mathbf{P}_0\right) \exp\left\{\frac{1}{k_B} S(\mathbf{R}, \mathscr{E})\right\}$$

$$(17.3.4)$$

where

$$S(\mathbf{R}, \mathbf{P}, \mathscr{E}) = k_B \ln \int dz \prod_\mu^M \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z))\delta(\mathbf{P}_\mu - \hat{\mathbf{P}}_\mu(z))\delta(\mathscr{E}_\mu - \hat{\mathscr{E}}_\mu(z)) \qquad (17.3.5)$$

The entropy function does not depend on the momenta $\mathbf{P}_\mu$, as can be easily seen by performing the change of variables from $\mathbf{v}_{i_\mu} \to \mathbf{v}_{i_\mu} - \mathbf{v}_\mu$.

17.4 Reversible part of the dynamics

The conditional averages are of the form

$$\langle \cdots \rangle^x = \frac{1}{\Omega(x)} \int dz \prod_\mu^M \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z))\delta(\mathbf{P}_\mu - \hat{\mathbf{P}}_\mu(z))\delta(\mathscr{E}_\mu - \hat{\mathscr{E}}_\mu(z)) \cdots \qquad (17.4.1)$$

The conditional averages of the time derivatives of the CG variables give the reversible part of the dynamics, this is

$$\left\langle L\hat{\mathbf{R}}_\mu \right\rangle^x = \mathbf{V}_\mu$$

$$\left\langle L\hat{\mathbf{P}}_\mu \right\rangle^x = \sum_\nu^M \mathbf{F}_{\mu\nu}$$

$$\left\langle L\hat{\mathscr{E}}_\mu \right\rangle^x = -\frac{1}{2} \sum_\nu^M \mathbf{F}_{\mu\nu} \cdot \mathbf{V}_{\mu\nu} \qquad (17.4.2)$$

The conditional average of the time derivative $L\hat{\mathbf{R}}_\mu$ of the position of the center of mass of the molecule μ is exactly the velocity. The mean force is defined as

$$\mathbf{F}_{\mu\nu} = \left\langle \hat{\mathbf{F}}_{\mu\nu} \right\rangle^x \qquad (17.4.3)$$

This force depends on $\mathbf{R}, \mathscr{E}$ but not on the momenta because the force $\hat{\mathbf{F}}_{\mu\nu}(z)$ depends only on the positions of the atoms. In a similar way, the conditional average of $\hat{Q}_{\mu\nu}(z)$ vanishes because, after a Galilean transformation to the CoM, it is a linear function of

velocities that vanishes on integrating over momenta.

The final reversible part of the dynamics has the form

$$\dot{\mathbf{R}}_\mu\Big|_{\text{rev}} = \mathbf{V}_\mu$$

$$\dot{\mathbf{P}}_\mu\Big|_{\text{rev}} = \sum_\nu^M \mathbf{F}_{\mu\nu}$$

$$\dot{\mathscr{E}}_\mu\Big|_{\text{rev}} = -\frac{1}{2}\sum_\nu^M \mathbf{F}_{\mu\nu}\cdot\mathbf{V}_{\mu\nu} \qquad (17.4.4)$$

These equations are manifestly Galilean invariant. They conserve total momentum because Newton's Third Law $\mathbf{F}_{\mu\nu} = -\mathbf{F}_{\nu\mu}$ is satisfied, a property inherited from the microscopic forces defined in (17.2.9).

They also conserve total energy

$$\frac{dE}{dt}(x) = \sum_\mu^M \left[\frac{\partial E}{\partial \mathbf{R}_\mu}\mathbf{V}_\mu + \frac{\partial E}{\partial \mathbf{P}_\mu}\mathbf{F}_\mu + \frac{\partial E}{\partial \epsilon_\mu}\left\langle L\hat{\mathscr{E}}_\mu\right\rangle^x\right] = 0$$

$$\sum_\mu^M \left[\mathbf{V}_\mu\mathbf{F}_\mu + \left\langle L\hat{\mathscr{E}}_\mu\right\rangle^x\right] = 0 \qquad (17.4.5)$$

The reversible equations (17.4.4) can be cast in the GENERIC form

$$\begin{pmatrix} \dot{\mathbf{R}}_\mu \\ \dot{\mathbf{P}}_\mu \\ \dot{\mathscr{E}}_\mu \end{pmatrix} = \sum_\nu^M \begin{pmatrix} \mathbf{0} & \delta_{\mu\nu}\mathbf{1} & \mathbf{0} \\ -\delta_{\mu\nu}\mathbf{1} & \mathbf{0} & \frac{1}{2}\left(\mathbf{F}_{\mu\nu} + \delta_{\mu\nu}\sum_\sigma \mathbf{F}_{\mu\sigma}\right) \\ \mathbf{0} & \frac{1}{2}\left(\mathbf{F}_{\mu\nu} - \delta_{\mu\nu}\sum_\sigma \mathbf{F}_{\mu\sigma}\right) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{V}_\nu \\ 1 \end{pmatrix}$$
$$(17.4.6)$$

Note that the structure of the L matrix contains not only pure kinematics, but also contains information about the thermodynamics of the system, in a way very much similar of what happens in the GENERIC formulation of hydrodynamics. If the L matrix satisfies the Jacobi identity, then the time structure of the reversible part of the dynamics is preserved [?]. Note, however, that it has not been possible to prove from the microscopic definition of the reversible matrix that it satisfy the Jacobi identity and that a microscopic foundation of this property is still lacking. The Jacobi identity involves the Poisson bracket

$$\{A,\{B,C\}\} + \{B,\{C,A\}\} + \{C,\{A,B\}\} = 0 \qquad (17.4.7)$$

for arbitrary phase functions $\hat{A}(z), \hat{B}(z), C(z)$. Let us introduce the vector

$$\mathbf{\Lambda}_{\mu\nu}^{\alpha} = \frac{1}{2}\left(\mathbf{F}_{\mu\nu}^{\alpha} + \delta_{\mu\nu}\sum_{\sigma}\mathbf{F}_{\mu\sigma}^{\alpha}\right) \qquad (17.4.8)$$

After a long calculation in which one uses that the functions A, B, C are arbitrary, the Jacobi identity reduces to the following restriction

$$\mathbf{\Lambda}_{\mu\nu}^{\alpha}\frac{\partial\mathbf{\Lambda}_{\mu'\nu'}^{\alpha'}}{\partial\mathscr{E}_{\nu}} - \frac{\partial\mathbf{\Lambda}_{\mu'\nu'}^{\alpha'}}{\partial\mathbf{R}_{\mu}^{\alpha}} = \mathbf{\Lambda}_{\mu'\nu}^{\alpha'}\frac{\partial\mathbf{\Lambda}_{\mu'\nu}^{\alpha}}{\partial\mathscr{E}_{\nu}} - \frac{\partial\mathbf{\Lambda}_{\mu\nu'}^{\alpha}}{\partial\mathbf{R}_{\mu'}^{\alpha}} \qquad (17.4.9)$$

Let us consider now the restriction (**??**)

$$k_B\frac{\partial L}{\partial x}\frac{\partial E}{\partial x} - \frac{\partial E}{\partial x}L\frac{\partial S}{\partial x} = 0 \qquad (17.4.10)$$

We evaluate each term in this equation

$$\frac{\partial E}{\partial x} \to \begin{pmatrix} \mathbf{0} \\ \mathbf{V}_{\nu} \\ 1 \end{pmatrix} \qquad \frac{\partial S}{\partial x} \to \begin{pmatrix} \frac{\partial S}{\partial\mathbf{R}_{\mu}} \\ 0 \\ \frac{\partial S}{\partial\mathscr{E}_{\mu}} \end{pmatrix} = \begin{pmatrix} \frac{\partial S}{\partial\mathbf{R}_{\mu}} \\ 0 \\ \frac{1}{T_{\mu}} \end{pmatrix}$$

$$L\frac{\partial S}{\partial x} \to \begin{pmatrix} \mathbf{0} \\ -\frac{\partial S}{\partial\mathbf{R}_{\mu}} + \sum_{\nu}^{M}\mathbf{F}_{\mu\nu}\frac{1}{2}\left(\frac{1}{T_{\mu}}+\frac{1}{T_{\nu}}\right) \\ 0 \end{pmatrix}$$

$$k_B\frac{\partial L}{\partial x} \to \begin{pmatrix} \mathbf{0} & \delta_{\mu\nu}\mathbf{1} & \mathbf{0} \\ -\delta_{\mu\nu}\mathbf{1} & \mathbf{0} & \frac{1}{2}\left(\mathbf{F}_{\mu\nu}+\delta_{\mu\nu}\sum_{\sigma}\mathbf{F}_{\mu\sigma}\right) \\ \mathbf{0} & \frac{1}{2}\left(\mathbf{F}_{\mu\nu}-\delta_{\mu\nu}\sum_{\sigma}\mathbf{F}_{\mu\sigma}\right) & \mathbf{0} \end{pmatrix}\begin{pmatrix} \frac{\partial}{\partial\mathbf{R}_{\mu}} \\ \frac{\partial}{\partial\mathbf{P}_{\mu}} \\ \frac{\partial}{\partial\mathscr{E}_{\mu}} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{0} \\ \sum_{\mu}^{M}\frac{\partial}{\partial\mathscr{E}_{\mu}}\left[\frac{1}{2}\left(\mathbf{F}_{\mu\nu}+\delta_{\mu\nu}\sum_{\sigma}\mathbf{F}_{\mu\sigma}\right)\right] \\ \mathbf{0} \end{pmatrix} \qquad (17.4.11)$$

In this way, (17.4.10) becomes

$$\sum_{\mu}^{M}\mathbf{V}_{\mu}\left[\frac{\partial S}{\partial\mathbf{R}_{\mu}} - \frac{1}{2}\sum_{\nu}^{M}\mathbf{F}_{\mu\nu}\left(\frac{1}{T_{\mu}}+\frac{1}{T_{\nu}}\right)\right] + k_B\sum_{\nu}^{M}\frac{1}{2}\left[\frac{\partial}{\partial\mathscr{E}_{\mu}}+\frac{\partial}{\partial\mathscr{E}_{\nu}}\right]\mathbf{F}_{\mu\nu}\right] = 0 \quad (17.4.12)$$

The term under square brackets is independent of the velocity implying that this term should vanish for each μ, this is

$$\frac{\partial S}{\partial \mathbf{R}_\mu} = \sum_\nu^M \mathbf{F}_{\mu\nu} \frac{1}{2} \left(\frac{1}{T_\mu} + \frac{1}{T_\nu} \right) + k_B \sum_\nu^M \frac{1}{2} \left[\frac{\partial}{\partial \mathscr{E}_\mu} + \frac{\partial}{\partial \mathscr{E}_\nu} \right] \mathbf{F}_{\mu\nu} \qquad (17.4.13)$$

This exact and rigorous equation reflects the condition that the reversible part of the dynamics has as stationary solution the equilibrium distribution (17.3.4). This result is important as it connects the derivatives of the entropy of the thermal blob level of description, with the average forces between blobs (multiplied with certain temperature factors). There is an additional small term, of the order of k_B related to the derivatives of the average force with respect to the internal energies. The condition (17.4.13) is a powerful modeling guide because *it needs to be satisfied for any model of the entropy, and any approximate calculation of the average pair force* $\mathbf{F}_{\mu\nu}$ *that we may propose.*

17.5   Irreversible part of the dynamics

We now consider the irreversible part of the dynamics, and we need to compute the friction matrix M, which will have the form

$$\mathbf{M}_{\mu\nu}(R, P, \mathscr{E}) = \frac{1}{k_B} \begin{pmatrix} \int_0^\infty dt \langle \delta L\mathbf{R}_\mu \delta L\mathbf{R}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathbf{R}_\mu \delta L\mathbf{P}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathbf{R}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x \\ \int_0^\infty dt \langle \delta L\mathbf{P}_\mu \delta L\mathbf{R}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathbf{P}_\mu \delta L\mathbf{P}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathbf{P}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x \\ \int_0^\infty dt \langle \delta L\mathscr{E}_\mu \delta L\mathbf{R}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathscr{E}_\mu \delta L\mathbf{P}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathscr{E}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x \end{pmatrix}$$

$$= \frac{1}{k_B} \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \int_0^\infty dt \langle \delta \mathbf{F}_\mu \delta \mathbf{F}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta \mathbf{F}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x \\ \mathbf{0} & \int_0^\infty dt \langle \delta L\mathscr{E}_\mu \delta \mathbf{F}_\nu(t) \rangle^x & \int_0^\infty dt \langle \delta L\mathscr{E}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x \end{pmatrix} \qquad (17.5.1)$$

In order to figure out the structure of the matrix M in more detail, let us consider the element in the second row last column, that involves the correlation $\langle \delta \mathbf{F}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x$.

$$\langle \delta \mathbf{F}_\mu \delta L\mathscr{E}_\nu(t) \rangle^x = \sum_{\nu'}^M \left\langle \delta \mathbf{F}_\mu \left[\delta Q_{\nu\nu'}(t) - \frac{1}{2} \delta \mathbf{F}_{\nu\nu'}(t) \mathbf{v}_{\nu\nu'}(t) \right] \right\rangle^x$$

$$= \langle \delta \mathbf{F}_\mu \delta Q_\nu(t) \rangle^x - \sum_{\nu'}^M \frac{1}{2} \langle \delta \mathbf{F}_\mu \delta \mathbf{F}_{\nu\nu'}(t) \rangle^x \mathbf{v}_{\nu\nu'} \qquad (17.5.2)$$

where in the last equality we have assumed that the velocities of the CoM hardly change in the time scale in which the correlations decay, allowing to approximate $\mathbf{v}_{\nu\nu'}(t) \approx \mathbf{v}_{\nu\nu'}$. This is consistent with the Markovian approximation implicit in the use of the FPE that assumes a clear time scale separation between the CG variables (i.e. the velocity) and the memory kernel given by the Green-Kubo correlations.

In a similar way, we may compute the energy-energy Green-Kubo correlation

$$\langle \delta L \mathcal{E}_\mu \delta L \mathcal{E}_\nu(t) \rangle^x = \sum_{\mu'\nu'}^M \left\langle \left[\delta Q_{\mu\mu'} - \frac{1}{2}\delta \mathbf{F}_{\mu\mu'}\mathbf{v}_{\mu\mu'} \right] \left[\delta Q_{\nu\nu'}(t) - \frac{1}{2}\delta \mathbf{F}_{\nu\nu'}(t)\mathbf{v}_{\nu\nu'}(t) \right] \right\rangle^x$$

$$= \sum_{\mu'\nu'}^M \langle \delta Q_{\mu\mu'} \delta Q_{\nu\nu'}(t) \rangle^x - \frac{1}{2}\sum_{\mu'\nu'}^M \langle \delta Q_{\mu\mu'}\delta \mathbf{F}_{\nu\nu'}(t) \rangle^x \, \mathbf{v}_{\nu\nu'} - \frac{1}{2}\sum_{\mu'\nu'}^M \mathbf{v}_{\mu\mu'} \langle \delta \mathbf{F}_{\mu\mu'}\delta Q_{\nu\nu'}(t) \rangle^x$$

$$+ \frac{1}{4}\sum_{\mu'\nu'}^M \mathbf{v}_{\mu\mu'} \cdot \langle \delta \mathbf{F}_{\mu\mu'}\delta \mathbf{F}_{\nu\nu'}(t) \rangle^x \cdot \mathbf{v}_{\nu\nu'} \qquad (17.5.3)$$

Let us construct now the irreversible part of the dynamics, which is of the form $M\frac{\partial S}{\partial x}$. The gradient of the entropy is

$$\frac{\partial S}{\partial x} \rightarrow \begin{pmatrix} \frac{\partial S}{\partial \mathbf{R}_\nu} \\ \frac{\partial S}{\partial \mathbf{P}_\nu} \\ \frac{\partial S}{\partial E_\nu} \end{pmatrix} = \begin{pmatrix} \frac{\partial S}{\partial \mathbf{R}_\nu} \\ \mathbf{0} \\ \frac{1}{T_\nu} \end{pmatrix} \qquad (17.5.4)$$

where we have introduced the temperature T_ν of molecule ν by analogy with the usual definition of temperature for macroscopic systems. By multiplying the friction matrix (17.5.1) with this vector we obtain the irreversible part of the dynamics as

$$\dot{\mathbf{R}}_\mu \Big|_{irr} = 0$$

$$\dot{\mathbf{P}}_\mu \Big|_{irr} = \sum_{\nu\mu'\nu'}^M \mathbf{\Lambda}_{\mu\mu'\nu\nu'} \frac{1}{T_\nu} - \frac{1}{2}\sum_{\nu\mu'\nu'}^M \mathbf{\Gamma}_{\mu\mu'\nu\nu'} \cdot \mathbf{v}_{\nu\nu'} \frac{1}{T_\nu}$$

$$\dot{\mathcal{E}}_\mu \Big|_{irr} = \sum_{\nu\mu'\nu'}^M \sigma_{\mu\mu'\nu\nu'} \frac{1}{T_\nu} - \frac{1}{2}\sum_{\nu\mu'\nu'}^M \left[\mathbf{\Lambda}_{\nu\nu'\mu\mu'} \cdot \mathbf{v}_{\nu\nu'} + \frac{1}{2}\sum_{\mu'\nu'}^M \mathbf{v}_{\mu\mu'} \cdot \mathbf{\Lambda}_{\mu\mu'\nu\nu'} \right] \frac{1}{T_\nu} + \frac{1}{4}\sum_{\nu\mu'\nu'}^M \mathbf{v}_{\mu\mu'} \cdot \mathbf{\Gamma}_{\mu\mu'\nu\nu'} \cdot \mathbf{v}_{\nu\nu'} \frac{1}{T_\nu}$$

$$(17.5.5)$$

where we have introduced the Green-Kubo expressions

$$\mathbf{\Gamma}_{\mu\mu'\nu\nu'} \equiv \frac{1}{k_B}\int_0^\infty dt \, \langle \delta \mathbf{F}_{\mu\mu'}\delta \mathbf{F}_{\nu\nu'}(t) \rangle^\alpha$$

$$\mathbf{\Lambda}_{\mu\mu'\nu\nu'} \equiv \frac{1}{k_B}\int_0^\infty dt \, \langle \delta \mathbf{F}_{\mu\mu'}\delta Q_{\nu\nu'}(t) \rangle^\alpha$$

$$\sigma_{\mu\mu'\nu\nu'} \equiv \frac{1}{k_B}\int_0^\infty dt \, \langle \delta Q_{\mu\mu'}\delta Q_{\nu\nu'}(t) \rangle^x \qquad (17.5.6)$$

By collecting the reversible (17.4.4) and irreversible parts (17.5.5) of the dynamics we

arrive at the final equations

$$\dot{\mathbf{R}}_\mu = \mathbf{V}_\mu$$

$$\dot{\mathbf{P}}_\mu = \sum_\nu^M \mathbf{F}_{\mu\nu} + \sum_{\nu\mu'\nu'}^M \mathbf{\Lambda}_{\mu\mu'\nu\nu'} \frac{1}{T_\nu} - \frac{1}{2} \sum_{\nu\mu'\nu'}^M \mathbf{\Gamma}_{\mu\mu'\nu\nu'} \cdot \mathbf{V}_{\nu\nu'} \frac{1}{T_\nu}$$

$$\dot{\hat{\mathscr{E}}}_\mu = -\frac{1}{2} \sum_\nu^M \mathbf{F}_{\mu\nu} \cdot \mathbf{V}_{\mu\nu} + \sum_{\nu\mu'\nu'}^M \sigma_{\mu\mu'\nu\nu'} \frac{1}{T_\nu} - \frac{1}{2} \sum_{\nu\mu'\nu'}^M \left[\mathbf{\Lambda}_{\nu\nu'\mu\mu'} \cdot \mathbf{v}_{\nu\nu'} + \frac{1}{2} \sum_{\mu'\nu'}^M \mathbf{v}_{\mu\mu'} \cdot \mathbf{\Lambda}_{\mu\mu'\nu\nu'} \right] \frac{1}{T_\nu}$$

$$+ \frac{1}{4} \sum_{\nu\mu'\nu'}^M \mathbf{V}_{\mu\mu'} \cdot \mathbf{\Gamma}_{\mu\mu'\nu\nu'} \cdot \mathbf{V}_{\nu\nu'} \frac{1}{T_\nu} \tag{17.5.7}$$

These equations give the deterministic part of the dynamics. By adding the stochastic drift term $k_B \nabla M$ and the random terms $d\tilde{x}$ we would obtain the final SDE equations for this level of description. We refrain to do this until some modeling simplifications are done in the next section.

17.6   Models

The previous equations have been obtained under the sole hypothesis that the CG variables evolve in time scales well separated from the time scales of the memory kernels entering the *exact* dynamic equation for the CG variables. Nevertheless, they are still too complex to be of any use. The complexity arises essentially from the definition of the entropy function as a multidimensional integral in phase space, which is impossible to compute in general. The entropy or rather its derivatives are closely related to the mean force, which is also impossible to compute explicitly or exactly. Finally, the Green-Kubo expressions also involve conditional expectations, complicated by the fact that the dynamics appearing in the correlation functions is given in terms of the uncomputable projected dynamics. While the framework is very powerful as it gives the structure of the equations, there is still the need of performing several modeling assumptions. We are guided by essentially by physical intuition, although some exact results like (17.4.13) are useful to validate the intuitions. Nevertheless, the validity of these approximation can only be made a posteriori by comparing the results of the coarse-grained model with the results of MD simulations if available, or experiments.

17.6.1   A model for the entropy

We may write the internal energy in the form

$$\hat{\mathscr{E}}_\mu(z) = \sum_i^N \frac{m_i}{2} (\mathbf{v}_i - \hat{\mathbf{v}}_\mu(z))^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z) + \frac{1}{2} \sum_{\nu \neq \mu} \phi_{\mu\nu}^{\text{inter}}(z) \tag{17.6.1}$$

where the intra-molecular and inter-molecular potential contributions are defined as

$$\phi_\mu^{\text{intra}}(z) \equiv \frac{1}{2} \sum_{ij}^{N} \phi_{ij} \delta_\mu(i) \delta_\mu(j)$$

$$\phi_{\mu\nu}^{\text{inter}}(z) \equiv \frac{1}{2} \sum_{ij}^{N} \phi_{ij} \delta_\mu(i) \delta_\nu(j) \qquad (17.6.2)$$

The intramolecular potential $\phi_\mu^{\text{intra}}(z)$ captures the potential energy due to pairs of atoms within the same blob and only depends on the coordinates of the atoms of blob μ. The intermolecular potential $\phi_{\mu\nu}^{\text{inter}}(z)$ is the potential energy of the pairs of atoms that are in different blobs μ, ν.

Now, assume that we could approximate the microscopic intermolecular potential in the form

$$\frac{1}{2} \sum_{\nu \neq \mu}^{N} \sum_{ij} \phi_{ij} \delta_\mu(i) \delta_\nu(j) \approx \frac{1}{2} \sum_\nu V^{\text{eff}}(R_{\mu\nu}) = \Phi_\mu(\{R\}) \qquad (17.6.3)$$

where $V^{\text{eff}}(R)$ is an effective pair potential between the two blobs. In this way, all the atomic dependence of that part of the internal energy is assumed to be well represented just by the centers of mass of the two blobs.

By inserting this approximation in the entropy function (17.3.5)

$$S(\mathbf{R}, \mathbf{P}, \mathcal{E}) = k_B \ln \int dz \prod_\mu^{M} \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z)) \delta(\mathbf{P}_\mu - \hat{\mathbf{P}}_\mu(z)) \delta\left(\mathcal{E}_\mu - \Phi_\mu(R) - \sum_i^{N} \frac{m_i}{2}(\mathbf{v}_i - \hat{\mathbf{v}}_\mu(z))^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z) \right)$$

$$= k_B \ln \int dz \prod_\mu^{M} \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z)) \delta(\hat{\mathbf{P}}_\mu(z)) \delta\left(\mathcal{E}_\mu - \Phi_\mu(R) - \sum_i^{N} \frac{m_i}{2} \mathbf{v}_i^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z) \right)$$

$$= k_B \ln \prod_\mu^{M} \int dz_\mu \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu(z_\mu)) \delta(\hat{\mathbf{P}}_\mu(z_\mu)) \delta\left(\mathcal{E}_\mu - \Phi_\mu(R) - \sum_i^{N} \frac{m_i}{2} \mathbf{v}_i^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z_\mu) \right)$$

$$(17.6.4)$$

where the change of variables $\mathbf{v}_i - \hat{\mathbf{v}}_\mu(z) = \mathbf{v}_i'$ has been performed in the second line. In the above equation z_μ denote the degrees of freedom of molecule μ. It is apparent that the integrals are independent, leading to the simple result

$$S(\mathbf{R}, \mathbf{P}, \mathcal{E}) = \sum_\mu \mathscr{S}_\mu(\mathbf{R}, \mathscr{E} - \Phi) \qquad (17.6.5)$$

where the entropy of blob μ is given by

$$
\mathscr{S}_\mu(\mathbf{R}, \mathscr{E} - \Phi) = k_B \ln \int dz_\mu \delta(\mathbf{R}_\mu - \hat{\mathbf{R}}_\mu) \delta \left(\mathscr{E}_\mu - \Phi_\mu(R) - \sum_i^N \frac{m_i}{2} \mathbf{v}_i^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z_\mu) \right)
$$

(17.6.6)

Because all the potential terms are translation invariant, we may perform the change of variables $\mathbf{q}_i' = \mathbf{r}_i - \mathbf{R}_\mu$ with the result

$$
\mathscr{S}_\mu(\mathbf{R}, \mathscr{U}) = k_B \ln \int d\mathbf{r}_{i_\mu}' d\mathbf{p}_{i_\mu} \delta(\hat{\mathbf{R}}_\mu) \delta \left(\mathscr{U}_\mu - \sum_i^N \frac{m_i}{2} \mathbf{v}_i^2 \delta_\mu(i) + \phi_\mu^{\text{intra}}(z_\mu) \right) = \mathscr{S}_\mu(\mathscr{U}_\mu)
$$

(17.6.7)

This shows that the functional form of the entropy of the blob is explicitly independent of the position of center of mass and only depends on the intrinsic internal energy $\mathscr{U}_\mu = \mathscr{E}_\mu - \Phi_\mu(R)$.

We may now check whether this model of the entropy complies with the GENERIC restriction (17.4.13). We take the derivative of the entropy function with respect to the position of blob μ, by using the chain rule

$$
\frac{\partial S}{\partial \mathbf{R}_\mu} = \frac{\partial}{\partial \mathbf{R}_\mu} \sum_\nu \mathscr{S}_\nu(\mathscr{E}_\nu - \Phi_\nu(R)) = -\sum_\nu \frac{1}{T(\mathscr{E}_\nu - \Phi_\nu(R))} \frac{\partial \Phi_\nu}{\partial \mathbf{R}_\mu}(R)
$$

(17.6.8)

where we have introduced the temperature of the blob as

$$
\frac{1}{T_\mu(\mathscr{U}_\mu)} \equiv \frac{\partial \mathscr{S}_\mu}{\partial \mathscr{U}_\mu}(\mathscr{U}_\mu)
$$

(17.6.9)

We now look at the derivatives of the potential energy

$$
-\frac{\partial \Phi_\nu}{\partial \mathbf{R}_\mu}(R) = -\frac{\partial}{\partial \mathbf{R}_\mu} \frac{1}{2} \sum_{\nu'} V^{\text{eff}}(\mathbf{R}_{\nu\nu'}) = \frac{1}{2} \left[\mathbf{F}_{\mu\nu} + \delta_{\mu\nu} \sum_{\nu'} \mathbf{F}_{\mu\nu'} \right]
$$

(17.6.10)

where the mean force is defined as the gradient of the effective potential

$$
\mathbf{F}_{\mu\nu} = -\boldsymbol{\nabla} V^{\text{eff}}(R_{\mu\nu})
$$

(17.6.11)

This force is independent of the internal energies of the blobs. Therefore, by inserting (17.6.10) into (17.6.8) we see that the GENERIC restriction (17.4.13) is fulfilled by the model of the entropy that we propose in this section. Note that more complex models could be in principle devised as, for example, effective potentials with a dependence on internal energies. However, the restriction (17.4.13) is a rather stringent condition on the different modeling possibilities.

We need to fix two different functions of one variable $\mathscr{S}(\mathscr{U})$ and $V^{\text{eff}}(R)$ in order to complete the present model for the entropy. Note that the functional form of the entropy $\mathscr{S}(\mathscr{U})$ of a single blob can be obtained by running an MD simulation of a single

isolated molecule and computing the average kinetic energy as a function of the total energy. This gives, through the equipartition theorem, the temperature as a function of the total energy. Because the temperature is given in terms of the derivative of the entropy with respect to the energy, we may infer the functional form of the entropy up to an additive constant, which is not important. Once the function $\mathscr{S}(\mathscr{U})$ is determined, there are several possibilities for obtaining the functional form of the effective potential. For example, we may compute the average force between two blobs at a fixed distance and infer from it the effective potential, as was done in Ref. [?]. Or we may use the relative entropy method introduced by Scott Shell in order to find the best parametrized model for the effective potential that fits a given trajectory of an MD simulation [?].

17.6.2   Model for the friction matrix

The friction matrix involves the correlations of the projected currents $\delta\hat{\mathbf{F}}_{\mu\nu}(t)$ and $\delta\hat{Q}_{\mu\nu}(t)$. One way to model the friction matrix is through the modeling of the projected currents in terms of white noises. This is also compatible with modeling the thermal noise as linear combinations of white noise, that, through the Fluctuation-Dissipation theorem, have covariances given by the friction matrix. In many occasions, it is simpler to model the random terms (which are vectors) than the friction matrix (which is a matrix).

Consider the projected force $\delta\hat{\mathbf{F}}_{\mu\nu}$ which gives the fluctuations of the force (with respect to the average force $\langle\mathbf{F}_{\mu\nu}\rangle^x$) that blob ν exerts on blob μ. This is a rapidly fluctuating force that will be modeled as white noise. In principle, we expect that this force will have two components, one directed along the line joining the particles, and another one which is perpendicular to that line. Therefore, we model this fluctuating force as

$$\delta\hat{\mathbf{F}}_{\mu\nu} \mapsto \tilde{\mathbf{F}}_{\mu\nu} \equiv A_{\mu\nu}(x)\mathbf{e}_{\mu\nu}\mathscr{W}_{\mu\nu}(t) + B_{\mu\nu}(x)\mathbf{e}_{\mu\nu} \times \mathscr{T}_{\mu\nu}(t) \qquad (17.6.12)$$

where $\mathbf{e}_{\mu\nu} = (\mathbf{r}_\mu - \mathbf{r}_\nu)/|\mathbf{r}_\mu - \mathbf{r}_\nu|$ is the unit vector joining the centers of the blobs. The prefactors $A_{\mu\nu}(x), B_{\mu\nu}(x)$ may depend in general on the state of the system and will be determined later on. $\mathscr{W}_{\mu\nu}(t)$ is a scalar white noise, while $\mathscr{T}_{\mu\nu}(t)$ is a vector of white noises. In Ref. [?] we have postulated different random forces that, nevertheless, lead to the same friction matrix. However the form presented here is much simpler and requires a smaller number of random number generations. The structure (17.6.12) of the random forces has been proposed recently by Groot [?].

In a similar way, the stochastic energy transfer between two blobs will be modeled as

$$\delta\hat{Q}_{\mu\nu} \mapsto \tilde{Q}_{\mu\nu} \equiv C_{\mu\nu}(x)V_{\mu\nu}(t) \qquad (17.6.13)$$

where the prefactor $C_{\mu\nu}(x)$ may depend on the state and $V_{\mu\nu}(t)$ is a white noise.

In order to satisfy the microscopic symmetries $\hat{\mathbf{F}}_{\mu\nu} = -\hat{\mathbf{F}}_{\nu\nu}$, $\hat{Q}_{\mu\nu} = -\hat{Q}_{\nu\nu}$ we will

assume the symmetries

$$A_{\mu\nu} = A_{\nu\mu}$$
$$B_{\mu\nu} = B_{\nu\mu}$$
$$C_{\mu\nu} = -C_{\nu\mu}$$
$$\mathscr{W}_{\mu\nu}(t) = \mathscr{W}_{\nu\mu}(t)$$
$$\mathscr{T}_{\mu\nu}(t) = \mathscr{T}_{\nu\mu}(t)$$
$$\mathscr{V}_{\mu\nu}(t) = \mathscr{V}_{\nu\mu}(t) \tag{17.6.14}$$

As for the stochastic properties of the white noise, we postulate

$$\langle \mathscr{W}_{\mu\mu'}(t)\mathscr{W}_{\nu\nu'}(t')\rangle = [\delta_{\mu\nu}\delta_{\mu'\nu'} + \delta_{\mu\nu'}\delta_{\nu\mu'}]\,\delta(t-t')$$
$$\left\langle \mathscr{T}^{\alpha}_{\mu\mu'}(t)\mathscr{T}^{\beta}_{\nu\nu'}(t')\right\rangle = \delta^{\alpha\beta}\,[\delta_{\mu\nu}\delta_{\mu'\nu'} + \delta_{\mu\nu'}\delta_{\nu\mu'}]\,\delta(t-t')$$
$$\langle \mathscr{V}_{\mu\mu'}(t)\mathscr{V}_{\nu\nu'}(t')\rangle = [\delta_{\mu\nu}\delta_{\mu'\nu'} + \delta_{\mu\nu'}\delta_{\nu\mu'}]\,\delta(t-t')$$
$$\langle \mathscr{W}_{\mu\mu'}(t)\mathscr{T}(t')\rangle = 0$$
$$\langle \mathscr{W}_{\mu\mu'}(t)\mathscr{V}_{\nu\nu'}(t')\rangle = 0$$
$$\langle \mathscr{T}_{\mu\mu'}(t)\mathscr{V}_{\nu\nu'}(t')\rangle = 0 \tag{17.6.15}$$

Implicit in this stochastic properties is the assumption that the force (and also the stochastic energy transfer) for a given pair $\mu\mu'$ is uncorrelated from the force of any other pair $\nu\nu'$ except if this is the same pair, this is, either $\mu = \nu, \mu' = \nu'$ or $\mu = \nu', \nu = \mu'$. The last two identities reflect that forces and heat fluxes are uncorrelated to each other, which is also a strong modeling assumption. In fact, these two last identities imply that we are assuming that the following correlation is negligibly small

$$\langle \delta\mathbf{F}_{\mu}\delta Q_{\nu}(t)\rangle^{x} \approx 0 \tag{17.6.16}$$

Indeed at time $t = 0$ this correlation vanishes because Q_{μ} is linear in the relative velocities. Also, the equilibrium average (i.e. the result of averaging the conditional expectation (17.6.16) with respect to the equilibrium measure $P^{\text{eq}}(x)$) being isotropic would produce a vanishing result for the vectorial quantity being averaged. Although we do not have further, stronger arguments in favor of the vanishing at later times, the theory that will result after neglecting this correlations is much simpler, while retaining the main physical ingredients.

With these modeling assumptions the different elements of the friction matrix become

$$\mathbf{\Gamma}^{\alpha\beta}_{\mu\mu'\nu\nu'} = [\delta_{\mu\nu}\delta_{\mu'\nu'} + \delta_{\mu\nu'}\delta_{\nu\mu'}]\left\{A_{\mu\mu'}A_{\nu\nu'}\mathbf{e}^{\alpha}_{\mu\mu'}\mathbf{e}^{\beta}_{\nu\nu'} + B_{\mu\mu'}B_{\nu\nu'}(\delta^{\alpha\beta}\mathbf{e}_{\mu\mu'}\cdot\mathbf{e}_{\nu\nu'} - \mathbf{e}^{\beta}_{\mu\mu'}\mathbf{e}^{\beta}_{\nu\nu'}\right\}$$
$$\mathbf{\Lambda}_{\mu\mu'\nu\nu'} = 0$$
$$\sigma_{\mu\mu'\nu\nu'} = [\delta_{\mu\nu}\delta_{\mu'\nu'} + \delta_{\mu\nu'}\delta_{\nu\mu'}]\,C_{\mu\mu'}C_{\nu\nu'} \tag{17.6.17}$$

where the factor 2 disappears because the integral over the delta function goes not from $-\infty$ to ∞ but rather only from 0 to ∞.

We now consider the microscopic expression of the coefficients $A_{\mu\nu}, B_{\mu\nu}, C_{\mu\nu}$. This

microscopic expression can be obtained from the combination of (17.5.6) and the model (17.6.17). Let the indices be $\mu' = \nu$ and $\nu' = \mu$ with $\mu \neq \nu$. Then, take the trace of $\mathbf{\Gamma}^{\alpha\beta}_{\mu\nu\mu\nu}$ and also contract $\mathbf{\Gamma}^{\alpha\beta}_{\mu\nu\mu\nu}$ with the dyadic $\mathbf{e}^\alpha_{\mu\nu}\mathbf{e}^\beta_{\mu\nu}$. The trace gives

$$\sum_\alpha \mathbf{\Gamma}^{\alpha\alpha}_{\mu\nu\nu\mu} = \sum_\alpha \left[\delta_{\mu\nu}\delta_{\nu\mu} + \delta_{\mu\mu}\delta_{\nu\nu}\right]\left\{A_{\mu\nu}A_{\nu\mu}\mathbf{e}^\alpha_{\mu\nu}\mathbf{e}^\alpha_{\nu\mu} + B_{\mu\nu}B_{\nu\mu}(\delta^{\alpha\alpha}\mathbf{e}_{\mu\nu}\cdot\mathbf{e}_{\nu\mu} - \mathbf{e}^\alpha_{\mu\nu}\mathbf{e}^\alpha_{\nu\mu}\right\} = -A^2_{\mu\nu} - 2B^2_{\mu\nu}$$

$$(17.6.18)$$

Therefore, we have

$$A^2_{\mu\nu} = \frac{1}{k_B}\int_0^\infty dt \, \langle\delta\mathbf{F}_{\mu\nu}\cdot\mathbf{e}_{\mu\nu}\delta\mathbf{F}_{\mu\nu}(t)\cdot\mathbf{e}_{\mu\nu}\rangle^x$$

$$A^2_{\mu\nu} + 2B^2_{\mu\nu} = \frac{1}{k_B}\int_0^\infty dt \, \langle\delta\mathbf{F}_{\mu\nu}\cdot\delta\mathbf{F}_{\mu\nu}(t)\rangle^x$$

$$C^2_{\mu\nu} = \frac{1}{k_B}\int_0^\infty dt \, \langle\delta Q_{\mu\nu}\delta Q_{\mu\nu}(t)\rangle^x \qquad (17.6.19)$$

These expressions provide the microscopic expressions to be used in the noise amplitudes $A_{\mu\nu}, B_{\mu\nu}, C_{\mu\nu}$.

Due to the conditional expectation, the noise amplitudes depend in general on the macrostate x of the system. This means that these transport coefficients are many body functions. The curse of dimensionality makes unpractical the brute force calculation of the functional form of this many body function by perming simulations at different values of the macrostate x. It is obvious that we need to approximate these many body functions and the first option is to substitute the conditional expectation by unconditional, equilibrium averages. While the dependence on the momenta and internal energies in the conditional expectation may be approximated as if they were equilibrium averages, the dependence on the positions of the centers of mass seems to be unavoidable, because the friction between two molecules should vanish as the two molecules become apart. In this way, we approximate the equilibrium conditional average over positions, momenta, and energies with an equilibrium conditional average over positions alone. Further, we will assume that the transport coefficients will depend only on the distance $|\mathbf{R}_\mu - \mathbf{R}_\nu|$ between blobs μ, ν and not on the positions of any other blob. This transform the many body functional form into a simple pair-wise form that depend on the distance of the pair. The transport coefficients (17.6.19) can be computed by running an MD simulation suitably modified for letting the position of the centers of mass of the blobs fixed [?]. By considering all the different pairs $\mu\nu$ of blobs and sorting them by the distance $R_{\mu\nu}$ between them one can extract the dependence on the distance of the different transport coefficients (17.6.19).

17.6.3   The approximate SDE

The assumption that the friction matrix (17.5.1) depends only on the positions and not momenta and internal energies, implies that the stochastic drift $k_B \nabla M(x)$ vanish identically, resulting in a further simplification of the final equations.

We may now introduce the expressions (17.6.17) for the transport coefficients in the different terms of the irreversible part of the dynamics

$$\frac{1}{2}\sum_{\nu\mu'\nu'}^{M} \mathbf{\Gamma}_{\mu\mu'\nu\nu'}^{\alpha\beta} \mathbf{V}_{\nu\nu'}^{\beta}\frac{1}{T_\nu} = \frac{1}{2}\sum_{\nu\mu'\nu'}^{M}[\delta_{\mu\nu}\delta_{\mu'\nu'}+\delta_{\mu\nu'}\delta_{\nu\mu'}]\left\{A_{\mu\mu'}A_{\nu\nu'}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\nu\nu'}^{\beta}+B_{\mu\mu'}B_{\nu\nu'}(\delta^{\alpha\beta}\mathbf{e}_{\mu\mu'}\cdot\mathbf{e}_{\nu\nu'}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\nu\nu'}^{\beta})\right\}\mathbf{V}_{\nu\nu'}^{\beta}\frac{1}{T_\nu}$$

$$=\frac{1}{2}\sum_{\mu'}^{M}\left\{A_{\mu\mu'}A_{\mu\mu'}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu\mu'}^{\beta}+B_{\mu\mu'}B_{\mu\mu'}(\delta^{\alpha\beta}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\mu\mu'}^{\beta}\right\}\mathbf{V}_{\mu\mu'}^{\beta}\frac{1}{T_\mu}$$

$$-\frac{1}{2}\sum_{\mu'}^{M}\left\{A_{\mu\mu'}A_{\mu'\mu}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu'\mu}^{\beta}+B_{\mu\mu'}B_{\mu'\mu}(\delta^{\alpha\beta}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\mu\mu'}^{\beta}\right\}\mathbf{V}_{\mu'\mu}^{\beta}\frac{1}{T_{\mu'}}$$

$$=\sum_{\mu'}^{M}\frac{1}{2}\left[\frac{1}{T_\mu}+\frac{1}{T_{\mu'}}\right]\left\{A_{\mu\mu'}^2\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu\mu'}^{\beta}+B_{\mu\mu'}^2(\delta^{\alpha\beta}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\mu\mu'}^{\beta})\right\}\mathbf{V}_{\mu\mu'}^{\beta}$$

$$=\sum_{\mu'}^{M}\left[\gamma_{\mu\mu'}^{||}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu\mu'}^{\beta}+\gamma_{\mu\mu'}^{\perp}(\delta^{\alpha\beta}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\mu\mu'}^{\beta})\right]\mathbf{V}_{\mu\mu'}^{\beta}$$

$$=\sum_{\mu'}^{M}\left[\gamma_{\mu\mu'}^{||}\mathbf{e}_{\mu\mu'}^{\alpha}(\mathbf{e}_{\mu\mu'}\cdot\mathbf{V}_{\mu\mu'})+\gamma_{\mu\mu'}^{\perp}(\mathbf{V}_{\mu\mu'}^{\alpha}-\mathbf{e}_{\mu\mu'}^{\alpha}(\mathbf{e}_{\mu\mu'}\cdot\mathbf{V}_{\mu\mu'}))\right]$$

$$\sum_{\nu\mu'\nu'}^{M}\sigma_{\mu\mu'\nu\nu'}\frac{1}{T_\nu}=\sum_{\nu\mu'\nu'}^{M}[\delta_{\mu\nu}\delta_{\mu'\nu'}+\delta_{\mu\nu'}\delta_{\nu\mu'}]C_{\mu\mu'}C_{\nu\nu'}\frac{1}{T_\nu}=\sum_{\mu'}^{M}C_{\mu\mu'}^2\left[\frac{1}{T_\mu}-\frac{1}{T_{\mu'}}\right]=\sum_{\mu'}^{M}\kappa_{\mu\mu'}\left[T_{\mu'}-T_\mu\right]$$

$$\frac{1}{4}\sum_{\nu\mu'\nu'}^{M}\mathbf{v}_{\mu\mu'}^{\alpha}\mathbf{\Gamma}_{\mu\mu'\nu\nu'}^{\alpha\beta}\mathbf{v}_{\nu\nu'}^{\beta}\frac{1}{T_\nu}=\frac{1}{4}\sum_{\nu\mu'\nu'}^{M}\mathbf{V}_{\mu\mu'}\cdot[\delta_{\mu\nu}\delta_{\mu'\nu'}+\delta_{\mu\nu'}\delta_{\nu\mu'}]\left\{A_{\mu\mu'}A_{\nu\nu'}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\nu\nu'}^{\beta}+B_{\mu\mu'}B_{\nu\nu'}(\delta^{\alpha\beta}\mathbf{e}_{\mu\mu'}\cdot\mathbf{e}_{\nu\nu'}-\mathbf{e}_{\mu\mu'}^{\beta}\mathbf{e}_{\nu\nu'}^{\beta}\right\}\cdot\mathbf{V}_{\nu\nu'}$$

$$=\frac{1}{2}\sum_{\mu'}^{M}\mathbf{V}_{\mu\mu'}^{\alpha}\left\{\gamma_{\mu\mu'}^{||}\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu\mu'}^{\beta}+\gamma_{\mu\mu'}^{\perp}(\delta^{\alpha\beta}-\mathbf{e}_{\mu\mu'}^{\alpha}\mathbf{e}_{\mu\mu'}^{\beta}\right\}\mathbf{V}_{\mu\mu'}^{\beta}$$

$$=\frac{1}{2}\sum_{\mu'}^{M}\left\{\gamma_{\mu\mu'}^{||}(\mathbf{e}_{\mu\mu'}\cdot\mathbf{V}_{\mu\mu'})^2+\gamma_{\mu\mu'}^{\perp}(\mathbf{V}_{\mu\mu'}^2-(\mathbf{e}_{\mu\mu'}\cdot\mathbf{V}_{\mu\mu'})^2)\right\}$$

$$(17.6.20)$$

we have introduced the friction coefficients

$$\gamma_{\mu\mu'}^{||} = \frac{1}{2} \left[\frac{1}{T_\mu} + \frac{1}{T_{\mu'}} \right] A_{\mu\mu'}^2$$

$$\gamma_{\mu\mu'}^{\perp} = \frac{1}{2} \left[\frac{1}{T_\mu} + \frac{1}{T_{\mu'}} \right] B_{\mu\mu'}^2 \qquad (17.6.21)$$

and the thermal transfer coefficient

$$\kappa_{\mu\nu} \equiv \frac{C_{\mu\mu'}^2}{T_\mu T_{\mu'}} \qquad (17.6.22)$$

If we now insert these models for the friction matrix elements, we have the following dynamic equations

$$\dot{\mathbf{R}}_\mu = \mathbf{V}_\mu$$

$$\dot{\mathbf{P}}_\mu = \sum_\nu^M \mathbf{F}_{\mu\nu} - \sum_{\mu'}^M \left[\gamma_{\mu\mu'}^{||} \mathbf{e}_{\mu\mu'} (\mathbf{e}_{\mu\mu'} \cdot \mathbf{V}_{\mu\mu'}) + \gamma_{\mu\mu'}^{\perp} (\mathbf{V}_{\mu\mu'}^\alpha - \mathbf{e}_{\mu\mu'} (\mathbf{e}_{\mu\mu'} \cdot \mathbf{V}_{\mu\mu'})) \right] + \sum_\nu \tilde{\mathbf{F}}_{\mu\nu}$$

$$\dot{\mathscr{E}}_\mu = -\frac{1}{2} \sum_\nu^M \mathbf{F}_{\mu\nu} \cdot \mathbf{V}_{\mu\nu} + \sum_{\mu'}^M \kappa_{\mu\mu'} \left[T_{\mu'} - T_\mu \right]$$

$$+ \frac{1}{2} \sum_{\mu'}^M \left\{ \gamma_{\mu\mu'}^{||} (\mathbf{e}_{\mu\mu'} \cdot \mathbf{V}_{\mu\mu'})^2 + \gamma_{\mu\mu'}^{\perp} (\mathbf{V}_{\mu\mu'}^2 - (\mathbf{e}_{\mu\mu'} \cdot \mathbf{V}_{\mu\mu'})^2) \right\} + \sum_\nu \left[\tilde{Q}_{\mu\nu} - \tilde{\mathbf{F}}_{\mu\nu} \cdot \mathbf{V}_{\mu\nu} \right]$$

$$(17.6.23)$$

These are the final SDE that govern the dynamics of the position, momentum and internal energy of the blobs and the main result of this paper. They are obviously Galilean invariant. They conserve exactly the total momentum (17.2.7) and total energy (17.2.6) of the system (one needs to use Ito calculus, of course). Total angular momentum is not conserved though, because the friction forces are not central. It has been claimed [?] that inclusion of a spin variable as suggested earlier [?] can produce better results. We plan to derive the equations for the level of description given by the position, momenta, internal energy, and spin, of complex molecules in the future.

The physical meaning of the different terms in the equations is rather transparent. For example, in the momentum equation, there are reversible forces due to the mean force $\mathbf{F}_{\mu\nu}$ and also friction forces that have two components, one which is central and another one that goes in the normal direction and produces shear forces. These friction forces reduce any velocity difference between complex molecules. The kinetic energy that is lost due to the stopping effect of friction is transformed into internal energy through the terms which are quadratic in the velocity in the equation for the internal energy. These term reflect, therefore, as a friction heating effect, analogous to the viscous heating in the hydrodynamic equations. Looking at the internal energy equation, the first term can be interpreted as the reversible work per unit time done on the blobs by the mean forces. The term proportional to the temperature difference is a heat conduction term that tries to reduce temperature differences between blobs. Finally, there are stochastic

terms, both in the momentum equation and in the internal energy equation that keep the variables of this level of description in continuous agitation due to the eliminated degrees of freedom.

The structure of the equations is similar to the energy conserving Dissipative Particle Dynamics model[?],[?], which was introduced on purely phenomenological grounds. In the present derivation, all the different terms in the equations have a microscopic definition that allows for its explicit calculation through molecular dynamic simulations. The equations and methodology presented should have use in the coarse-grained non-isothermal descriptions of complex molecules.

18

Falling apples and spinning planets

Macroscopic bodies like falling apples and planets are made of an enormous number of particles. Nevertheless, one speaks of the trajectory of the apple and the orbit of the planet, as if some mathematical curves where traversed by a single point. How can this be? From what we have seen so far, we should treat these macroscopic bodies in a Statistical Mechanics sense, and apply the theory of coarse-graining to them. In this chapter we discuss the application of the ToCG, which is based on Classical Mechanics for point particles, to Classical Mechanics of macroscopic bodies. In this doing, we will obtain the conditions under which Classical Mechanics could be applied to macroscopic bodies and treat them as point particles.

18.1   A falling apple

From standard textbooks on mechanics, we know that there is a privileged point in bodies to which the laws of mechanics apply. This point is the center of mass. It seems therefore, that the appropriate relevant variable is the position and momentum of the center of mass.

The time independent Hamiltonian of the apple is, near the earth surface

$$\hat{H}(z) = \sum_i^N \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1, \cdots, \mathbf{r}_N) + \sum_i^N m\mathbf{g}\cdot\mathbf{r}_i \qquad (18.1.1)$$

where $V(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ is the potential of interaction of the atoms of the apple binding them together and the last term is the potential of gravitation, where $\mathbf{g}$ is the acceleration of gravity (which is directed in the vertical direction

We chose as CG variables the total energy of the system $\hat{H}(z)$ and the position $\mathbf{R}$ and momenta $\mathbf{P}$ of the CoM. We choose the probabilistic flavor, in which we will ask for the probability of these CG variables

The relevant ensemble in the probabilistic flavor is

$$\overline{\rho}_t(z) = \frac{P(\mathbf{R}, \mathbf{P}, E}{\Omega((\mathbf{R}, \mathbf{P}, E} \tag{18.1.2}$$

The SDE equation for the CoM is

$$d\mathbf{R}(t) = \frac{\mathbf{P}}{M}$$
$$d\mathbf{P}(t) = \left\langle \hat{\mathbf{F}} \right\rangle^{\mathbf{R}, \mathbf{P}} + m\mathbf{g} \tag{18.1.3}$$

18.2   Spinning planets

Consider two large macroscopic bodies like two planets or a binary star system that is self-gravitating in a region of space far away from other masses in such a way that they may be regarded as an isolated system. From a conceptual point of view this problem is somewhat challenging for the theory of coarse-graining as regarding to the idea of equilibrium state.

Gravitation does not have ordinary equilibrium states

Think of a region of space containing many particles homogeneously distributed and interacting with gravity. Because gravity is attractive, we expect that the system will shrink. It does but, in general, not by keeping the density homogeneous, but generating inhomogeneities that may eventually form accretion disks and generate galaxies and planetary systems. The long term fate of a gravitational system is not what we have in mind when we think of a system "equilibrating until reaching a state of rest". In fact, it is well-known that a system of many particles interacting through *long range* interactions (i.e. potential functions decaying as the inverse of the distance) cannot have an equilibrium state [?].

A stable planetary system cannot be taken as an equilibrium system because, for example, the density field has a very peculiar form, concentrated in spherical regions (the planets) that move following orbits around the star of the planetary system. It is a non-homogeneous time- dependend field, very different from what we expect of a system at equilibrium. Nevertheless, the matter in the planet itself, which is interacting with short range potentials do has a well defined equilibrium state.

The problem we want to address in this section is that of the coupling of the spin between to self-gravitating bodies.

[Perhaps this goes into the chapter on Phase transitions] While the technical details are complicated, the intuitive reason for the non-existence of an equilibrium state can be understood by looking at a level of description given by the density field. Microscopically, the potential of interaction between the particles that form the system has two contributions. One due to the cohesive, non-interpenetrating potential and another one due to the gravitational interaction with the rest of particles in the system. The former avoids that a planet collapses onto itself due to its own weight and prevents two particles to be on top of each other, thus avoiding the singularity of the gravitational

potential at short distances. The Hamiltonian of the system can be expressed as

$$\hat{H}(z) = \sum_i^N \frac{\mathbf{p}_i^2}{2m_i} + V^{\text{short}}(\mathbf{r}_1, \cdots, \mathbf{r}_N) - \frac{1}{2} \sum_{ij} \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{18.2.1}$$

where V^{short} is the cohesive, non-interpenetrating potential of short range nature and the last term is the gravitational potential. This last term may be written in terms of the number density field $\hat{n}_\mathbf{r} = \sum_i^N \delta(\mathbf{r} - \mathbf{r}_i)$ as

$$-\frac{1}{2} \sum_{ij} \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} = \int d\mathbf{r} \frac{Gm_i m_j}{|\mathbf{r} - \mathbf{r}'|} \hat{n}_\mathbf{r}(z) \hat{n}_{\mathbf{r}'}(z) \tag{18.2.2}$$

Because the gravitational part of the Hamiltonian is expressed in terms of a relevant variable, this implies that the entropy of the system will take the form

$$S[E, n] = S^{\text{short}}[E, n] + \int d\mathbf{r} \frac{Gm_i m_j}{|\mathbf{r} - \mathbf{r}'|} \hat{n}_\mathbf{r}(z) \hat{n}_{\mathbf{r}'}(z) \tag{18.2.3}$$

where $S^{\text{short}}[E, n]$ is the entropy of the system when the gravitational interactions are switched off. For the sake of the argument, we may stick to the simplest model for the short range which is the local model

$$S[E, n] = \int d\mathbf{r} f(E, n(\mathbf{r}) + \beta E + \int d\mathbf{r} \frac{Gm_i m_j}{|\mathbf{r} - \mathbf{r}'|} \hat{n}(\mathbf{r}) \hat{n}(\mathbf{r}') \tag{18.2.4}$$

where $f(E, n)$ is the local free energy and β is the inverse temperature (which is a function of E and a functional of $n(\mathbf{r})$.

Then, we ask to ourselves, is a constant density profile possible for a self-gravitating system of particles? The answer is no

19

Discrete diffusion: $\hat{n}_\mu$

20

Discrete Non-isothermal hydrodynamics: $\hat{\rho}_\mu, \hat{\mathbf{g}}_\mu, \hat{e}_\mu$

Up to now we have derived a number of *field theories* for the transport properties of fluids and solids. These theories are characterized by the fact that the resulting dynamic equations are partial differential equations. These PDE's govern the evolution of the *average* value of fields that are defined, at the microscopic level, through the Dirac delta function. The prototype of these fields is given by the density field

$$\hat{\rho}_{\mathbf{r}}(z) = \sum_{i=1}^{N} m_i \delta(\mathbf{r} - \mathbf{q}_i) \tag{20.0.1}$$

We may be as well interested on not only the average value of the field variables but also on their *fluctuations*. Fluctuations are important in fluids, as they are the responsible for the blue color of the sky, due to light scattering, and they are particularly important near critical points. Also, they become important at length scales below the micron, and then they are the responsible for the diffusion of suspended particles.

In order to describe fluctuations we should take one of the following two paths. Either we introduce covariances of fields, and construct their dynamic equations by following identical procedures as we have done in the derivation of transport equations for the averages (at the end covariances are just averages of products of fields). This results in *mode coupling* theories. The second route is to look for the probability distribution function of the fields themselves. This is tantamount of looking at the probabilistic flavor of field theories, i.e. at the Fokker-Planck equation governing the probability of the hydrodynamic fields. Because to the Fokker-Planck description it corresponds a Stochastic Differential Equation, the resulting field theory with fluctuations will be given by a Stochastic Partial Differential Equation. Perhaps the most notable example of such a theory is the Landau and Lifshitz theory of hydrodynamic fluctuations. In this theory, to the usual stress and heat flux, additional stochastic contributions are introduced. However, a little of reflection shows that the Landau-Lifshitz theory is really the result of an abuse of notation.

We cannot use the probabilistic flavor for the delta defined hydrodynamic variables. This implies that we cannot answer questions about fluctuations of hydrodynamic variables.

How one can then deal with fluctuations in fluids? It turns out that one has to abandon the notion of continuum fields and opt for a definition of the hydrodynamic variables which are defined for regions of space with finite extension. It should be mention that which makes use of continuum hydrodynamic fields obeying partial differential equations should be taken as an abuse of notation. The "fields"

We aim in the following chapters at a consistent description of the coarse-graining of a fluid that permits to treat fluctuations in a proper manner.

21

Boundary conditions in Fluctuating Hydrodynamics

22

Discrete Isothermal Kinetic theory, Lattice Boltzmann: $f_{\mu\nu}$

Part III

Appendices

23

Gaussian integrals

23.0.1   CG probability

Consider a probability distributions of the form

$$\Omega(y, J) = \frac{1}{\Omega_0(J)} \int dx \exp\left\{-\frac{1}{2}x^T Ax + J^T x\right\} \delta(\Pi x - y) \qquad (23.0.1)$$

where $x \in \mathbb{R}^N$, $y \in \mathbb{R}^M$, $A \in \mathbb{R}^{N \times N}$ is invertible, $\Pi \in \mathbb{R}^{M \times N}$ is rectangular matrix, and $J \in \mathbb{R}^N$. The normalization factor is given by Eq. (??). We start first with the numerator integral and we introduce the Fourier transform of the Dirac delta function in order to compute

$$\int dx \exp\left\{-\frac{1}{2}x^T Ax + J^T x\right\} \delta(\Pi x - y)$$

$$= \int \frac{d\lambda}{(2\pi)^M} \exp\{-i\lambda^T y\} \int dx \exp\left\{-\frac{1}{2}x^T Ax + J^T x\right\} \exp\{i\lambda^T \Pi x\}$$

$$= \int \frac{d\lambda}{(2\pi)^M} \exp\{-i\lambda y\} \int dx \exp\left\{-\frac{1}{2}x^T Ax + (J + i\Pi^T \lambda)^T x\right\}$$

$$= \int \frac{d\lambda}{(2\pi)^M} \exp\{-i\lambda y\} \frac{(2\pi)^{N/2}}{\det A^{1/2}} \exp\left\{\frac{1}{2}(J + i\Pi^T \lambda)^T A^{-1}(J + i\Pi^T \lambda)\right\}$$

$$= \frac{(2\pi)^{N/2}}{\det A^{1/2}} \int \frac{d\lambda}{(2\pi)^M} \exp\{-i\lambda y\} \exp\left\{\frac{1}{2}(J + i\Pi^T \lambda)^T A^{-1}(J + i\Pi^T \lambda)\right\}$$

$$= \frac{(2\pi)^{N/2}}{\det A^{1/2}} \int \frac{d\lambda}{(2\pi)^M} \exp\{-i\lambda(y + \Pi A^{-1}J)\} \exp\left\{-\frac{1}{2}\lambda^T \Pi A^{-1}\Pi^T \lambda + \frac{1}{2}J^T A^{-1}J\right\}$$

$$= \frac{(2\pi)^{N/2}}{\det A^{1/2}} \frac{\det \overline{A}^{1/2}}{(2\pi)^{M/2}} \exp\left\{-\frac{1}{2}(y + \Pi A^{-1}J)^T \overline{A}(y + \Pi A^{-1}J) + \frac{1}{2}J^T A^{-1}J\right\} \qquad (23.0.2)$$

where we have introduced the *coarse-grained inverse matrix* $\overline{A} \in \mathbb{R}^{M \times M}$ as

$$\overline{A} \equiv \left[\Pi A^{-1} \Pi^T \right]^{-1} \tag{23.0.3}$$

provided that the inverse exists. For future reference, consider the eigenvectors $\overline{v}_\mu$ of the matrix $\overline{A}^{-1} = \Pi A^{-1} \Pi^T$, this is $\overline{A}^{-1} \overline{v}_\mu = \overline{d}_\mu \overline{v}_\mu$. We may write the matrix $\overline{A}$ as

$$\overline{A} = \sum_{\mu=1}^{M} \overline{v}_\mu \frac{1}{\overline{d}_\mu} \overline{v}_\mu^T \tag{23.0.4}$$

Finally, the probability (23.0.1) is given by

$$\Omega(y, J) = \frac{\det \overline{A}^{1/2}}{(2\pi)^{M/2}} \exp \left\{ -\frac{1}{2} (y + \Pi A^{-1} J)^T \overline{A} (y + \Pi A^{-1} J) \right\} \tag{23.0.5}$$

24

SDE for pedestrians

Supongamos que el estado macroscópico de un sistema está descrito por un vector $\mathbf{x} \in \mathbb{R}^M$ y que la dinámica de este vector está descrito por un proceso difusivo que obedece la ecuación de Fokker-Planck

$$\partial_t P(\mathbf{x}, t) = -\frac{\partial}{\partial \mathbf{x}_i} \mathbf{A}_i(\mathbf{x}) P(\mathbf{x}, t) + \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_i} \frac{\partial}{\partial \mathbf{x}_j} \mathbf{D}_{ij}(\mathbf{x}) P(\mathbf{x}, t)$$

$$(24.0.1)$$

La ecuación de Fokker-Planck gobierna la densidad de probabilidad $P(\mathbf{x}, t)$ de que una variable estocástica $\mathbf{x}$ tome valores en una vecindad infinitesimal de $\mathbf{x}$. Es posible demostrar que también la densidad de probabilidad condicional $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ de que la variable estocástica tome valores entre $\mathbf{x}$ y $\mathbf{x} + d\mathbf{x}$ en el tiempo t *dado* que en el tiempo t_0 tomó el valor $\mathbf{x}_0$, obedece la misma ecuación diferencial [1], es decir

$$\partial_t P(\mathbf{x}, t | \mathbf{x}_0, t_0) = -\frac{\partial}{\partial \mathbf{x}_i} \mathbf{A}_i(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) + \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_i} \frac{\partial}{\partial \mathbf{x}_j} \mathbf{D}_{ij}(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0)$$

$$(24.0.2)$$

Es evidente que la probabilidad condicional debe cumplir que

$$P(\mathbf{x}, t_0 | \mathbf{x}_0, t_0) = \delta(\mathbf{x} - \mathbf{x}_0) \tag{24.0.3}$$

ya que la probabilidad de que en el instante inicial el estado sea $\mathbf{x}$ es nula a menos que $\mathbf{x}$ sea precisamente el estado inicial, $\mathbf{x} = \mathbf{x}_0$.

24.1 Solución de la ecuación de Fokker-Planck con coeficientes constantes

Vamos a resolver la siguiente ecuación de Fokker-Planck con coeficientes constantes

$$\partial_t P(\mathbf{x}, t | \mathbf{x}_0, t_0) = -\mathbf{A}_i(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{x}_i} P(\mathbf{x}, t | \mathbf{x}_0, t_0) + \frac{1}{2} \mathbf{D}_{ij}(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{x}_i} \frac{\partial}{\partial \mathbf{x}_j} P(\mathbf{x}, t | \mathbf{x}_0, t_0)$$

$$(24.1.1)$$

Consideremos la transformada de Fourier

$$P(\mathbf{k}, t | \mathbf{x}_0, t_0) = \int d\mathbf{x} \exp\{-i\mathbf{k} \cdot \mathbf{x}\} P(\mathbf{x}, t | \mathbf{x}_0, t_0)$$

$$P(\mathbf{x}, t | \mathbf{x}_0, t_0) = \frac{1}{(2\pi)^M} \int d\mathbf{x} \exp\{i\mathbf{k} \cdot \mathbf{x}\} P(\mathbf{k}, t | \mathbf{x}_0, t_0)$$

$$(24.1.2)$$

La ecuación para las componentes de Fourier es

$$\partial_t P(\mathbf{k}, t | \mathbf{x}_0, t_0) = -i\mathbf{A}(\mathbf{x}_0) \cdot \mathbf{k} P(\mathbf{k}, t) - \frac{1}{2} \mathbf{k}^T \cdot \mathbf{D}(\mathbf{x}_0) \cdot \mathbf{k} P(\mathbf{k}, t | \mathbf{x}_0, t_0)$$

$$(24.1.3)$$

que tiene por solución

$$P(\mathbf{k}, t | \mathbf{x}_0, t_0) = \exp t \left\{ -i\mathbf{A}(\mathbf{x}_0) \cdot \mathbf{k} - \frac{1}{2} \mathbf{k}^T \cdot \mathbf{D}(\mathbf{x}_0) \cdot \mathbf{k} \right\} P(\mathbf{k}, t_0 | \mathbf{x}_0, t_0) \qquad (24.1.4)$$

Así, que volviendo a las variables originales

$$P(\mathbf{x}, t | \mathbf{x}_0, t_0) = \frac{1}{(2\pi)^M} \int d\mathbf{k} \exp\{i\mathbf{k} \cdot \mathbf{x}\} \exp t \left\{ -i\mathbf{A}(\mathbf{x}_0) \cdot \mathbf{k} - \frac{1}{2} \mathbf{k}^T \cdot \mathbf{D}(\mathbf{x}_0) \cdot \mathbf{k} \right\} P(\mathbf{k}, 0 | \mathbf{x}_0, t_0)$$

$$(24.1.5)$$

Si inicialmente $P(\mathbf{x}, 0 | \mathbf{x}_0, t_0) = \delta(\mathbf{x} - \mathbf{x}_0)$, eso significa que $P(\mathbf{k}, 0 | \mathbf{x}_0, t_0) = \exp\{-i\mathbf{k} \cdot \mathbf{x}_0\}$ y, por tanto,

$$P(\mathbf{x}, t | \mathbf{x}_0, t_0) = \frac{1}{(2\pi)^M} \int d\mathbf{k} \exp\{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0)\} \exp t \left\{ -i\mathbf{A}(\mathbf{x}_0) \cdot \mathbf{k} - \frac{1}{2} \mathbf{k}^T \cdot \mathbf{D}(\mathbf{x}_0) \cdot \mathbf{k} \right\}$$

$$= \frac{1}{(2\pi)^M} \int d\mathbf{k} \exp\{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{x}_0 - t\mathbf{A}(\mathbf{x}_0))\} \exp \left\{ -\frac{t}{2} \mathbf{k}^T \cdot \mathbf{D}(\mathbf{x}_0) \cdot \mathbf{k} \right\}$$

$$(24.1.6)$$

Cuando la inversa de $\mathbf{D}$ existe, tenemos la solución explícita

$$
\begin{aligned}
P(\mathbf{x}, t|\mathbf{x}_0, t_0) &= \exp\left\{-\frac{1}{2t}(\mathbf{x} - \mathbf{x}_0 - t\mathbf{A}(\mathbf{x}_0))^T\mathbf{D}^{-1}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0 - t\mathbf{A}(\mathbf{x}_0))\right\}\frac{1}{(2\pi)^M}\int d\mathbf{k}\exp\left\{-\frac{t}{2}\mathbf{k}^T\cdot\mathbf{D}(\mathbf{x}_0)\cdot\mathbf{k}\right\} \\
&= \exp\left\{-\frac{1}{2t}(\mathbf{x} - \mathbf{x}_0 - t\mathbf{A}(\mathbf{x}_0))^T\mathbf{D}^{-1}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0 - t\mathbf{A}(\mathbf{x}_0))\right\}\frac{1}{(2\pi t)^{M/2}}\frac{1}{(\det\mathbf{D}(\mathbf{x}_0))^{1/2}}
\end{aligned}
$$

$$(24.1.7)$$

24.2 Ecuación diferencial estocástica

Ahora consideremos la siguiente equación en diferencias finitas

$$\mathbf{x}' = \mathbf{x} + \mathbf{A}(\mathbf{x})\Delta t + \mathbf{B}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta} \qquad (24.2.1)$$

donde $\mathbf{A}(\mathbf{x})$ es una función vectorial de $\mathbb{R}^M \to \mathbb{R}^M$ y $\mathbf{B}(\mathbf{x})$ es una matriz $\mathbb{R}^M \times \mathbb{R}^M$. Aquí, $\boldsymbol{\zeta} \in \mathbb{R}^M$ es una vector de variables aleatorias gausianas de media nula y varianza unidad. Es decir,

$$
\begin{aligned}
\langle\boldsymbol{\zeta}_k\rangle &= 0 \\
\langle\boldsymbol{\zeta}_k\boldsymbol{\zeta}_{k'}\rangle &= \delta_{kk'}
\end{aligned}
\qquad (24.2.2)
$$

La densidad de probabilidad de $\boldsymbol{\zeta}$ es [2]

$$G(\boldsymbol{\zeta}) = \frac{1}{(2\pi)^{M/2}}\exp\left\{-\frac{1}{2}\boldsymbol{\zeta}^T\cdot\boldsymbol{\zeta}\right\} \qquad (24.2.3)$$

Cuando decimos que $\boldsymbol{\zeta}$ es independiente, estamos presuponiendo que las variables $\mathbf{x}', \mathbf{x}$ y $\boldsymbol{\zeta}$ son estadísticamente independientes.

La ecuación (24.2.1) debe entenderse de la siguiente forma. Suponemos que el valor del estado $\mathbf{x}$ en un instante inicial es conocido con certeza. Entonces (24.2.1) nos da el valor $\mathbf{x}'$ en un instante posterior Δt. Como $\boldsymbol{\zeta}$ es un número aleatorio, $\mathbf{x}'$ también lo será. Nos interesa calcular la densidad de probabilidad $P(\mathbf{x}', \Delta t|\mathbf{x})$ de esta variables $\mathbf{x}'$, que será una probabilidad condicional sujeta a que conocemos $\mathbf{x}$. Para calcular esta probabilidad, podemos entender la ecuación (24.2.1) como que nos da $\mathbf{x}'$ como una función (lineal) de $\boldsymbol{\zeta}$. De la sección **??** sabemos cómo se relaciona la densidad de probabilidad de una función de una variable estocástica con la densidad de probabilidad de dicha variable estocástica, que es a través de la (**??**). En nuestro caso esta ecuación toma la forma

$$P(\mathbf{x}', \Delta t|\mathbf{x}) = \int d\boldsymbol{\zeta}\, G(\boldsymbol{\zeta})\delta\left(\mathbf{x}' - \mathbf{x} - \mathbf{A}(\mathbf{x})\Delta t - \mathbf{B}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}\right) \qquad (24.2.4)$$

Si hacemos un cambio de variable $\mathbf{y} = \mathbf{B}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}$. Usando la propiedad (**??**) podemos

hacer la integral explícitamente y el resultado es [3]

$$P(\mathbf{x}', \Delta t|\mathbf{x}) = \frac{1}{\det \mathbf{B}(\mathbf{x})} \int d\mathbf{y} \frac{1}{(2\pi)^{M/2}} \exp\left\{-\frac{1}{2\Delta t}\mathbf{y}^T \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \mathbf{y}\right\} \delta\left(\mathbf{x}' - \mathbf{x} - \mathbf{A}(\mathbf{x})\Delta t - \mathbf{y}\right)$$

$$= \frac{1}{(2\pi\Delta t)^{M/2}(\det \mathbf{D}(\mathbf{x}))^{1/2}}$$

$$\times \exp\left\{-\frac{1}{2}\left(\frac{\mathbf{x}'-\mathbf{x}}{\Delta t} - \mathbf{A}(\mathbf{x})\right)^T \cdot \mathbf{D}^{-1}(\mathbf{x}) \cdot \left(\frac{\mathbf{x}'-\mathbf{x}}{\Delta t} - \mathbf{A}(\mathbf{x})\right)\right\} \tag{24.2.5}$$

donde hemos definido la matrix $\mathbf{D}(\mathbf{x}) = \mathbf{B}^T(\mathbf{x})\mathbf{B}(\mathbf{x})$ que es una matriz $M \times M$. En componentes tenemos

$$\mathbf{D}_{ij}(\mathbf{x}) = \mathbf{B}_{ik}(\mathbf{x})\mathbf{B}_{jk}(\mathbf{x}) \tag{24.2.6}$$

con suma sobre índices repetidos.

Notemos que la probabilidad condicional (24.2.5) es idéntica a (24.1.7). De esta manera, vemos que la variable aleatoria $\mathbf{x}'$ está distribuida de acuerdo con la probabilidad condicional que es solución de la ecuación de Fokker-Planck. Por tanto, lanzando muchas realizaciones de la variable aleatoria gausiana $\boldsymbol{\zeta}$ podemos generar un montón de valores de $\mathbf{x}'$ cuyo histograma está dado precisamente por $P(\mathbf{x}', \Delta t|\mathbf{x})$. De esta manera, tenemos un método para "resolver" la ecuación de Fokker-Planck a través de la generación de números aleatorios gausianos.

24.3 Cálculo de Ito

De nuevo, sin excesivo rigor, vamos a introducir el cálculo de Ito. Imaginemos que nos preguntamos cuanto cambia una función $F(\mathbf{x}), \mathbb{R}^M \to \mathbb{R}$ al hacer una actualización de $\mathbf{x}$ de acuerdo con la ecuación en diferencias finitas (24.2.1). La variación estará dada por

$$\Delta F(\mathbf{x}) = F(\mathbf{x}') - F(\mathbf{x})$$

$$= F(\mathbf{x} + \mathbf{A}(\mathbf{x})\Delta t + \mathbf{B}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}) - F(\mathbf{x}) \tag{24.3.1}$$

Si suponemos que la función es diferenciable, podemos expander en serie de Taylor alrededor de $\mathbf{x}$ hasta segundo orden (veremos porqué hasta segundo orden). Obtendremos

$$\Delta F(\mathbf{x}) = \frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i} \left(\mathbf{A}_i(\mathbf{x})\Delta t + \mathbf{B}_{ij}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}_j \right)$$

$$+ \frac{1}{2}\frac{\partial^2 F(\mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \left(\mathbf{A}_i(\mathbf{x})\Delta t + \mathbf{B}_{ik}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}_k \right) \left(\mathbf{A}_i(\mathbf{x})\Delta t + \mathbf{B}_{jk'}(\mathbf{x})\Delta t^{1/2}\boldsymbol{\zeta}_{k'} \right)$$

$$= \frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i}\mathbf{B}_{ij}(\mathbf{x})\boldsymbol{\zeta}_j\Delta t^{1/2}$$

$$+ \left(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i}\mathbf{A}_i(\mathbf{x}) + \frac{1}{2}\frac{\partial^2 F(\mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\mathbf{B}_{ik}(\mathbf{x})\boldsymbol{\zeta}_k\mathbf{B}_{jk'}(\mathbf{x})\boldsymbol{\zeta}_{k'} \right)\Delta t + \mathscr{O}(\Delta t^{3/2}) \quad (24.3.2)$$

donde hemos agrupado los distintos términos en orden creciente en potencias de Δt.

Está claro que como $\boldsymbol{\zeta}$ es un vector aleatorio, $\Delta F(\mathbf{x})$ también será una variable aleatoria. Podemos calcular, por tanto, su valor medio que está dado por

$$\langle \Delta F(\mathbf{x}) \rangle = \left(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i}\mathbf{A}_i(\mathbf{x}) + \frac{1}{2}\frac{\partial^2 F(\mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\mathbf{D}_{ij}(\mathbf{x}) \right)\Delta t + \mathscr{O}(\Delta t^{3/2}) \quad (24.3.3)$$

donde hemos usado (24.2.2). Vemos, por tanto, que a la hora de calcular valores medios de las variaciones de una función, por el hecho de que la ecuación en diferencias (24.2.1) va acompañada de términos de orden $\Delta t^{1/2}$, nos obliga a considerar derivadas *segundas* de la función incluso para el término de *primer* orden en Δt!

Una ecuación diferencial estocástica se puede pensar como el límite cuando $\Delta t \to 0$ de la ecuación en diferencias finitas (24.2.1). En este caso, se escribe dicha ecuación en la forma [?]

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{x}_t)dt + \mathbf{B}(\mathbf{x}_t)d\mathbf{W}_t \quad (24.3.4)$$

Aquí, $d\mathbf{W}_t$ se denomina el incremento independiente del proceso de Wiener, que esencialmente es un número aleatorio gausiano de media nula y varianza dt. Para calcular variaciones de funciones $F(\mathbf{x})$ (es decir, la diferencial $dF(\mathbf{x})$), el cálculo de Ito deviene

$$dF(\mathbf{x}) = \left(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i}\mathbf{A}_i(\mathbf{x}) + \frac{1}{2}\frac{\partial^2 F(\mathbf{x})}{\partial \mathbf{x}_i \partial \mathbf{x}_j}\mathbf{D}_{ij}(\mathbf{x}) \right)dt + \frac{\partial F(\mathbf{x})}{\partial \mathbf{x}_i}\mathbf{B}_{ij}(\mathbf{x})d\mathbf{W}_t \quad (24.3.5)$$

que se diferencia del cálculo ordinario en que hay que expandir la función hasta segundo orden y usar la regla mnemotécnica

$$dW_t dW_t = dt \quad (24.3.6)$$

25

Change of variables in the FPE

In some applications it is useful to express the dynamics in terms of new variables which are not the CG variables that have been used to construct the FPE. For example, we may want to express the dynamics in terms of temperatures instead of energies. While it is not possible to derive directly from Hamilton's equations a FPE for the temperatures, because the temperature is not a phase function, the dynamics looks sometimes simpler in terms of the temperature. Because the temperature is a function of the energy, we may use then a change of variables. In this section, we consider how the FPE transforms under a change of variables.

26

Functionals, functional derivatives, conditioned extrema

En esta sección presentamos un resumen sobre funcionales. Aunque no es matemáticamente muy riguroso, en la práctica esto es todo lo que uno necesita saber para salir del paso.

Un funcional I, es una aplicación del conjunto de las funciones, C^∞ en $\mathbb{R}$, es decir, un funcional asocia a cada función un número.

- **Ejemplo 1:** la integral de una función es un funcional $I_{[a,b]}$

$$I_{[a,b]} : C^\infty \longrightarrow \qquad\qquad\qquad\qquad\qquad \mathbb{R}$$
$$f \longrightarrow \qquad\qquad I_{[a,b]}[f] = \int_a^b f(x)dx \qquad\qquad (26.0.1)$$

  Este es un funcional que depende de dos parámetros a, b. El argumento de los funcionales (la función) se suele escribir entre corchetes cuadrados, $I[f]$.

- **Ejemplo 2:** Otro ejemplo un poco más sutil de funcional es aquel funcional que asocia a una función dada $f(x)$ el valor que toma esa función en un punto particular x_0, es decir

$$I_{x_0} : C^\infty \longrightarrow \qquad\qquad\qquad\qquad \mathbb{R}$$
$$f \longrightarrow \qquad\qquad I_{x_0}[f] = f(x_0) \qquad\qquad (26.0.2)$$

  Está claro que este último funcional se puede escribir $I_{x_0}[f] = \int \delta(x - x_0)f(x)dx$. Así, una función $f(x)$ se puede pensar como un funcional o, más propiamente, como una familia de funcionales parametrizada por la variable argumento de la función.

- **Ejemplo 3:** Otro ejemplo más es la composición de funciones, $I_x^g[f] = g(f(x))$

donde $g(x)$ es una función particular. De nuevo podemos pensar la composición de funciones como una familia de funcionales parametrizada por x.

26.1   La derivada funcional

Intuitivamente, podemos pensar una función como un "vector de infinitas componentes" (el valor de cada componente es el valor de la función en cada punto del espacio) y un funcional, como una "función" de este "vector de infinitas componentes". Tiene sentido, por tanto, preguntarse por las derivadas de esta "función". En este sentido, la *derivada funcional* es la generalización de la derivada parcial. La derivada parcial de una función $\mathcal{I}(x)$

$$
\begin{aligned}
\mathcal{I} : \mathbb{R}^n &\longrightarrow & \mathbb{R} \\
(x_1, \cdots, x_n) &\longrightarrow & \mathcal{I}(x_1, \cdots, x_n) \\
\{x_i\} &\longrightarrow & \mathcal{I}(\{x_i\})
\end{aligned}
\tag{26.1.1}
$$

se define como

$$
\begin{aligned}
\frac{\partial \mathcal{I}}{\partial x_k} &= \lim_{\epsilon \to 0} \frac{\mathcal{I}(x_1, \ldots, x_k + \epsilon, \ldots, x_n) - \mathcal{I}(x_1, \ldots, x_k, \ldots, x_n)}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{\mathcal{I}(\{x_i + \delta_{ik}\epsilon\}) - \mathcal{I}(\{x_i\})}{\epsilon}
\end{aligned}
\tag{26.1.2}
$$

Cuando hacemos tender $n \to \infty$, podemos heurísticamente tomar el índice discreto i como un índice contínuo y, es decir $i \to y$, el vector $\{x_i\} = (x_1, \cdots, x_n)$ de $\mathbb{R}^n$ deviene una función de C^∞, es decir $\{x_i\} \to X(y)$, la delta de Kroenecker δ_{ij} deviene una delta de Dirac $\delta(x-y)$ y la función $\mathcal{I}$ deviene un funcional $\mathcal{I}(\{x_i\}) \to I[X]$. Con estas intuiciones, podemos definir por analogía con (26.1.2), la derivada funcional de un funcional $I[f]$ de la siguiente forma

$$
\frac{\delta I}{\delta f(x)} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[I[f(y) + \epsilon \delta(x - y)] - I[f(y)]] \right.
\tag{26.1.3}
$$

Esta definición nos permite ya calcular las derivadas funcionales de los funcionales que ya hemos visto. Por ejemplo, la derivada funcional del funcional $I_{x_0}[f] = f(x)$. Según la definición tendremos

$$
\frac{\delta I_{x_0}}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[\int (f(x) + \epsilon \delta(x - y)) \, \delta(x - x_0) dx - \int f(x) \delta(x - x_0) dx \right] = \delta(x_0 - y)
\tag{26.1.4}
$$

que escribiremos de la forma más simple

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x - y) \tag{26.1.5}$$

Esta es la versión funcional de la expresión

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \tag{26.1.6}$$

Podemos también calcular la derivada funcional del funcional $I_{[a,b]}$, que vendrá dada, usando (28.5.1) por

$$\frac{\delta I_{[a,b]}}{\delta f(y)} = \int_a^b \frac{\delta f(x)}{\delta f(y)} dx = \int_a^b dx \delta(x - y) \tag{26.1.7}$$

que vale 1 is $y \in [a, b]$ y 0 si y no está dentro del intervalo de integración $[a, b]$. Finalmente, nos podemos preguntar cuánto vale la derivada funcional de una función compuesta. Es decir,

$$\frac{\delta I_x^g}{\delta f(x)} = \frac{\delta g(f(x))}{\delta f(y)}$$
$$= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[g\left(f(x) + \epsilon \delta(x - y)\right) - g\left(f(x)\right) \right] = g'(f(x))\delta(x - y) \tag{26.1.8}$$

donde hemos expandido en Taylor la función g alrededor de $\epsilon = 0$. Por tanto, g' es la derivada de la función g.

26.2   Extremos de un funcional

De la misma manera que el máximo (o mínimo) x_0 de una función $\mathcal{I}(x)$ (con $x \in \mathbb{R}^n$) se obtiene a partir de la condición de que todas las derivadas parciales se anulan,

$$\frac{\partial \mathcal{I}}{\partial x_i}(x_0) = 0, \qquad i = 1, \cdots, n \tag{26.2.1}$$

el máximo (o mínimo) de un funcional $I[f]$ ocurre para aquella función f_0 para la cual la derivada funcional se anula

$$\frac{\delta I}{\delta f(x)}[f_0] = 0 \tag{26.2.2}$$

Por ejemplo, consideremos el funcional $I[f] = \int_0^1 [f(x) - e^{-x}]^2 dx$. Este funcional tiene su extremo (máximo o mínimo, no sabemos todavía) para la función $f_0(x)$ que cumple

(26.2.2), es decir

$$\frac{\delta}{\delta f(y)} \int_0^1 [f_0(x) - e^{-x}]^2 dx = \int_0^1 \frac{\delta}{\delta f(y)} [f_0(x) - e^{-x}]^2 dx$$

$$= 2 \int_0^1 [f(x) - e^{-x}] \frac{\delta}{\delta f(y)} [f(x) - e^{-x}] dx$$

$$= 2 \int_0^1 [f(x) - e^{-x}] \delta(x - y) dx \qquad (26.2.3)$$

Por tanto, la derivada funcional de este funcional vale $f(y) - e^{-y}$ si $y \in [0,1]$ y cero si y no está en este intervalo. En cualquier caso, vemos que la derivada funcional se anula para la función $f_0(x) = e^{-x}$, con $x \in [0,1]$ que es donde el funcional $I[f]$ toma su valor extremal.

26.3 Extremos condicionados

Por lo que respecta a encontrar extremos *condicionados* de un funcional, la estrategia es prácticamente idéntica a la que se da en el caso de encontrar el extremo condicionado de una función de muchas variables, que es el método de los multiplicadores de Lagrange.

Si tenemos una función $\mathcal{I} : \mathbb{R}^n \to \mathbb{R}$ y queremos hallar su extremo bajo la condición de que se cumpla una cierta condición $g(x) = 0$, sobre el posible extremo, entonces el método de Lagrange nos asegura que basta con hallar el extremo de la función $\mathcal{I}(x) + \lambda g(x)$, donde λ es un multiplicador indeterminado. El extremo x_0 está dado por la solución de la ecuación

$$\frac{\partial}{\partial x} \left[\mathcal{I}(x_0) + \lambda g(x_0) \right] = 0 \qquad (26.3.1)$$

Claramente, esta solución dependerá paramétricamente del valor que tenga λ y, por tanto, podemos decir que x_0 es una función de λ, $x_0(\lambda)$. La solución verdadera, que es la que permite determinar λ, es la que cumple la restricción, es decir $g(x_0) = 0$, o bien, $g(x_0(\lambda)) = 0$, que es una ecuación implícita para hallar λ. Si tenemos M restricciones es decir, $g_1(x) = 0, \cdots, g_M(x) = 0$, entonces se introducen tantos multiplicadores como restricciones y se maximiza $\mathscr{I}(x) + \sum_i^M \lambda_i g_i(x)$. El valor de los M multiplicadores λ_i se obtiene de pedir que la solución de

$$\frac{\partial}{\partial x} \left[\mathcal{I}(x_0) + \sum_i^M \lambda_i g_i(x_0) \right] = 0 \qquad (26.3.2)$$

que dependerá en general de $\lambda_i, i = 1, \cdots, M$, cumpla las M restricciones $g_i(x_0(\lambda)) = 0$.

En el caso de que hablemos de funcionales en lugar de funciones en $\mathbb{R}^n$, el extremo condicionado de un funcional $I[f]$ sujeto a que otro funcional $G[f] = 0$, se obtiene simplemente hallando el extremo del funcional $I[f] + \lambda G[f]$. Este extremo ocurrirá para cierta función f_0, que dependerá paramétricamente del multiplicador λ. El valor final del multiplicador λ se obtiene pidiendo que $G[f_0(\lambda)] = 0$.

Ejemplo: Veamos con un poco de detalle el problema típico que aparece en la teoría de la información. En este caso, se introduce un funcional entropía $S[\rho]$, definido sobre el espacio de densidades de probabilidad $\rho(z)$, que tiene la siguiente forma

$$S[\rho] = -k_B \int dz \rho(z) \ln \left[\frac{\rho(z)}{\rho^{\text{ref}}(z(} \right] \tag{26.3.3}$$

donde k_B es la constante de Boltzmann y $\rho^{\text{ref}}(z)$ es una densidad de probabilidad de referencia. Como estamos hablando de densidades de probabilidad, queremos que estén normalizadas a uno, es decir, se debe cumplir la restricción

$$\int \rho(z) dz = 1 \tag{26.3.4}$$

Si queremos hallar el extremo de $S[\rho]$ sujeto a la restricción (26.3.4), tenemos que introducir un multiplicador, llamémosle μ, y maximizaremos $S[\rho] + \mu \int \rho(z) dz$. La derivada funcional de este nuevo funcional está dada por

$$\frac{\delta}{\delta\rho(z)} \left[S[\rho] + \mu \int \rho(z) dz \right] = -k_B \int dz' \frac{\delta}{\delta\rho(z)} \left(\rho(z') \ln \left[\frac{\rho(z')}{\rho^{\text{ref}}(z')} \right] \right) + \mu \int \frac{\delta\rho(z')}{\delta\rho(z)} dz'$$

$$= -k_B \int dz' \delta(z - z') \ln \left[\frac{\rho(z')}{\rho^{\text{ref}}(z')} \right]$$

$$- k_B \int dz' \rho(z') \frac{\delta}{\delta\rho(z)} \ln \left[\frac{\rho(z')}{\rho^{\text{ref}}(z')} \right] + \mu$$

$$= -k_B \ln \left[\frac{\rho(z)}{\rho^{\text{ref}}(z)} \right] - k_B \int dz' \rho(z') \frac{1}{\rho(z')} \delta(z - z') + \mu$$

$$= -k_B \left(\ln \left[\frac{\rho(z)}{\rho^{\text{ref}}(z)} \right] + 1 \right) + \mu \tag{26.3.5}$$

El máximo de $S[\rho]$ se obtiene igualando a cero esta derivada funcional

$$\ln \left[\frac{\rho_0(z)}{\rho^{\text{ref}}(z)} \right] + 1 = \mu \tag{26.3.6}$$

es decir

$$\rho_0(z) = \rho^{\text{ref}}(z) e^{\mu - 1} \tag{26.3.7}$$

Vemos que, en principio, para cada μ tenemos una solución. Pero el valor de μ está fijado

cuando requerimos que la solución cumpla la restricción (26.3.4), es decir

$$1 = \int \rho_0(z)dz = \int \rho^{\text{ref}}(z)e^{1+\mu}dz \qquad (26.3.8)$$

de donde

$$e^{1+\mu} = \frac{1}{\int \rho^{\text{ref}}(z)dz} \qquad (26.3.9)$$

de donde, finalmente,

$$\rho_0(z) = \frac{\rho^{\text{ref}}(z)}{\int \rho^{\text{ref}}(z')dz'} = \rho^{\text{ref}}(z) \qquad (26.3.10)$$

al ser $\rho^{\text{ref}}(z)$ una densidad de probabilidad que está, de por sí, normalizada.

Lo que acabamos de mostrar es que de todas las densidades de probabilidad $\rho(z)$ normalizadas a uno, la que hace máximo el valor de $S[\rho]$ definido en (26.3.3) es precisamente $\rho^{\text{ref}}(z)$. En general, si pedimos más restricciones (además de la normalización) a $\rho(z)$, tendremos que añadir más multiplicadores de Lagrange.

27

Appendix: Macrocanonical phase space

We have noted in the previous section how the impossibility to know the initial state of a system of many degrees of freedom forces us to use probabilistic concepts. We now note that not only it is impossible to know the initial microstate of the system, but also we often *do not know exactly how many particles constitute our system*. By performing a collection of experiments with identical preparation, we may end up with systems with slightly different number of particles. After all, atoms and molecules are so tiny that we may skip a few of them from experiment to experiment. Of course, we should aim at having controlled at least the *average* number of particles in our system because the behaviour of a system is very different for different number of particles, the rest of conditions on the system being the same. Reproducible experiments require certain degree of control. But if we only know the average number of particles, not the actual number, what phase space should we use for the microscopic description of the system? Note that the phase space has a dimension equal to the number of particles of the system, which is a number that it is not exactly known. We should emphasize that our system is isolated, with a conserved Hamiltonian, and does not exchange particles with any other system. Not knowing the exact number of particles does not imply, in this perspective, that the number of particles is a dynamic quantity that fluctuates in time, nor the result of interchanging particles with a bath.

The idea is to work with *many* phase spaces at the same time, each one weighted with the probability $P(N)$ that the system has exactly N particles. In each phase space we have a probability density $\rho(z_N)$ that the microscopic state of the system is z_N. Physical quantities will be averages in which, in addition to average over the probability $\rho(z_N)$ we also average over $P(N)$. One way to get this construction at work is by introducing the **macrocanonical phase space** as the infinite set of the phase space of systems of $N = 0, 1, \cdots, \infty$ particles. Any microstate z of the macrocanonical phase space is of the form $z_N = \{\mathbf{q}_1, \mathbf{p}_1, \cdots, \mathbf{q}_N, \mathbf{p}_N\}$ for some N, where $\mathbf{q}_i, \mathbf{p}_i$ are the position and momentum of particle i. The probability density $\rho(z)$ of finding a particular microstate z in the macrocanonical phase space is given by the collection of probability densities

$\rho(z_N)$ defined in each phase space of given number N of particles,

$$\rho = \{\rho(z_1), \rho(z_2), \cdots, \rho(z_N), \cdots\} \tag{27.0.1}$$

The probability density is normalized in the following way

$$\sum_{N=0}^{\infty} \int dz_N \rho(z_N) = 1 \tag{27.0.2}$$

In this way, we may interpret

$$P(N) = \int dz_N \rho(z_N) \tag{27.0.3}$$

as the probability that the system has exactly N particles. Note that we have included the case $N = 0$ in order to accept the case that the system contains no particle with certain probability.

A phase function $\hat{A}(z)$ in the macrocanonical phase space should be understood as an infinite collection of phase functions $\hat{A}(z) = \{\hat{A}(z_1), \cdots, \hat{A}(z_N), \cdots\}$, each one defined in the phase space Γ_N of fixed number N of particles. Usually the functions $\hat{A}(z_N)$ are all "identical in form" involving sums over particles, and they differ from each other just by the number of particles. A typical example is the Hamiltonian

$$\hat{H}(z_N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \sum_{i<j=1}^{N} V(q_{ij}) \tag{27.0.4}$$

where q_{ij} is the distance between two particles. Another phase function that we may construct in phase space is **the number of particles** $\hat{N}(z_N)$ which is defined as a function that takes the same value N for all the microstates of the phase space Γ_N, this is

$$\hat{N}(z_N) = N \tag{27.0.5}$$

Sometimes it is convenient to think of this function as a dynamical invariant, in the same footing as the Hamiltonian.

The average of a phase function in the macrocanonical phase space is defined as

$$\langle \hat{A} \rangle = \text{Tr}[\rho \hat{A}] = \sum_{N=0}^{\infty} \int dz_N \rho(z_N) \hat{A}(z_N) \tag{27.0.6}$$

where we introduce the shorthand trace notation $\text{Tr}[\cdots]$ borrowed from Quantum Mechanics to denote a sum over the number of particles and an integral over the microstates of the phase space of N particles

$$\text{Tr}[\cdots] = \sum_{N=0}^{\infty} \int dz_N \rho(z_N) \cdots \tag{27.0.7}$$

Unless explicitly stated, in the rest of this book, we will assume that the system is described microscopically with a macrocanonical phase space. The situation in which we know with certainty the number of particles of our system, as it happens in molecular dynamics simulations, for example, can be obtained as a particular case of a macrocanonical description for which all $\rho(z_N)$ are zero except for one particular N.

The equilibrium ensemble in the macrocanonical phase space

In the macrocanonical phase space introduced above, the equilibrium ensemble is, in fact, a collection of ensembles in each individual phase space $\rho^{\mathrm{eq}}(z) = \{\rho_0^{\mathrm{eq}}, \rho^{\mathrm{eq}}(z_1), \cdots, \rho^{\mathrm{eq}}(z_N), \cdots\}$. Each ensemble $\rho^{\mathrm{eq}}(z_N)$ is of the form (1.8.1), this is, a function of the dynamical invariants of the N-th phase space. We assume for simplicity of presentation that the energy is the only dynamic invariant in the system but what follows is easily generalized to the case that more dynamical invariants exist. In this case, the limiting effective equilibrium ensemble that results from the infinite time evolution of the densities in the macrocanonical phase space are

$$\lim_{t \to \infty} \rho(z_N, t) = \rho^{\mathrm{eq}}(z_N) = g(\hat{H}(z_N)). \tag{27.0.8}$$

The function $g(E)$ is connected to the probability distribution $P(N, E)$ of dynamical invariants (which is itself invariant) in the N-th phase space. By definition,

$$P(N, E) = \int dz_N \rho^{\mathrm{eq}}(z_N) \delta(\hat{H}(z_N) - E) \tag{27.0.9}$$

In sloppy but significative words, in order to obtain $P(N, E)$ we are summing (hence the $\int$ symbol) the probability (hence the bit $dz_N \rho(z_N)$) of those microstates of N particles that have exactly the energy E (hence the Dirac delta function $\delta(\hat{H}(z_N) - E)$). By inserting (27.0.8) into (27.0.9) we have

$$P(N, E) = \int dz_N g(\hat{H}(z_N)) \delta(\hat{H}(z_N) - E) = g(E) \frac{1}{\rho_N^0} \Omega(N, E), \tag{27.0.10}$$

where we have introduced the **structure function**[1] $\Omega(N, E)$ that gives the measure of the region of the N-th phase space compatible with a given set of dynamical invariants

$$\Omega(N, E) \equiv \int dz_N \rho_N^0 \delta(\hat{H}(z_N) - E). \tag{27.0.11}$$

The factor ρ_N^0 introduced in both (27.0.10) and (27.0.11) is given in (1.8.19) and renders $\Omega(N, E)$ and $P(N, E)$ with the same dimensions of $[\text{energy}]^{-1}$.

Equation (27.0.10) allows one to identify $g(E)$ and (27.0.8) becomes

$$\rho^{\mathrm{eq}}(z_N) = \rho_N^0 \frac{P(N, \hat{H}(z_N))}{\Omega(N, \hat{H}(z_N))}. \tag{27.0.12}$$

This is the most general ensemble in the macrocanonical phase space that results from the

---

[1] This terminology was introduced by A.I.Khinchin in [?].

mixing property of the Hamiltonian dynamics. We need to invent a name for the equilibrium ensemble (27.0.12) and perhaps one sufficiently suggestive is the **equilibrium mesocanonical ensemble**. The mesocanonical equilibrium ensemble is normalized because the probability $P(N, E)$ is normalized itself

$$\sum_{N=0}^{\infty} \int dz_N \rho^{\text{eq}}(z_N) = \sum_{N=0}^{\infty} \int dz_N \frac{P(N, \hat{H}(z_N))}{\Omega(N, \hat{H}(z_N))} \int dE \delta(\hat{H}(z_N) - E)$$

$$= \sum_{N=0}^{\infty} \int dE \frac{P(N, E)}{\Omega_0(N, E)} \int dz_N \delta(\hat{H}(z_N) - E)$$

$$= \sum_{N=0}^{\infty} \int dE P(N, E) = 1 \qquad (27.0.13)$$

where we have introduced the identity

$$\int dE \delta(\hat{H}(z_N) - E) = 1 \qquad (27.0.14)$$

The mesocanonical ensemble is defined in terms of the unspecified probability density $P(N, E)$. The particular form of this probability is dictated by the initial preparation of the system. In Sec **??**, we will consider two particular preparations of the system at the initial time (thermalization with a bath, with or without exchange of particles) that lead to specific forms for $P(N, E)$. These preparations will lead to the equilibrium canonical and macrocanonical ensembles, respectively. In MD simulations where the number of particles and the energy are exactly known, the probability $P(N, E) = \delta_{NN_0}\delta(E - E_0)$ and the mesocanonical ensemble becomes the usual **equilibrium microcanonical ensemble**

$$\rho^{\text{mic}}_{N_0 E_0}(z_N) \equiv \rho_N^0 \delta_{NN_0} \frac{\delta(\hat{H}(z_N) - E_0)}{\Omega(N_0, E_0)} \qquad (27.0.15)$$

In the microcanonical ensemble the probability density of realization of the microstate z_N is zero unless $N = N_0$ and the energy of z_N equals E_0.

In this way, we depart from usual interpretations in which a macrocanonical ensemble is regarded as an ensemble describing a system in contact with a reservoir of particles. We regard the macrocanonical ensemble as just a way to describe our ignorance about the specific number of particles that make our system. However, even though we do not may know the exact number of particles, the system is always assumed to have a fixed number of particles. From a statistical mechanics point of view, this approach is due to the difficulty of working with a phase space of changing number of dimensions.

28

Appendix: Discretizing continuum equations

28.1 Motivation

For the numerical solution of partial differential equations in Physics, it is necessary to convert the problem in continuum space into a problem in a discrete space, amenable of treatment with a computer. Usual procedures rely on assigning values of the fields to nodes of a grid. We are interested in discretizations on *arbitrary* distributions (not necesarily regular) of points because generating grids is the most human time consuming part of the simulations. Lagrangian grids in fluid mechanics have some advantadges in free surface flows over Eulerian grids. Lagrangian grids are essentially irregular. Also, remeshing is an issue that appears often. Remeshing is the art of redistributing the values of the field in a grid onto another grid, perhaps more regular. Remeshing has associated an unacceptable numerical diffusion.

In this note, we formalize the process of discretizing a field, which is previous to any procedure for solving dynamic equations for this field. We also consider implications in the discretization of partial differential equations.

28.2 Discretization and Continuation operators

Consider a field $\alpha(\mathbf{r})$ which is a differentiable function of space $\mathbf{r} \in \mathbb{R}^3$. We want to discretize this field in a set of points $x = \{\mathbf{r}_1, \cdots, \mathbf{r}_M\}$. This set of points x is called a *configuration* or, also, a *cloud* and, if connectivity between points is defined, they may form a *mesh*. On every point $\mathbf{r}_\mu$ we associate a number a_μ which "represents" the field $\alpha(\mathbf{r})$ at this point. One possibility, of course is $a_\mu = \alpha(\mathbf{r}_\mu)$ but there are other possibilities. We call the collection a_x of the pairs $a_\mu, \mathbf{r}_\mu$ as the *discrete field*, while the original $\alpha(\mathbf{r})$ field is called the *continuous field*. It is obvious that a discrete field contains less information than a continuum field.

Introduce the *discretization operator* D_x. It takes any continuous field and produces

a discrete field, this is

$$D_x[\alpha] = a_x \qquad (28.2.1)$$

The square brackets denote that D_x is a functional of $\alpha(\mathbf{r})$ and the subindex x denotes that the discretization depends on the configuration x. Note that the discretization operator may take many different continuum fields into the same discrete field. In that sense, it *destroys* information from the continuum field.

Introduce the *continuation operator* C that takes a discrete field and transforms it into a continuum field, this is

$$C(a_x) = \alpha(\mathbf{r}) \qquad (28.2.2)$$

The rounded parenthesis denote that C is a function of a_x. The continuation operator *creates* information.

28.2.1   Required properties for D_x and C

A first property that we require on these operators is that if we continue a discrete field a_x and then we discretize it on the configuration x we should get the same answer a_x. Formally

$$D_x[C(a_x)] = a_x \qquad (28.2.3)$$

We can call this property as *projection*. It seems that this is a basic property to be satisfied by D_x and C in order to not degrade a discrete field under the action of the continuation.

The second property that seems natural to ask is that the result of continuing a discrete field a_x and then discretizing at another cloud x', which is $a'_{x'}$, should be such that if we continue $a'_{x'}$ and discretize it at the cloud x we get back a_x. Formally

$$a'_{x'} = D_{x'}[C(a_x)]$$
$$a_x = D_x[C(a'_{x'})] \qquad (28.2.4)$$

We can call this property as *reversibility*. The reversibility property is a basic requirement that appears when considering the possibility of *remeshing* a field. Remeshing is the process by which a discrete field a_x, defined in a configuration x is mapped onto a discrete field $a'_{x'}$ defined in another configuration x' (another mesh). Therefore, the reversibility property states that the result of remeshing a discrete field into another configuration and remeshing again back to the original configuration should lead to the original discrete field. It seems to us that this property will be essential to reduce the "numerical diffusion" that appears in remeshing techniques.

While the projective property may always be satisfied, the reversible property cannot be satisfied in general. Actually, for Eq. (28.2.4) to happen it must be true that

$$D_x[C(D_{x'}[C(a_x)])] = a_x \qquad (28.2.5)$$

Eq. (28.2.3) shows that the operator $D_x C$ that takes discrete fields onto discrete fields is the identity operator, this is, $D_x C = \mathbf{1}$. This suggest that D_x and C may be inverse of each other. However, for a true inverse property, we would need to have also that $CD_x = \mathbf{1}$. The operator CD_x takes continuum fields into continuum fields, for example,

$$C(D_{x'}[\alpha]) = \overline{\alpha}(\mathbf{r}) \tag{28.2.6}$$

It is apparent that, in general, $\overline{\alpha}(\mathbf{r}) \neq \alpha(\mathbf{r})$ because many different continuum field may discretize into the same discrete field. Note that Eq. (28.2.5) would be satisfied if CD_x would be the identity operator. However, the operator CD_x is in general a *filter* because, necessarily the information contained in $\overline{\alpha}(\mathbf{r})$ is smaller than that in $\alpha(\mathbf{r})$. If you first destroy and then create, you loose information.

What we learn from this discussion is that we may get both projection *and* reversibility to the extent that the operators D_x and C are inverse of each other, and that this is not possible for arbitrary continuum fields. In fact, we have the following theorem. Define $f(\mathbf{r})$ as an *exactly continued field of C* if it satisfies

$$C(f(\mathbf{r}_\mu)) = f(\mathbf{r}) \tag{28.2.7}$$

We may say that the operator C *exactly continues* these fields. The projective property (28.2.3) then implies that

$$D_x[f(\mathbf{r})] = f(\mathbf{r}_\mu) \tag{28.2.8}$$

As a consequence, for exactly continued fields, we have

$$CD_x[f(\mathbf{r})] = f(\mathbf{r}) \tag{28.2.9}$$

Therefore, given the projective property (28.2.3), the reversible property (28.2.4) is satisfied on all the continuum fields that are exactly continued by C.

For example, assume that the exactly continued fields of C are quadratic functions. Then

$$C(a + \mathbf{b} \cdot \mathbf{r}_\mu + \mathbf{c} : \mathbf{r}_\mu \mathbf{r}_\mu) = a + \mathbf{b} \cdot \mathbf{r} + \mathbf{c} : \mathbf{rr} \tag{28.2.10}$$

where $a, \mathbf{b}, \mathbf{c}$ are constant. Then, by requiring that $D_x C = \mathbf{1}$ we obtain that the discretization operator D_x must satisfy

$$\begin{aligned} D_x[1] &= 1 \\ D_x[\mathbf{r}] &= \mathbf{r}_\mu \\ D_x[\mathbf{rr}] &= \mathbf{r}_\mu \mathbf{r}_\mu \end{aligned} \tag{28.2.11}$$

These properties, in turn, imply that,

$$CD_x[a + \mathbf{b} \cdot \mathbf{r} + \mathbf{c} : \mathbf{rr}] = a + \mathbf{b} \cdot \mathbf{r} + \mathbf{c} : \mathbf{rr} \tag{28.2.12}$$

As a consequence, $CD_x = \mathbf{1}$ *if applied to quadratic functions.*

The main conclusion of this discussion is that given the operators D_x, C, it is very

important to find out which are their exactly continued fields, because on these fields it is possible to have both, the projective and the reversibility properties.

28.2.2   Local operators

Up to now we have not specified the form of the operators D_x, C. We introduce now a set of functions $\delta_\mu(\mathbf{r})$ and $\psi_\mu(\mathbf{r})$ associated to each point $\mathbf{r}_\mu$. We tend to think about these functions as "localised" around each point, but this is not necessary on the following. The discretization operator D_x is defined through the following action

$$a_\mu = \int d\mathbf{r}\delta_\mu(\mathbf{r})\alpha(\mathbf{r}) \qquad (28.2.13)$$

that converts an arbitrary field $\alpha(\mathbf{r})$ into a set of discrete values a_μ. The basis function $\delta_\mu(\mathbf{r})$ has dimensions of inverse of a volume. The continuation operator C is defined as

$$\alpha(\mathbf{r}) = \sum_\mu \psi_\mu(\mathbf{r})a_\mu \qquad (28.2.14)$$

The properties (28.2.3) and (28.2.5) now reflect onto properties for the functions ϕ_μ, ψ_μ. The first projection property in Eq. (28.2.3) becomes

$$a_\mu = \int d\mathbf{r}\delta_\mu(\mathbf{r})\alpha(\mathbf{r}) = \int d\mathbf{r}\delta_\mu(\mathbf{r})\sum_\nu \psi_\nu(\mathbf{r})a_\nu = \sum_\nu \left[\int d\mathbf{r}\delta_\mu(\mathbf{r})\psi_\nu(\mathbf{r})\right]a_\nu \qquad (28.2.15)$$

Because this must be true for all a_μ we obtain that the local versions of D_x and C will be projective if the basis functions satisfy the orthogonality condition

$$\int d\mathbf{r}\delta_\mu(\mathbf{r})\psi_\nu(\mathbf{r}) = \delta_{\mu\nu} \qquad (28.2.16)$$

The reversibility property in Eq. (28.2.5) now becomes

$$\int d\mathbf{r}\delta_\nu(\mathbf{r})\sum_{\mu'}\psi_{\mu'}(\mathbf{r})\int d\mathbf{r}'\delta_{\mu'}(\mathbf{r}')\sum_\mu \psi_\mu(\mathbf{r}')a_\mu = a_\nu$$

$$(28.2.17)$$

Note that here $\phi_{\mu'}(\mathbf{r})$ is the set of functions corresponding to the configuration x' and it is different from $\phi_\mu(\mathbf{r})$ which is defined on the configuration x. Note that Eq. (28.2.17) implies, because a_μ is arbitrary, that

$$\int d\mathbf{r}\int d\mathbf{r}'\psi_\mu(\mathbf{r})D(\mathbf{r},\mathbf{r}')\delta_\nu(\mathbf{r}') = \delta_{\mu\nu} \qquad (28.2.18)$$

where

$$D(\mathbf{r},\mathbf{r}') = \sum_\sigma \psi_\sigma(\mathbf{r})\delta_\sigma(\mathbf{r}') \qquad (28.2.19)$$

The explicit form (28.2.18) of the reversibility condition is not very important because we already know that it is not possible to have Eq. (28.2.18) satisfied exactly in general. However, this identity will be satisfyed automatically for the exactly continued fields.

28.2.3 The basis functions δ_μ, ψ_μ are linearly related

We will assume that the local discretization weight functions $\delta_\mu(\mathbf{r})$ can be expressed as linear combinations of the local continuation weight functions $\psi_\mu(\mathbf{r})$, this is

$$\delta_\mu(\mathbf{r}) = \sum_\nu M^\delta_{\mu\nu} \psi_\nu(\mathbf{r}) \tag{28.2.20}$$

where $M^\delta_{\mu\nu}$ is the matrix of the linear combination. In this case, the projection property (28.2.3) turns out to be

$$\delta_{\mu\nu} = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \psi_\nu(\mathbf{r}) = \sum_{\mu'} M^\delta_{\mu\mu'} M^\psi_{\mu'\nu} \tag{28.2.21}$$

where the symmetric and positive definite mass matrix is defined as

$$M^\psi_{\mu\nu} = \int d\mathbf{r} \psi_\mu(\mathbf{r}) \psi_\nu(\mathbf{r}), \tag{28.2.22}$$

If we assume that the basis functions are dimensionless, the mass matrix $\mathbf{M}^\psi$ has dimensions of volume. According to (28.2.21) we have that the matrices $\mathbf{M}^\delta, \mathbf{M}^\psi$ are inverse of each other

$$M^\delta_{\mu\nu} = [M^\psi]^{-1}_{\mu\nu} \tag{28.2.23}$$

It is easy to show that the matrix $M^\delta_{\mu\nu}$ is given by

$$M^\delta_{\mu\nu} = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \delta_\nu(\mathbf{r}), \tag{28.2.24}$$

In summary, given a set of continuation weight functions $\psi_\mu(\mathbf{r})$ we can always construct a set of discretization weight functions $\delta_\mu(\mathbf{r})$ with

$$\delta_\mu(\mathbf{r}) = \sum_\nu [M^\psi]^{-1}_{\mu\nu} \psi_\nu(\mathbf{r}) \tag{28.2.25}$$

in such a way that the projective property (28.2.16) is satisfied.

We will often use the notation $\boldsymbol{\delta}(\mathbf{r}) = \{\delta_1(\mathbf{r}), \cdots, \delta_M(\mathbf{r})\}$ and $\boldsymbol{\psi}(\mathbf{r}) = \{\psi_1(\mathbf{r}), \cdots, \psi_M(\mathbf{r})\}$. In this way, we may write Eq. (28.2.25) as

$$\boldsymbol{\delta}(\mathbf{r}) = \mathbf{M}^\delta \boldsymbol{\psi}(\mathbf{r}) \tag{28.2.26}$$

The mass matrices $\mathbf{M}^{\delta}$ and $\mathbf{M}^{\psi}$ may be written as

$$\mathbf{M}^{\delta} = \left(\boldsymbol{\delta}^{T}, \boldsymbol{\delta}\right)$$
$$\mathbf{M}^{\psi} = \left(\boldsymbol{\psi}^{T}, \boldsymbol{\psi}\right) \tag{28.2.27}$$

with an obvious notation for the scalar product $(\,\cdot\,,\,\cdot\,)$ in functional space.

28.2.4   Linear consistency

We further assume that the functions $\psi_{\mu}(\mathbf{r})$ satisfy the partition of unity and linear consistency properties

$$\sum_{\mu} \psi_{\mu}(\mathbf{r}) = 1$$
$$\sum_{\mu} \mathbf{r}_{\mu} \psi_{\mu}(\mathbf{r}) = \mathbf{r} \tag{28.2.28}$$

In this way, a linear discrete field of the form $a_{\mu} = a + \mathbf{b} \cdot \mathbf{r}_{\mu}$ is continued into a linear field, this is

$$C(a + \mathbf{b} \cdot \mathbf{r}_{\mu}) = a + \mathbf{b} \cdot \mathbf{r} \tag{28.2.29}$$

It is a matter of some calculation to show that if Eq. (28.2.16) and Eq. (28.2.28) are satisfied, then the reversibility property (28.2.17) will apply *for linear discrete fields* of the form $a_{\mu} = a + \mathbf{b} \cdot \mathbf{r}_{\mu}$.

As a consequence of Eq. (28.2.16) and Eq. (28.2.28) we have that the functions $\delta_{\mu}(\mathbf{r})$ satisfy

$$\int d\mathbf{r}\delta_{\mu}(\mathbf{r}) = 1$$
$$\int d\mathbf{r}\mathbf{r}\delta_{\mu}(\mathbf{r}) = \mathbf{r}_{\mu} \tag{28.2.30}$$

Let us investigate to what extent we may have properties similar to (28.2.28) for the discretization basis function $\delta_{\mu}(\mathbf{r})$. To this end, we first introduce the volume and center of mass of the node μ through the definitions

$$\mathscr{V}_{\mu} \equiv \int d\mathbf{r}\psi_{\mu}(\mathbf{r})$$
$$\overline{\mathbf{r}}_{\mu} \equiv \frac{1}{\mathscr{V}_{\mu}} \int d\mathbf{r}\psi_{\mu}(\mathbf{r})\mathbf{r} \tag{28.2.31}$$

The linear consistency (28.2.28) has the following implications on the mass matrix defined

in Eq. (28.2.22)

$$\sum_\nu M_{\mu\nu}^\psi = \mathscr{V}_\nu$$

$$\sum_\nu \mathbf{r}_\nu M_{\mu\nu}^\psi = \mathscr{V}_\nu \bar{\mathbf{r}}_\nu \qquad (28.2.32)$$

These two equations can be thought as matrix equations that may be inverted, leading to

$$\sum_\nu \mathscr{V}_\nu M_{\mu\nu}^\delta = 1$$

$$\sum_\nu \mathscr{V}_\nu \mathbf{r}_\nu M_{\mu\nu}^\delta = \bar{\mathbf{r}}_\nu \qquad (28.2.33)$$

By multiplying with $\psi_\nu(\mathbf{r})$ and summing over ν, we have

$$\sum_\nu \mathscr{V}_\nu \delta_\nu(\mathbf{r}) = 1$$

$$\sum_\nu \mathscr{V}_\nu \mathbf{r}_\nu \delta_\nu(\mathbf{r}) = \sum_\nu \bar{\mathbf{r}}_\nu \psi_\nu(\mathbf{r}) \qquad (28.2.34)$$

If the center of mass $\bar{\mathbf{r}}_\nu$ of cell ν coincides with the node position $\mathbf{r}_\nu$, then we have a similar linear consistency relation for the discretization basis functions $\delta_\mu(\mathbf{r})$, this is

$$\sum_\mu \mathscr{V}_\mu \delta_\mu(\mathbf{r}) = 1$$

$$\sum_\mu \mathscr{V}_\mu \mathbf{r}_\mu \delta_\mu(\mathbf{r}) = \mathbf{r} \qquad (28.2.35)$$

Therefore, for regular grids, the destruction operators are also linearly consistent. This implies that we could use, in a regular grid, a Delaunay finite element for $\delta_\mu(\mathbf{r})$ and the scheme will be still linearly consistent.

For a linearly consistent scheme, we note the following identities

$$\int d\mathbf{r}' D(\mathbf{r}, \mathbf{r}') = 1$$

$$\int d\mathbf{r}' \mathbf{r}' D(\mathbf{r}, \mathbf{r}') = \mathbf{r} \qquad (28.2.36)$$

where $D(\mathbf{r}, \mathbf{r}')$ has been introduced in Eq. (28.2.19). The proof is as follows

$$1 = \sum_\mu \psi_\mu(\mathbf{r}) = \sum_{\mu\nu} M_{\mu\nu}^\psi \delta_\nu(\mathbf{r}) = \sum_{\mu\nu} \int d\mathbf{r}' \psi_\mu(\mathbf{r}')\psi_\nu(\mathbf{r}')\delta_\nu(\mathbf{r}) = \sum_\nu \int d\mathbf{r}' \psi_\nu(\mathbf{r}')\delta_\nu(\mathbf{r}) = \int d\mathbf{r}' D(\mathbf{r}', \mathbf{r})$$

$$(28.2.37)$$

and

$$\mathbf{r} = \sum_\mu \mathbf{r}_\mu \psi_\mu(\mathbf{r}) = \sum_{\mu\nu} \mathbf{r}_\mu M^\psi_{\mu\nu} \delta_\nu(\mathbf{r}) = \sum_{\mu\nu} \int d\mathbf{r}' \mathbf{r}_\mu \psi_\mu(\mathbf{r}') \psi_\nu(\mathbf{r}') \delta_\nu(\mathbf{r}) = \sum_\nu \int d\mathbf{r}' \mathbf{r}' \psi_\nu(\mathbf{r}') \delta_\nu(\mathbf{r}) = \int d\mathbf{r}' \mathbf{r}' D(\mathbf{r}', \mathbf{r})$$

$$(28.2.38)$$

28.2.5   Delaunay finite elements

One specific choice for the interpolation basis function $\psi_\mu(\mathbf{r})$ that satisfies (28.2.28) is the finite element defined on the Delaunay triangulation and denoted with $\phi_\mu(\mathbf{r})$. The corresponding discretization basis function $\boldsymbol{\delta}(\mathbf{r})$ is given by (28.2.26).

Delaunay tessellation in 1D is explicitly given by:

$$\phi_\mu(x) = \theta(x - x_{\mu-1})\theta(x_\mu - x)\frac{x - x_{\mu-1}}{x_\mu - x_{\mu-1}}$$

$$+ \theta(x - x_\mu)\theta(x_{\mu+1} - x)\frac{x_{\mu+1} - x}{x_{\mu+1} - x_\mu}$$

$$(28.2.39)$$

where $\theta(x)$ is the usual Heaviside step function. Figure 28.1 shows function $\phi_\mu(x)$ for three Delaunay cells.

Figure 28.1: Delaunay tessellation in 1D. Each node μ has two elements, e^l_μ and e^r_μ shared with its neighbor nodes.

In a centroidal grid in which the center of mass of a cell is located at the node, we may use for $\delta_\mu(\mathbf{r}) = \phi_\mu(\mathbf{r})/\mathcal{V}_\mu$, where $\mathcal{V}_\mu = \frac{1}{2}(x_{\mu+1} - x_{\mu-1})$ in 1D. In this way, the corresponding continuation basis functions will interpolate exactly linear functions. It should be clear that if the grid is not regular then the use of the Delaunay finite element for the discretization operator $\delta_\mu(\mathbf{r})$ will not allow to interpolate exactly the linear functions. Therefore, what we have done in our previous JCP paper with Arturo, where we used the Delaunay finite element for $\delta_\mu(\mathbf{r})$ works only because we used a regular grid. It is expected that in irregular grids we get bad results.

As an example we may consider a 1D regular grid in which $x_\mu - x_{\mu-1} = a$ for every node μ. In that case, the mass matrix will be given by

$$\mathbf{M}^\psi = \frac{a}{6}\begin{pmatrix} 4 & 1 & 0 & 0 & \dots & 1 \\ 1 & 4 & 1 & 0 & \dots & 0 \\ 0 & 1 & 4 & 1 & \dots & 0 \\ & & & \ddots & & \end{pmatrix}$$

$$(28.2.40)$$

We introduce a finite element shape function $\phi_\mu(\mathbf{r})$ whose support is the Delaunay cell of node μ, which is the convex hull of the neighbours of μ, according to the Delaunay triangulation. In 1D this shape function has the shape of a tent and in 2D has the shape of a piramid. The details of this construction are given in Ref. [?] and we follow the same

notation. Suffices to state here that the shape functions satisfies the following properties

$$0 \le \phi_\mu(\mathbf{r}) \le 1$$

$$\phi_\mu(\mathbf{r}_\nu) = \delta_{\mu\nu}$$

$$\sum_\mu^M \phi_\mu(\mathbf{r}) = 1$$

$$\sum_\mu^M \phi_\mu(\mathbf{r})\mathbf{r}_\mu = \mathbf{r} \qquad (28.2.41)$$

The third property is called the partition of unity while the last one is usually referred to as linear consistency. These properties show that we may use $\phi_\mu(\mathbf{r})$ as a good interpolating function. We finally introduce the volume of the Delaunay cell at $\mathbf{r}_\mu$ as

$$\mathcal{V}_\mu = \int d\mathbf{r}\phi_\mu(\mathbf{r}) \qquad (28.2.42)$$

In 2D, this volume is the sum over all the triangles of $1/3$ of the area of each triangle of the Delaunay cell, while in 3D is $1/4$ of the volume of the tetrahedra of the Delaunay cell. The shape function $\phi_\mu(\mathbf{r})$ of node μ has the form of a pyramid, as shown in Fig. **??** for 2D. The explicit functional form of this pyramid is given by

$$\phi_\mu(\mathbf{r}) = \sum_\nu \left[a_{\mu\nu\sigma} + \mathbf{b}_{\mu\nu\sigma}{\cdot}\mathbf{r}\right]\theta_{\mu\nu\sigma}(\mathbf{r}) \qquad (28.2.43)$$

where the sum goes over the neighbours ν of μ, and σ is the clock-wise neighbour of μ, ν, as in Fig. **??**. Here $\theta_{\mu\nu\sigma}(\mathbf{r})$ is the characteristic function of the triangle $\mu\nu\sigma$ taking the value 1 inside the triangle and zero outside. We have introduced the constants

$$a_{\mu\nu\sigma} = \frac{\mathbf{r}_\nu{\cdot}\boldsymbol{\omega}{\cdot}\mathbf{r}_\sigma}{2A_{\mu\nu\sigma}}$$

$$\mathbf{b}_{\mu\nu\sigma} = \frac{\boldsymbol{\omega}{\cdot}\mathbf{r}_{\nu\sigma}}{2A_{\mu\nu\sigma}}$$

$$A_{\mu\nu\sigma} = \frac{1}{2}\mathbf{r}_{\mu\nu}{\cdot}\boldsymbol{\omega}{\cdot}\mathbf{r}_{\mu\sigma}$$

$$\mathbf{r}_{\mu\nu} = \frac{\mathbf{r}_\mu + \mathbf{r}_\nu}{2}$$

$$\boldsymbol{\omega} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad (28.2.44)$$

$A_{\mu\nu\sigma}$ is the (signed) area of the triangle of vertexes $\mathbf{r}_\mu, \mathbf{r}_\nu, \mathbf{r}_\sigma$. Note that $\boldsymbol{\omega}^T\boldsymbol{\omega} = \mathbf{1}$. For every sub-element e, whose nodes are μ, ν, σ there are three vectors, $\mathbf{b}_{\mu\nu\sigma}, \mathbf{b}_{\nu\sigma\mu}, \mathbf{b}_{\sigma\mu\nu}$ depicted in Fig. **??** and are directed towards the node appearing in the first index of the vector $\mathbf{b}$. According to its definition in Eq. (28.2.44) the three $\mathbf{b}$ vectors of a sub-element

satisfy the following identity

$$\mathbf{b}_{\mu\nu\sigma} + \mathbf{b}_{\nu\sigma\mu} + \mathbf{b}_{\sigma\mu\nu} = 0 \qquad (28.2.45)$$

In a similar way we have the identity

$$a_{\mu\nu\sigma} + a_{\nu\sigma\mu} + a_{\sigma\mu\nu} = 1 \qquad (28.2.46)$$

We may switch to an "element notation" where the three vectors of element e that point to the nodes $\mu, \nu, \sigma \in e$ are $\mathbf{b}_{e_\mu}, \mathbf{b}_{e_\nu}, \mathbf{b}_{e_\sigma}$ (the element is the same, the pointing node changes). The identity (28.2.45) now becomes

$$\sum_{\mu \in e} \mathbf{b}_{e_\mu} = 0 \qquad (28.2.47)$$

which simply expresses that if we sum all the $\mathbf{b}$ vectors of element e pointing towards the nodes of that element we get zero. The identity (28.2.46) can be expressed as

$$\sum_{\mu \in e} a_{e_\mu} = 1 \qquad (28.2.48)$$

The pyramid itself can be written in element notation as

$$\phi_\mu(\mathbf{r}) = \sum_{e \in \mu} \left[a_{e_\mu} + \mathbf{b}_{e_\mu} \cdot \mathbf{r} \right] \theta_{e_\mu}(\mathbf{r}) \qquad (28.2.49)$$

In what follows we will need to transform repeatedly sums over nodes into sums over elements. We can get some familiarity by computing the sum $\sum_\mu \phi_\mu(\mathbf{r})$ which we know that it sums up to one. Let us see how this result is obtained by switching from sums over nodes to sums over elements

$$\sum_\mu \phi_\mu(\mathbf{r}) = \sum_\mu \sum_{e \in \mu} \left[a_{e_\mu} + \mathbf{b}_{e_\mu} \cdot \mathbf{r} \right] \theta_{e_\mu}(\mathbf{r}) = \sum_e \theta_e(\mathbf{r}) \sum_{\mu \in e} \left[a_{e_\mu} + \mathbf{b}_{e_\mu} \cdot \mathbf{r} \right] = \sum_e \theta_e(\mathbf{r}) = 1$$

$$(28.2.50)$$

where in the second equality we have expressed the sum over all elements of all nodes in terms of the sum over al nodes of all elements, in the third equality we have used (28.2.47),(28.2.48), and, in the last identity we have used that the characteristic function of the elements is indeed a partition of unity. The gradient of the discrete delta function defined in (**??**) has the following form

$$\nabla \delta_\mu(\mathbf{r}) = M^\delta_{\mu\nu} \sum_{e_\nu} \mathbf{b}_{e_\nu} \theta_{e_\nu}(\mathbf{r}) \qquad (28.2.51)$$

where e_ν labels each of the sub-elements of the node ν, $\mathbf{b}_{e_\nu}$ is a constant vector within the sub-element e_ν that is pointing towards the node ν and $\theta_{e_\nu}(\mathbf{r})$ is the characteristic function of the sub-element e_ν. In Fig. 28.2 we show the sub-elements e_ν of the node ν and the corresponding vectors $\mathbf{b}_{e_\nu}$.
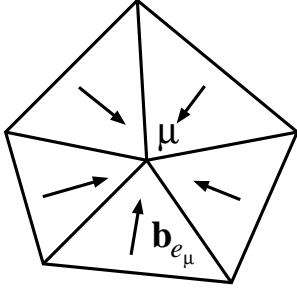
Figure 28.2: The node μ and its neighbours define five sub-elements e_μ in which the vector $\mathbf{b}_{e_\mu}$ is constant. The characteristic function $\theta_{e_\mu}(\mathbf{r})$ takes the value 1 inside the triangle e_μ and zero outside.

28.3   Spatial Derivatives

Given the discrete version of a function a_x we may want to estimate spatial derivatives at the configuration x. One possibility is to continue a_x, compute the derivatives and then discretize the derivative function at x. For the case of the gradient, for example, we may define

$$(\boldsymbol{\nabla} a)_x \equiv D_x \left[\boldsymbol{\nabla} C \left(a_x \right) \right] \tag{28.3.1}$$

For local operators we have the following estimate for the gradient

$$(\boldsymbol{\nabla} a)_\mu = \sum_\nu \boldsymbol{\Omega}_{\mu\nu} a_\nu \tag{28.3.2}$$

with

$$\boldsymbol{\Omega}_{\mu\nu} \equiv \int d\mathbf{r} \delta_\mu(\mathbf{r}) \boldsymbol{\nabla} \psi_\nu(\mathbf{r}) \tag{28.3.3}$$

Second order derivatives are

$$(\boldsymbol{\nabla}\boldsymbol{\nabla} a)_x \equiv D_x \left[\boldsymbol{\nabla}\boldsymbol{\nabla} C \left(a_x \right) \right] \tag{28.3.4}$$

which gives

$$(\boldsymbol{\nabla}\boldsymbol{\nabla} a)_\mu = -\sum_\nu \boldsymbol{\Delta}_{\mu\nu} a_\nu \tag{28.3.5}$$

with

$$\boldsymbol{\Delta}_{\mu\nu} \equiv - \int d\mathbf{r} \delta_\mu(\mathbf{r}) \boldsymbol{\nabla}\boldsymbol{\nabla} \psi_\nu(\mathbf{r})$$
$$= \int d\mathbf{r} \boldsymbol{\nabla}\delta_\mu(\mathbf{r}) \boldsymbol{\nabla} \psi_\nu(\mathbf{r}) \tag{28.3.6}$$

We may use the linear connection between the two sets of basis functions and write in compact matrix form the following identities

$$\mathbf{\Delta} = \mathbf{M}^\delta \mathbf{L}^\psi = \mathbf{L}^\delta \mathbf{M}^\psi \tag{28.3.7}$$

where we have introduced the stiffness matrices in each basis set

$$\mathbf{L}^\psi = \left(\boldsymbol{\nabla}\boldsymbol{\psi}, \boldsymbol{\nabla}\boldsymbol{\psi}^T\right)$$
$$\mathbf{L}^\delta = \left(\boldsymbol{\nabla}\boldsymbol{\delta}, \boldsymbol{\nabla}\boldsymbol{\delta}^T\right) \tag{28.3.8}$$

The relation between both stiffness matrices is

$$\mathbf{L}^\delta = \mathbf{M}^\delta \mathbf{L}^\psi \mathbf{M}^\delta$$
$$\mathbf{L}^\psi = \mathbf{M}^\psi \mathbf{L}^\delta \mathbf{M}^\psi \tag{28.3.9}$$

Note that the matrix $\mathbf{L}^\delta$ satisfies, thanks to (28.2.35) the following identities

$$\sum_\mu \mathbf{L}^\psi_{\mu\nu} = 0$$
$$\sum_\mu \mathscr{V}_\mu \mathbf{L}^\delta_{\mu\nu} = 0 \tag{28.3.10}$$

which shows that the vector $\boldsymbol{V} = (\mathscr{V}_1, \cdots, \mathscr{V}_N)$ is eigenvector of $\mathbf{L}^\delta$ of null eigenvalue.

In a 1D regular grid in which $x_\mu - x_{\mu-1} = a$ for every node μ the stiffness matrix will be given by

$$\mathbf{L}^\psi = \frac{1}{a}\begin{pmatrix} 2 & -1 & 0 & 0 & \dots & -1 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ & & & \ddots & & \end{pmatrix} \tag{28.3.11}$$

In principle, we may also define second derivatives in an alternative way as the result of two iterations of first derivatives. In particular, we could also define

$$(\boldsymbol{\nabla}\boldsymbol{\nabla}a)_x = D_x\left[\boldsymbol{\nabla}C\left((\boldsymbol{\nabla}a)_x\right)\right]$$
$$= D_x\left[\boldsymbol{\nabla}C\left(D_x\left[\boldsymbol{\nabla}C\left(a_x\right)\right]\right)\right] \tag{28.3.12}$$

Note the presence of the operator CD_x. As we have discussed previously, this operator *is not* the identity in general. It acts as the identity only when applied on exactly continued fields. Indeed, for exactly continued fields, we have $C\left(D_x\left[\boldsymbol{\nabla}C\left(a_x\right)\right]\right) = \boldsymbol{\nabla}C\left(a_x\right)$ and, therefore, (28.3.12) becomes again (28.3.4).

In terms of the discrete operators $\mathbf{\Omega}_{\mu\nu}$, the second derivative constructed iteratively is

$$(\boldsymbol{\nabla}\boldsymbol{\nabla}a)_\mu = \sum_{\nu\sigma} \boldsymbol{\Omega}_{\mu\nu}\boldsymbol{\Omega}_{\nu\sigma}a_\sigma \tag{28.3.13}$$

It is apparent that $\sum_\nu \boldsymbol{\Omega}_{\mu\nu}\boldsymbol{\Omega}_{\nu\sigma} \neq \boldsymbol{\Delta}_{\mu\sigma}$ in general. However, when applied onto exactly continued fields, both discretizations of second order derivatives coincide.

28.4   Partial differential equations

Up to now we have just a language to discretize and interpolate functions and its derivatives, in a local form. We use now this language to discretize partial differential equations on irregular meshes. A general method for discretizing partial differential equations is the Weighted Residual method [**?**]. We illustrate two procedures with the prototypical example is the diffusion equation

$$\partial_t c(\mathbf{r}, t) = \nabla^2 c(\mathbf{r}, t) \tag{28.4.1}$$

where we have taken units in which the diffusion coefficient takes the value 1. We will write this equation in the obvious form

$$\partial_t c(\mathbf{r}, t) - \nabla^2 c(\mathbf{r}, t) = 0 \tag{28.4.2}$$

28.4.1   Weighted residuals

The idea of weighted residuals is to use inside the partial differential equation an interpolated version of the function solution. This interpolated function depends on the discrete values of the solution. The interpolated function does not fullfill exactly the PDE. The error is called the residual and the method weights the residual with a set of weight functions that allow to obtain the discrete values of the solution in an algebraic manner.

Define $\mathbf{c}(t) = \{c_1(t), \cdots, c_M(t)\}$ as the discrete version of a field on a cloud of M points. Consider the interpolated field

$$\overline{c}(\mathbf{r}, t) = \boldsymbol{\psi}^T(\mathbf{r}) \cdot \mathbf{c}(t) \tag{28.4.3}$$

Substitution of this interpolated field into the diffusion equation allows to define the residual

$$R(\mathbf{c}, \dot{\mathbf{c}}, \mathbf{r}) = \boldsymbol{\psi}^T(\mathbf{r}) \cdot \dot{\overline{\mathbf{c}}} - \nabla^2 \boldsymbol{\psi}^T(\mathbf{r}) \cdot \overline{\mathbf{c}} \tag{28.4.4}$$

If the interpolated function $\overline{c}$ was an exact solution of the diffusion equation, the residual would be zero, but this is not the case. The idea is to use a set of M weight functions $\omega_\mu(\mathbf{r})$ (denoted collectively with the vector $\boldsymbol{\omega}(\mathbf{r}) = \{\omega_1(\mathbf{r}), \cdots, \omega_M(\mathbf{r})\}$) and integrate over the whole domain in order to obtain M conditions that allow to obtain a set of

ODE for the vector $\mathbf{c}(t)$. In this way, we have

$$\left(\boldsymbol{\psi}, \boldsymbol{\omega}^T\right) \dot{\overline{\mathbf{c}}} = \left(\nabla^2 \boldsymbol{\psi}, \boldsymbol{\omega}^T\right) \cdot \overline{\mathbf{c}} \tag{28.4.5}$$

Depending on the choice of the weight functions $\omega_\mu(\mathbf{r})$ we obtain different methods of discretization. The collocation method takes $\omega_\mu(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_\mu)$, the least-square Galerkin method uses

$$\omega_\mu(\mathbf{r}) = \frac{\partial R}{\partial c_\mu}(\mathbf{r}) \tag{28.4.6}$$

in such a way that the solution $\mathbf{c}$ is, in fact, the one that minimizes the square of the residual error over the whole domain

$$I(\mathbf{c}) = \int d\mathbf{r} R(\mathbf{r})^2 \tag{28.4.7}$$

The Galerkin method uses $\omega_\mu(\mathbf{r}) = \psi_\mu(\mathbf{r})$ in order to make the residual function orthogonal (in functional sense) to the basis functions. If the basis functions form a complete set of the functional space, then the residual would be exactly zero. In the Galerkin

$$\mathbf{M}^\psi \dot{\mathbf{c}}_\nu(t) = -\mathbf{L}^\psi \mathbf{c} \tag{28.4.8}$$

If we choose as weight functions the consistent basis function $\boldsymbol{\omega}(\mathbf{r}) = \boldsymbol{\delta}(\mathbf{r})$ we have

$$\dot{\overline{\mathbf{c}}} = -\boldsymbol{\Delta}\overline{\mathbf{c}} \tag{28.4.9}$$

where the discrete Laplacian is given by (28.3.6). When the basis functions are related linearly, the resulting discretized equation (**??**) is in compact form

$$\frac{d}{dt}\mathbf{c}(t) = -\mathbf{M}^\delta \mathbf{L}^\psi \mathbf{c}(t) \tag{28.4.10}$$

which can also be writen as

$$\mathbf{M}^\psi \frac{d}{dt}\mathbf{c}(t) = -\mathbf{L}^\psi \mathbf{c}(t) \tag{28.4.11}$$

which is exactly the finite element discretization of the diffusion equation. In this way, we may also write Eq. (28.4.10) in the form

$$\frac{d}{dt}\mathbf{c}(t) = -\mathbf{L}^\delta \mathbf{M}^\psi \mathbf{c}(t) \tag{28.4.12}$$

We see that in this case, the method is identical to the Galerkin method.

28.5   Non-linear diffusion

Let us move to a more interesting situation that involves the following non-linear diffusion equation

$$\partial_t c(\mathbf{r}, t) = \boldsymbol{\nabla} \mu(c(\mathbf{r}, t)) \boldsymbol{\nabla} \frac{\delta \mathscr{F}}{\delta c(\mathbf{r})}[c] \tag{28.5.1}$$

where $\mathscr{F}[c]$ is the free energy functional and $\mu(c)$ is a mobility coefficiente given by

$$\mu(c(\mathbf{r}, t)) = \frac{D(c(\mathbf{r}, t))}{k_B T} c(\mathbf{r}, t) \tag{28.5.2}$$

with D the diffusion coefficient. This equation is paradigmatic in that it captures two essential features:

1. Conservation of the number of particles

2. H-theorem

In fact, it is immediate to show that, in an infinite or periodic system, the total number $N = \int d\mathbf{r} c(\mathbf{r}, t)$ is conserved. The time derivative of $\mathscr{F}[c]$ is always negative provided that the diffusion coefficient $D(c)$ is positive.

We will be considering free energy functionals under the local approximation, this is, of the form

$$\mathscr{F}[c] = \int d\mathbf{r} f(c(\mathbf{r})) \tag{28.5.3}$$

where $f(c)$ is the free energy *density*. In this case, the diffusion equation becomes

$$\partial_t c(\mathbf{r}, t) = \boldsymbol{\nabla} \mu(c(\mathbf{r}, t)) \boldsymbol{\nabla} \lambda(c(\mathbf{r}, t)) \tag{28.5.4}$$

where $\lambda(c) = f'(c)$.

The residual of the PDE (28.5.1) is obtained after substitution of the interpolated or trial function $\overline{c}(\mathbf{r}, t) = \boldsymbol{\psi}^T(\mathbf{r}) \mathbf{c}(t)$

$$R(\mathbf{r}) = \partial_t \overline{c}(\mathbf{r}, t) - \boldsymbol{\nabla} \mu(\overline{c}(\mathbf{r}, t)) \boldsymbol{\nabla} \frac{\delta \mathscr{F}}{\delta c(\mathbf{r})}[\overline{c}] \tag{28.5.5}$$

By weigthing the residual with weights $\boldsymbol{\omega}(\mathbf{r})$ and requiring the weighted residual to vanish we obtain

$$0 = (\boldsymbol{\omega}, R)$$
$$= \partial_t (\boldsymbol{\omega}, \overline{c}(t)) - \left(\boldsymbol{\omega}, \boldsymbol{\nabla} \mu(\overline{c}(\mathbf{r}, t)) \boldsymbol{\nabla} \frac{\delta \mathscr{F}}{\delta c} [\overline{c}(t)] \right) \tag{28.5.6}$$

Formally, this is a set of M ordinary differential equations for the M unknowns $\mathbf{c}(t)$. It is apparent that we cannot proceed until we have a way to compute the functional

derivative $\frac{\delta \mathscr{F}}{\delta c}$. To this end *define* the discrete free energy function $F(\mathbf{c})$ as

$$F(\mathbf{c}) \equiv \mathscr{F}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \tag{28.5.7}$$

this is, the free energy *function* of the discrete field is obtained by evaluating the free energy *functional* at the interpolated field. For example, for a local free energy functional, we have

$$F(\mathbf{c}) = \int d\mathbf{r} f\left(\boldsymbol{\psi}^T(\mathbf{r})\mathbf{c}\right) \tag{28.5.8}$$

But what we need is not only a discrete approximation for the functional, but also a discrete approximation for its functional derivative. To this end, let us compute the derivative of the function (28.5.7)

$$\frac{\partial F}{\partial c_\mu}(\mathbf{c}) = \int d\mathbf{r}' \frac{\delta \mathscr{F}}{\delta c(\mathbf{r}')}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \psi_\mu(\mathbf{r}') \tag{28.5.9}$$

where the functional chain rule has been used. Let us multiply Eq. (28.5.9) with a set $\overline{\boldsymbol{\varpi}}(\mathbf{r})$ that we will later particularize to be $\boldsymbol{\delta}(\mathbf{r})$ or $\boldsymbol{\psi}(\mathbf{r})$. In this way, we get

$$\overline{\boldsymbol{\varpi}}^T(\mathbf{r})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) = \int d\mathbf{r}' \frac{\delta \mathscr{F}}{\delta c(\mathbf{r}')}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \mathscr{D}(\mathbf{r}, \mathbf{r}') \tag{28.5.10}$$

where we have introduced the smoothing function $\mathscr{D}(\mathbf{r}, \mathbf{r}')$ as

$$\mathscr{D}(\mathbf{r}, \mathbf{r}') = \overline{\boldsymbol{\varpi}}^T(\mathbf{r})\boldsymbol{\psi}(\mathbf{r}') \tag{28.5.11}$$

If the functional derivative does not change appreciably within the range of $\mathscr{D}(\mathbf{r}, \mathbf{r}')$, we may simply write from Eq. (28.5.10)

$$\int d\mathbf{r}' \frac{\delta \mathscr{F}}{\delta c(\mathbf{r}')}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \mathscr{D}(\mathbf{r}, \mathbf{r}') = \frac{\delta \mathscr{F}}{\delta c(\mathbf{r})}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \int d\mathbf{r}' \mathscr{D}(\mathbf{r}, \mathbf{r}')$$
$$= \frac{\delta \mathscr{F}}{\delta c(\mathbf{r})}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \sum_\mu \mathscr{V}_\mu \overline{\varpi}_\mu(\mathbf{r}) \tag{28.5.12}$$

and, therefore, we have an approximate expression for the functional derivative

$$\frac{\delta \mathscr{F}}{\delta c(\mathbf{r})}\left[\boldsymbol{\psi}^T \mathbf{c}\right] \approx \frac{\overline{\boldsymbol{\varpi}}^T(\mathbf{r})}{\sum_\mu \mathscr{V}_\mu \overline{\varpi}_\mu(\mathbf{r})} \frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \tag{28.5.13}$$

Now that we have an approximation for the functional derivative, we may introduce (28.5.13) into (28.5.6) and obtain

$$(\boldsymbol{\omega}, \boldsymbol{\psi}) \, \dot{\mathbf{c}}(t) = -\mathbf{D}^{\omega\overline{\omega}}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \tag{28.5.14}$$

where the dissipative matrix $\mathbf{D}^{\omega\overline{\omega}}(\mathbf{c})$ is defined through

$$\mathbf{D}^{\omega\overline{\omega}}_{\mu\nu}(\mathbf{c}) \equiv \int d\mathbf{r} \boldsymbol{\nabla}\omega_\mu(\mathbf{r})\mu(c(\mathbf{r}))\boldsymbol{\nabla}\left(\frac{\overline{\omega}_\nu(\mathbf{r})}{\sum_\mu \mathscr{V}_{\mu'}\overline{\omega}_{\mu'}(\mathbf{r})}\right) \tag{28.5.15}$$

We have now the freedom to choose for the basis functions $\boldsymbol{\omega}$ and $\overline{\boldsymbol{\omega}}$ either $\boldsymbol{\delta}$ or $\boldsymbol{\psi}$. Of the four different possibilities it is apparent that the only sensible one is $\overline{\boldsymbol{\omega}} = \boldsymbol{\delta}$ because then

$$\frac{\overline{\omega}_\nu(\mathbf{r})}{\sum_\mu \mathscr{V}_\mu\overline{\omega}_\mu(\mathbf{r})} = \delta_\nu(\mathbf{r}) \tag{28.5.16}$$

In this case, the smoothing function $\mathscr{D}(\mathbf{r}, \mathbf{r}')$ defined in Eq. (28.5.11) is precisely $D(\mathbf{r}, \mathbf{r}')$ introduced in Eq. (28.2.19).

For the basis function $\boldsymbol{\omega}$ we may choose either $\boldsymbol{\delta}$ or $\boldsymbol{\psi}$, as both give exactly the same final dynamic equation. For example, with the Galerkin prescription $\boldsymbol{\omega} = \boldsymbol{\psi}$ we have in Eq. (28.5.14)

$$\mathbf{M}^\psi \dot{\mathbf{c}}(t) = -\mathbf{D}^{\psi\delta}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \tag{28.5.17}$$

which is

$$\begin{aligned}
\dot{\mathbf{c}}(t) &= -\mathbf{M}^\delta \mathbf{D}^{\psi\delta}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \\
&= -\mathbf{D}^{\delta\delta}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c})
\end{aligned} \tag{28.5.18}$$

where with an obvious notation we have introduced

$$\mathbf{D}^{\delta\delta}(\mathbf{c}) = \int d\mathbf{r}\boldsymbol{\nabla}\delta_\mu(\mathbf{r})\mu(c(\mathbf{r}))\boldsymbol{\nabla}\delta_\nu(\mathbf{r}) \tag{28.5.19}$$

On the other hand, if instead in Eq. (28.5.14) we use $\boldsymbol{\omega} = \boldsymbol{\psi}$ we again obtain exactly the same equation (28.5.18) which we now simply write as

$$\dot{\mathbf{c}}(t) = -\mathbf{D}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \tag{28.5.20}$$

where now the matrix

$$\mathbf{D}_{\mu\nu}(\mathbf{c}) = \int d\mathbf{r}\boldsymbol{\nabla}\delta_\mu(\mathbf{r})\mu\left(\sum_\sigma \psi_\sigma(\mathbf{r})c_\sigma\right)\boldsymbol{\nabla}\delta_\nu(\mathbf{r}) \tag{28.5.21}$$

When $\mu(\mathbf{c}) = \mu_0$ is a constant mobility coefficient we simply have

$$\mathbf{D}_{\mu\nu}(\mathbf{c}) = \mu_0 \mathbf{L}^\delta_{\mu\nu} \tag{28.5.22}$$

Note that we have used the interpolated discrete field as argument in the diffusion coefficient within the integral in Eq. (28.5.15) and, therefore, the dissipative matrix depends on $\mathbf{c}$. The matrix $\mathbf{D}_{\mu\nu}(\mathbf{c})$ is symmetric and positive semidefinite. For this reason, the time derivative of the discrete free energy $F(\mathbf{c}(t))$, which is given by

$$\frac{dF}{dt}(\mathbf{c}(t)) = -\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c})\mathbf{D}(\mathbf{c})\frac{\partial F}{\partial \mathbf{c}}(\mathbf{c}) \tag{28.5.23}$$

is always negative or zero.

The only assumption that we have taken is that the functional derivative of the free energy hardly changes in the range of $D(\mathbf{r}, \mathbf{r}')$ defined in (28.2.19). We have plotted this function in Fig. 28.3 and, therefore, we observe that if the average lattice spacing is much smaller than the length scale of variation of the field, then the approximation (28.5.13) will be fulfilled.

Figure 28.3: Plot of $\mathscr{D}(\mathbf{r}, \mathbf{r}')$ in a regular lattice of size $a = 1/20.$. The function $\mathscr{D}(\mathbf{r}, \mathbf{r}')$ is non-vanishing only for points differing by a few (2 or 3) lattice spacings.

28.6   Isothermal hydrodynamics

For continuum hydrodynamics we have the reversible part given as

$$\partial_t \rho_{\mathbf{r}} = -\nabla \mathbf{g}_r$$
$$\partial_r \mathbf{g_r} = -\nabla \mathbf{g} \mathbf{v} - \nabla P \tag{28.6.1}$$

The CG Hamiltonian is

$$\hat{H} = F[\rho] + \int d\mathbf{r} \frac{\mathbf{g}_{\mathbf{r}}^2}{2\rho_{\mathbf{r}}} \tag{28.6.2}$$

where $F[\rho]$ is the free energy functional. It is taken of the local form

$$F[\rho] = \int d\mathbf{r} f(\rho_{\mathbf{r}}) \tag{28.6.3}$$

where $f(\rho)$ is the free energy density. Note that we may write the pressure gradient in the form

$$\boldsymbol{\nabla} P = \rho_{\mathbf{r}} \nabla \frac{\delta F}{\delta \rho_{\mathbf{r}}} [\rho] \tag{28.6.4}$$

on due account of the identity

$$P(\rho) = f(\rho) - \rho f'(\rho) \tag{28.6.5}$$

The functional (28.6.2) is conserved by the reversible dynamics (28.6.1). The functional derivatives are given by

$$\frac{\delta \hat{H}}{\delta \rho_{\mathbf{r}}} = \frac{\delta \mathscr{F}}{\delta \rho_{\mathbf{r}}} [\rho] - \frac{\mathbf{g}_{\mathbf{r}}^2}{2\rho_{\mathbf{r}}^2}$$
$$\frac{\delta \hat{H}}{\delta \mathbf{g_r}} = \frac{\mathbf{g_r}}{\rho_{\mathbf{r}}} \tag{28.6.6}$$

We may write the reversible equation (28.6.1) in the form

$$\partial_t \rho_{\mathbf{r}} = -\nabla \mathbf{g}_r$$
$$\partial_r \mathbf{g_r} = -\nabla \mathbf{g} \mathbf{v} - \rho \nabla \frac{\mathbf{v}^2}{2} + \mathbf{g} \nabla \mathbf{v} - \nabla P \tag{28.6.7}$$

And in matrix form

$$
\begin{pmatrix} \partial_t \rho_{\mathbf{r}} \\ \partial_t \mathbf{g}_{\mathbf{r}}^{\alpha} \end{pmatrix} = \int d\mathbf{r}' \begin{pmatrix} 0 & -\rho_{\mathbf{r}'}(\boldsymbol{\nabla}_{\mathbf{r}}^{\beta}\delta_{\mathbf{rr}'}) \\ \rho_{\mathbf{r}}\left(\boldsymbol{\nabla}_{\mathbf{r}'}^{\alpha}\delta_{\mathbf{rr}'}\right) & \left(\boldsymbol{\nabla}_{\mathbf{r}'}^{\beta}\delta_{\mathbf{rr}'}\right)\mathbf{g}_{\mathbf{r}'}^{\alpha} - \mathbf{g}_{\mathbf{r}}^{\beta}\left(\boldsymbol{\nabla}_{\mathbf{r}}^{\alpha}\delta_{\mathbf{rr}'}\right) \end{pmatrix} \begin{pmatrix} \frac{\delta\hat{H}}{\delta\rho_{\mathbf{r}'}} \\ \frac{\delta\hat{H}}{\delta\mathbf{g}_{\mathbf{r}'}^{\beta}} \end{pmatrix}
$$

$$(28.6.8)$$

where $\delta_{\mathbf{rr}'}$ is the Dirac delta function. The equations (28.6.8) are writen in compact form

$$
\partial_t a_{\mathbf{r}} = \int d\mathbf{r} L_{\mathbf{rr}'} \frac{\delta\hat{H}}{\delta a_{\mathbf{r}'}}
\tag{28.6.9}
$$

28.6.1   Discretizing the equations

Now, let us introduce the CG Dirac delta function

$$D(\mathbf{r}, \mathbf{r}') = \psi_\sigma(\mathbf{r})\delta_\sigma(\mathbf{r}') \qquad (28.6.10)$$

that satisfies the linear consistent relations

$$\int d\mathbf{r} D(\mathbf{r}, \mathbf{r}') = 1$$

$$\int d\mathbf{r}\mathbf{r} D(\mathbf{r}, \mathbf{r}') = \mathbf{r}' \qquad (28.6.11)$$

In this way it behaves exactly like the ordinary Dirac delta function for linear functions. For non-linear functions that change slowly on the cell size, the difference with the Dirac delta function is small.

In order to discretize 28.6.9) we choose to multiply with $\delta_\mu(\mathbf{r})$, integrate over $\mathbf{r}$ and introduce the unity as follows

$$\partial_t a_\mu = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \underbrace{\int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta\hat{H}}{\delta a_{\mathbf{r}'}}}_{1} \qquad (28.6.12)$$

We now make the physical assumption that the intensive fields $\frac{\delta\hat{H}}{\delta a_{\mathbf{r}'}}$ are smooth on the length scale of the cells, i.e. in the length scale of $D(\mathbf{r}, \mathbf{r}')$. If this is the case, then we may approximate

$$\frac{\delta\hat{H}}{\delta a_{\mathbf{r}'}} \approx \int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta\hat{H}}{\delta a_{\mathbf{r}''}} \qquad (28.6.13)$$

in such a way that

$$\partial_t a_\mu \approx \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta\hat{H}}{\delta a_{\mathbf{r}''}} = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \int d\mathbf{r}'' \psi_\nu(\mathbf{r}'') \delta_\nu(\mathbf{r}') \frac{\delta\hat{H}}{\delta a_{\mathbf{r}''}}$$

$$= \underbrace{\int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \delta_\nu(\mathbf{r}')}_{L_{\mu\nu}} \int d\mathbf{r}'' \psi_\nu(\mathbf{r}'') \frac{\delta\hat{H}}{\delta a_{\mathbf{r}''}}$$

$$= L_{\mu\nu} \frac{\partial\hat{H}}{\partial a_\nu}(a) \qquad (28.6.14)$$

with the discrete Hamiltonian described by

$$H(a) = \hat{H}[\psi a] \qquad (28.6.15)$$

In order to have a closed equation, any dependence on the field $a_\mathbf{r}$ in $L_{\mathbf{r}\mathbf{r}'}$ needs to be transformed into the interpolated form $a_\mathbf{r} \approx \psi_\mu(\mathbf{r})a_\mu$. After this assumption, Eq. (28.7.10) becomes a closed equation for the discrete values a_μ of the hydrodynamic vari-

ables.

If we apply this discretization scheme on Eq. (28.6.8) we obtain

$$
\begin{pmatrix} \partial_t \rho_\mu \\[2mm] \partial_t \mathbf{g}_\mu^\alpha \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{L}_{\mu\nu}^\beta \\[2mm] -\mathcal{L}_{\nu\mu}^\alpha & \mathscr{G}_{\mu\nu}^{\alpha\beta} - \mathscr{G}_{\nu\mu}^{\beta\alpha} \end{pmatrix} \begin{pmatrix} \frac{\partial \hat{H}}{\partial \rho_\nu} \\[2mm] \frac{\partial \hat{H}}{\partial \mathbf{g}_\nu^\beta} \end{pmatrix}
\tag{28.6.16}
$$

where

$$
\mathcal{L}_{\mu\nu}^\alpha \equiv -\int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) \delta_\nu(\mathbf{r}') \left(\boldsymbol{\nabla}_{\mathbf{r}}^\alpha \delta_{\mathbf{r}\mathbf{r}'} \right) \rho_{\mathbf{r}'} = \int d\mathbf{r} \rho_{\mathbf{r}} \delta_\nu(\mathbf{r}) \boldsymbol{\nabla}^\alpha \delta_\mu(\mathbf{r})
$$

$$
\mathscr{G}_{\mu\nu}^{\alpha\beta} \equiv \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) \delta_\nu(\mathbf{r}') \mathbf{g}_{\mathbf{r}'}^\alpha \left(\boldsymbol{\nabla}_{\mathbf{r}'}^\beta \delta_{\mathbf{r}\mathbf{r}'} \right) = \int d\mathbf{r} \mathbf{g}_{\mathbf{r}}^\alpha \delta_\nu(\mathbf{r}) \boldsymbol{\nabla}^\beta \delta_\mu(\mathbf{r})
\tag{28.6.17}
$$

Let us be more explicit now. The kinetic energy in the discrete CG Hamiltonian becomes, after substituting the linearly interpolated fields into the kinetic energy functional, the following kinetic function

$$
K(\rho, \mathbf{g}) = \frac{1}{2} \int d\mathbf{r} \mathbf{g}_\mu \psi_\mu(\mathbf{r}) \frac{1}{\psi_\sigma(\mathbf{r})\rho_\sigma} \mathbf{g}_\nu \psi_\nu(\mathbf{r}) = \frac{1}{2} \mathbf{g}_\mu \mathscr{A}_{\mu\nu}(\rho) \mathbf{g}_\nu
\tag{28.6.18}
$$

where we have introduced the density dependent matrix

$$
\mathscr{A}_{\mu\nu}(\rho) = \int d\mathbf{r} \psi_\mu(\mathbf{r}) \frac{1}{[\psi_\sigma(\mathbf{r})\rho_\sigma]} \psi_\nu(\mathbf{r})
\tag{28.6.19}
$$

Consider the matrix $\mathscr{M}_{\mu\nu}$ of the form

$$
\mathscr{M}_{\mu\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma \right] \delta_\nu(\mathbf{r})
\tag{28.6.20}
$$

The matrices $\mathscr{A}$ and $\mathscr{M}$ are approximate inverses of each other, this is

$$
\sum_\alpha \mathscr{A}_{\mu\alpha}(\rho) \mathscr{M}_{\alpha\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma \right] \sum_\alpha \delta_\alpha(\mathbf{r}) \int d\mathbf{r}' \psi_\alpha(\mathbf{r}') \frac{1}{[\psi_{\sigma'}(\mathbf{r}')\rho_{\sigma'}]} \psi_\nu(\mathbf{r}')
$$

$$
= \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma \right] D(\mathbf{r}, \mathbf{r}') \frac{1}{[\psi_{\sigma'}(\mathbf{r}')\rho_{\sigma'}]} \psi_\nu(\mathbf{r}')
\tag{28.6.21}
$$

If the interpolated density field changes slowly in the cell length scale, we may simply have

$$
\sum_\alpha \mathscr{A}_{\mu\alpha}(\rho) \mathscr{M}_{\alpha\nu}(\rho) = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) D(\mathbf{r}, \mathbf{r}') \psi_\nu(\mathbf{r}') = \delta_{\mu\nu}
\tag{28.6.22}
$$

where the last identity is exact when applied to exactly continued discrete fields. There-

fore, we may model the kinetic energy term as simply

$$K(\rho, \mathbf{g}) = \frac{1}{2} \mathbf{g}_\mu \mathcal{M}_{\mu\nu}^{-1} \mathbf{g}_\nu$$

$$\mathcal{M}_{\mu\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r}) \rho_\sigma \right] \delta_\nu(\mathbf{r}) \tag{28.6.23}$$

The CG Hamiltonian is, therefore,

$$H(\rho, \mathbf{g}) = \frac{1}{2} \mathbf{g}_\mu \mathcal{M}_{\mu\nu}^{-1} \mathbf{g}_\nu + F(\rho) \tag{28.6.24}$$

where $F(\rho) = \mathscr{F}[\psi\rho]$ is the discrete approximation to the free energy. The derivatives of the discrete Hamiltonian are

$$\frac{\partial \hat{H}}{\partial \rho_\mu} = \frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) + \frac{\partial F}{\partial \rho_\mu}(\rho)$$

$$\frac{\partial \hat{H}}{\partial \mathbf{g}_\mu} = \mathcal{M}_{\mu\nu}^{-1} \mathbf{g}_\nu \tag{28.6.25}$$

where the density derivative of the kinetic energy is

$$\frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) = \frac{1}{2} \mathbf{g}_{\mu'}^\alpha \frac{\partial \mathcal{M}_{\mu'\nu'}^{-1}}{\partial \rho_\mu} \mathbf{g}_{\nu'}^\alpha$$

$$= \frac{1}{2} \mathbf{g}_{\mu'}^\alpha \mathcal{M}_{\mu'\alpha}^{-1} \frac{\partial \mathcal{M}_{\alpha\beta}}{\partial \rho_\mu} \mathcal{M}_{\beta\nu'}^{-1} \mathbf{g}_{\nu'}^\alpha = \frac{1}{2} \mathbf{v}_{\mu'}^\alpha \frac{\partial \mathcal{M}_{\mu'\nu'}}{\partial \rho_\mu} \mathbf{v}_{\nu'}^\alpha$$

$$= \frac{1}{2} \mathbf{g}_{\mu'}^\alpha \mathcal{M}_{\mu'\alpha}^{-1} \int d\mathbf{r} \delta_\alpha(\mathbf{r}) \psi_\mu(\mathbf{r}) \delta_\beta(\mathbf{r}) \mathcal{M}_{\beta\nu'}^{-1} \mathbf{g}_{\nu'}^\alpha \tag{28.6.26}$$

The dynamic equations after the model Hamiltonian (28.7.20) are

$$\begin{pmatrix} \partial_t \rho_\mu \\ \partial_t \mathbf{g}_\mu^\alpha \end{pmatrix} = \begin{pmatrix} 0 & \mathcal{L}_{\mu\nu}^\beta \\ -\mathcal{L}_{\nu\mu}^\alpha & \mathcal{G}_{\mu\nu}^{\alpha\beta} - \mathcal{G}_{\nu\mu}^{\beta\alpha} \end{pmatrix} \begin{pmatrix} \frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) + \frac{\partial F}{\partial \rho_\nu}(\rho) \\ \mathcal{M}_{\nu\nu'}^{-1} \mathbf{g}_{\nu'}^\beta \end{pmatrix} \tag{28.6.27}$$

Leading to the following equations

$$\partial_t \rho_\mu = \mathcal{L}_{\mu\nu}^\beta \mathcal{M}_{\nu\nu'}^{-1} \mathbf{g}_{\nu'}^\beta$$

$$\partial_t \mathbf{g}_\mu^\alpha = -\mathcal{L}_{\nu\mu}^\alpha \frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) - \mathcal{L}_{\nu\mu}^\alpha \frac{\partial F}{\partial \rho_\nu}(\rho) + (\mathcal{G}_{\mu\nu}^{\alpha\beta} - \mathcal{G}_{\nu\mu}^{\beta\alpha}) \mathbf{v}_\nu^\beta \tag{28.6.28}$$

$$\mathcal{L}_{\nu\mu}^\alpha = \int d\mathbf{r} \rho_\mathbf{r} \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r}) = \rho_\sigma \int d\mathbf{r} \psi_\sigma(\mathbf{r}) \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r})$$

$$\mathcal{G}_{\mu\nu}^{\alpha\beta} = \int d\mathbf{r} \mathbf{g}_\mathbf{r}^\alpha \delta_\nu(\mathbf{r}) \boldsymbol{\nabla}^\beta \delta_\mu(\mathbf{r}) \tag{28.6.29}$$

Now, let us use the definitions (28.7.13)

$$\partial_t \mathbf{g}_\mu^\alpha = -\mathcal{L}_{\nu\mu}^\alpha \frac{\partial F}{\partial \rho_\nu}(\rho) + \int d\mathbf{r} \mathbf{g}_\mathbf{r}^\alpha \delta_\nu(\mathbf{r}) \nabla^\beta \delta_\mu(\mathbf{r}) \mathbf{v}_\nu^\beta$$

$$- \mathcal{L}_{\nu\mu}^\alpha \frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) - \int d\mathbf{r} \mathbf{g}_\mathbf{r}^\beta \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r}) \mathbf{v}_\nu^\beta \qquad (28.6.30)$$

Therefore

$$\partial_t \mathbf{g}_\mu^\alpha = -\underbrace{\int d\mathbf{r} \rho_\mathbf{r} \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{\partial F}{\partial \rho_\nu}(\rho)}_{\text{pressure gradient}} + \underbrace{\int d\mathbf{r} \mathbf{g}_\mathbf{r}^\alpha \delta_\nu(\mathbf{r}) \nabla^\beta \delta_\mu(\mathbf{r}) \mathbf{v}_\nu^\beta}_{\text{convective}}$$

$$- \int d\mathbf{r} \rho_\mathbf{r} \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{1}{2} \mathbf{v}_{\mu'}^\beta \frac{\partial \mathcal{M}_{\mu'\nu'}}{\partial \rho_\mu} \mathbf{v}_{\nu'}^\beta - \int d\mathbf{r} \mathbf{g}_\mathbf{r}^\beta \delta_\mu(\mathbf{r}) \nabla^\alpha \delta_\nu(\mathbf{r}) \mathbf{v}_\nu^\beta \qquad (28.6.31)$$

28.6.2   Mass and momentum conservation

The nice thing about this discretization scheme is that the matrix $L_{\mu\nu}$ is antisymmetric and, therefore, the CG Hamiltonian $H(a)$ is conserved by the dynamics in the same way as $\hat{H}[a]$ is conserved by the continuum dynamics. The total mass of the fluid

$$M_T \equiv \sum_\mu \mathcal{V}_\mu \rho_\mu \qquad (28.6.32)$$

is also nicely conserved. This follows from the identity

$$\sum_\mu \mathcal{V}_\mu \mathcal{L}_{\mu\nu} = \int d\mathbf{r} \rho_\mathbf{r} \delta_\nu(\mathbf{r}) \nabla \underbrace{\sum_\mu \mathcal{V}_\mu \delta_\mu(\mathbf{r})}_{1} = 0 \qquad (28.6.33)$$

Unfortunately, total momentum is not conserved in general. For this to happen, we would need that

$$0 = -\int d\mathbf{r} \rho_\mathbf{r} \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{\partial F}{\partial \rho_\nu}(\rho) - \int d\mathbf{r} \rho_\mathbf{r} \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{1}{2} \mathbf{v}_{\mu'}^\beta \frac{\partial \mathcal{M}_{\mu'\nu'}}{\partial \rho_\mu} \mathbf{v}_{\nu'}^\beta - \int d\mathbf{r} \mathbf{g}_\mathbf{r}^\beta \nabla^\alpha \delta_\nu(\mathbf{r}) \mathbf{v}_\nu^\beta$$

$$(28.6.34)$$

Because we expect that the terms with velocities need to vanish independently of the rest, we require, in fact the two equations

$$\int d\mathbf{r} \rho_\mathbf{r} \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{\partial F}{\partial \rho_\nu}(\rho) = 0$$

$$\int d\mathbf{r} \rho_\mathbf{r} \nabla^\alpha \delta_\nu(\mathbf{r}) \frac{1}{2} \mathbf{v}^\beta_{\mu'} \frac{\partial \mathcal{M}_{\mu'\nu'}}{\partial \rho_\mu} \mathbf{v}^\beta_{\nu'} = \int d\mathbf{r} \mathbf{g}^\beta_\mathbf{r} \nabla^\alpha \delta_\nu(\mathbf{r}) \mathbf{v}^\beta_\nu \qquad (28.6.35)$$

These are requirements for the way in which we may approximate $\rho_\mathbf{r}, \mathbf{g}_\mathbf{r}$ and $F(\rho)$ in terms of $\rho_\mu, \mathbf{g}_\mu$.

28.7   Non-isothermal hydrodynamics

Consider the reversible part of the dynamics for a non-isothermal fluid

$$\partial_t \rho_{\mathbf{r}} = -\nabla \rho_{\mathbf{r}} \mathbf{v}_r$$
$$\partial_t \mathbf{g_r} = -\nabla \left(\mathbf{g_r} \mathbf{v}_r \right) - \nabla \hat{P}_{\mathbf{r}}$$
$$\partial_t \epsilon_{\mathbf{r}} = -\nabla \left(\epsilon_{\mathbf{r}} \mathbf{v}_r \right) - P_{\mathbf{r}} \nabla \mathbf{v}_r \tag{28.7.1}$$

Introduce the energy functional

$$E(\rho, \mathbf{g}, \epsilon) = \int d\mathbf{r} \left[\frac{\mathbf{g_r^2}}{2\rho_{\mathbf{r}}} + \epsilon_{\mathbf{r}} \right] \tag{28.7.2}$$

with functional derivatives

$$\frac{\delta E}{\delta a_{\mathbf{r}'}} \rightarrow \begin{pmatrix} -\frac{\mathbf{v}_{\mathbf{r}'}^2}{2} \\ \mathbf{v}_{\mathbf{r}'} \\ 1 \end{pmatrix} \tag{28.7.3}$$

We may write the above equations as

$$\partial a_{\mathbf{r}} = \int d\mathbf{r}' L_{\mathbf{r}\mathbf{r}'} \frac{\delta E}{\delta a_{\mathbf{r}'}} \tag{28.7.4}$$

$$L_{\mathbf{r}\mathbf{r}'} = \begin{pmatrix} 0 & \rho_{\mathbf{r}'} \nabla' \delta_{\mathbf{r}\mathbf{r}'} & 0 \\ \rho_{\mathbf{r}} \nabla' \delta_{\mathbf{r}\mathbf{r}'} & \mathbf{g}_{\mathbf{r}'} \nabla' \delta_{\mathbf{r}\mathbf{r}'} + \nabla' \delta_{\mathbf{r}\mathbf{r}'} \mathbf{g_r} & \hat{\epsilon}_{\mathbf{r}} \nabla' \delta_{\mathbf{r}\mathbf{r}'} + \hat{P}_{\mathbf{r}'} \nabla' \delta_{\mathbf{r}\mathbf{r}'} \\ 0 & \hat{\epsilon}_{\mathbf{r}'} \nabla' \delta_{\mathbf{r}\mathbf{r}'} + \hat{P}_{\mathbf{r}} \nabla' \delta_{\mathbf{r}\mathbf{r}'} & 0 \end{pmatrix} \tag{28.7.5}$$

28.7.1   Discretizing the equations

Now, let us introduce the CG Dirac delta function

$$D(\mathbf{r}, \mathbf{r}') = \psi_\sigma(\mathbf{r}) \delta_\sigma(\mathbf{r}') \tag{28.7.6}$$

that satisfies the linear consistent relations

$$\int d\mathbf{r} D(\mathbf{r}, \mathbf{r}') = 1$$
$$\int d\mathbf{r} \mathbf{r} D(\mathbf{r}, \mathbf{r}') = \mathbf{r}' \tag{28.7.7}$$

In this way it behaves exactly like the ordinary Dirac delta function for linear functions. For non-linear functions that change slowly on the cell size, the difference with the Dirac delta function is small.

In order to discretize 28.6.9) we choose to multiply with $\delta_\mu(\mathbf{r})$, integrate over $\mathbf{r}$ and introduce the unity as follows

$$\partial_t a_\mu = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \underbrace{\int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta E}{\delta a_{\mathbf{r}'}}}_{1} \tag{28.7.8}$$

We now make the physical assumption that the intensive fields $\frac{\delta E}{\delta a_{\mathbf{r}'}}$ are smooth on the length scale of the cells, i.e. in the length scale of $D(\mathbf{r}, \mathbf{r}')$. If this is the case, then we may approximate

$$\frac{\delta E}{\delta a_{\mathbf{r}'}} \approx \int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta E}{\delta a_{\mathbf{r}''}} \tag{28.7.9}$$

in such a way that

$$\partial_t a_\mu \approx \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \int d\mathbf{r}'' D(\mathbf{r}'', \mathbf{r}') \frac{\delta E}{\delta a_{\mathbf{r}''}} = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \int d\mathbf{r}'' \psi_\nu(\mathbf{r}'') \delta_\nu(\mathbf{r}') \frac{\delta E}{\delta a_{\mathbf{r}''}}$$

$$= \underbrace{\int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) L_{\mathbf{r}\mathbf{r}'} \delta_\nu(\mathbf{r}')}_{L_{\mu\nu}} \int d\mathbf{r}'' \psi_\nu(\mathbf{r}'') \frac{\delta E}{\delta a_{\mathbf{r}''}}$$

$$= L_{\mu\nu} \frac{\partial E}{\partial a_\nu}(a) \tag{28.7.10}$$

with the discrete energy function defined as

$$E(a) = E[\psi a] \tag{28.7.11}$$

In order to have a closed equation, any dependence on the field $a_\mathbf{r}$ in $L_{\mathbf{r}\mathbf{r}'}$ needs to be transformed into the interpolated form $a_\mathbf{r} \approx \psi_\mu(\mathbf{r}) a_\mu$. After this assumption, Eq. (28.7.10) becomes a closed equation for the discrete values a_μ of the hydrodynamic variables.

If we apply this discretization scheme on Eq. (28.6.8) we obtain

$$\begin{pmatrix} \partial_t \rho_\mu \\ \partial_t \mathbf{g}_\mu^\alpha \\ \partial_t \epsilon_\mu \end{pmatrix} = \begin{pmatrix} 0 & \left(\rho \delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right) & 0 \\ -\left(\rho \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & \left(\mathbf{g}^\alpha \delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right) - \left(\mathbf{g}^\beta \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & -\left(\epsilon \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) + \left(P \delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right) \\ 0 & \left(\epsilon \delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right) - \left(P \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial E}{\partial \rho_\nu} \\ \frac{\partial E}{\partial \mathbf{g}_\nu^\beta} \\ \frac{\partial E}{\partial \epsilon_\nu} \end{pmatrix} \tag{28.7.12}$$

where we have introduced the notation

$$\left(\rho \delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right) \equiv \int d\mathbf{r} \rho_\mathbf{r} \delta_\nu(\mathbf{r}) \boldsymbol{\nabla}^\alpha \delta_\mu(\mathbf{r}) \tag{28.7.13}$$

etc. The kinetic energy in the discrete CG Hamiltonian becomes, after substituting

the linearly interpolated fields into the kinetic energy functional, the following kinetic function

$$K(\rho, \mathbf{g}) = \frac{1}{2} \int d\mathbf{r} \mathbf{g}_\mu \psi_\mu(\mathbf{r}) \frac{1}{\psi_\sigma(\mathbf{r})\rho_\sigma} \mathbf{g}_\nu \psi_\nu(\mathbf{r}) = \frac{1}{2} \mathbf{g}_\mu \mathscr{A}_{\mu\nu}(\rho) \mathbf{g}_\nu \qquad (28.7.14)$$

where we have introduced the density dependent matrix

$$\mathscr{A}_{\mu\nu}(\rho) = \int d\mathbf{r} \psi_\mu(\mathbf{r}) \frac{1}{[\psi_\sigma(\mathbf{r})\rho_\sigma]} \psi_\nu(\mathbf{r}) \qquad (28.7.15)$$

Consider the matrix $\mathscr{M}_{\mu\nu}$ of the form

$$\mathscr{M}_{\mu\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma\right] \delta_\nu(\mathbf{r}) \qquad (28.7.16)$$

The matrices $\mathscr{A}$ and $\mathscr{M}$ are approximate inverses of each other, this is

$$\sum_\alpha \mathscr{A}_{\mu\alpha}(\rho) \mathscr{M}_{\alpha\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma\right] \sum_\alpha \delta_\alpha(\mathbf{r}) \int d\mathbf{r}' \psi_\alpha(\mathbf{r}') \frac{1}{[\psi_{\sigma'}(\mathbf{r}')\rho_{\sigma'}]} \psi_\nu(\mathbf{r}')$$

$$= \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma\right] D(\mathbf{r}, \mathbf{r}') \frac{1}{[\psi_{\sigma'}(\mathbf{r}')\rho_{\sigma'}]} \psi_\nu(\mathbf{r}') \quad (28.7.17)$$

If the interpolated density field changes slowly in the cell length scale, we may simply have

$$\sum_\alpha \mathscr{A}_{\mu\alpha}(\rho) \mathscr{M}_{\alpha\nu}(\rho) = \int d\mathbf{r} \int d\mathbf{r}' \delta_\mu(\mathbf{r}) D(\mathbf{r}, \mathbf{r}') \psi_\nu(\mathbf{r}') = \delta_{\mu\nu} \qquad (28.7.18)$$

where the last identity is exact when applied to exactly continued discrete fields. Therefore, we may model the kinetic energy term as simply

$$K(\rho, \mathbf{g}) = \frac{1}{2} \mathbf{g}_\mu \mathscr{M}_{\mu\nu}^{-1} \mathbf{g}_\nu$$

$$\mathscr{M}_{\mu\nu}(\rho) = \int d\mathbf{r} \delta_\mu(\mathbf{r}) \left[\psi_\sigma(\mathbf{r})\rho_\sigma\right] \delta_\nu(\mathbf{r}) \qquad (28.7.19)$$

The CG Hamiltonian is, therefore,

$$E(\rho, \mathbf{g}) = \frac{1}{2} \mathbf{g}_\mu \mathscr{M}_{\mu\nu}^{-1} \mathbf{g}_\nu + \sum_\mu \mathscr{V}_\mu \epsilon_\mu \qquad (28.7.20)$$

The derivatives of the discrete Hamiltonian are

$$\frac{\partial E}{\partial \rho_\nu} = \frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g})$$

$$\frac{\partial E}{\partial \mathbf{g}_\nu} = \mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'}$$

$$\frac{\partial E}{\partial \epsilon_\nu} = \mathscr{V}_\nu \tag{28.7.21}$$

where the density derivative of the kinetic energy is

$$\frac{\partial K}{\partial \rho_\nu}(\rho, \mathbf{g}) = \frac{1}{2}\mathbf{g}_{\mu'}^\alpha \frac{\partial \mathscr{M}_{\mu'\nu'}^{-1}}{\partial \rho_\mu}\mathbf{g}_{\nu'}^\alpha$$

$$= \frac{1}{2}\mathbf{g}_{\mu'}^\alpha \mathscr{M}_{\mu'\alpha}^{-1}\frac{\partial \mathscr{M}_{\alpha\beta}}{\partial \rho_\mu}\mathscr{M}_{\beta\nu'}^{-1}\mathbf{g}_{\nu'}^\alpha = \frac{1}{2}\mathbf{v}_{\mu'}^\alpha \frac{\partial \mathscr{M}_{\mu'\nu'}}{\partial \rho_\mu}\mathbf{v}_{\nu'}^\alpha$$

$$= \frac{1}{2}\mathbf{g}_{\mu'}^\alpha \mathscr{M}_{\mu'\alpha}^{-1}\left(\delta_\alpha \psi_\mu \delta_\beta\right)\mathscr{M}_{\beta\nu'}^{-1}\mathbf{g}_{\nu'}^\alpha \tag{28.7.22}$$

The dynamic equations after the model Hamiltonian (28.7.20) are

$$\begin{pmatrix} \partial_t \rho_\mu \\ \partial_t \mathbf{g}_\mu^\alpha \\ \partial_t \epsilon_\mu \end{pmatrix} = \begin{pmatrix} 0 & \left(\rho\delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right) & 0 \\ -\left(\rho\delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & \left(\mathbf{g}^\alpha \delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right) - \left(\mathbf{g}^\beta \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & -\left(\epsilon\delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) + \left(P\delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right) \\ 0 & \left(\epsilon\delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right) - \left(P\delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right) & 0 \end{pmatrix}\begin{pmatrix} \frac{\partial K}{\partial \rho_\nu} \\ \mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'} \\ \mathscr{V}_\nu \end{pmatrix}$$
$$\tag{28.7.23}$$

This gives the following equations

$$\partial_t \rho_\mu = \left(\rho\delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'}$$

$$\partial_t \mathbf{g}_\mu^\alpha = \left(\mathbf{g}^\alpha \delta_\nu \boldsymbol{\nabla}^\beta \delta_\mu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'} + \left(P\boldsymbol{\nabla}^\alpha \delta_\mu\right) - \left(\rho\delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right)\frac{\partial K}{\partial \rho_\nu} - \left(\mathbf{g}^\beta \delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'}$$

$$\partial_t \epsilon_\mu = -\left(\epsilon\delta_\mu \boldsymbol{\nabla}^\alpha \delta_\nu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'} + \left(P\delta_\nu \boldsymbol{\nabla}^\alpha \delta_\mu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'} \tag{28.7.24}$$

Note that mass is automatically conserved, energy is automatically conserved because of the antisymmetry of the L matrix. Momentum seems to not be automatically conserved. It will be conserved provided that the following identity is satisfied

$$-\left(\rho\boldsymbol{\nabla}^\alpha \delta_\nu\right)\frac{\partial K}{\partial \rho_\nu} - \left(\mathbf{g}^\beta \boldsymbol{\nabla}^\alpha \delta_\nu\right)\mathscr{M}_{\nu\nu'}^{-1}\mathbf{g}_{\nu'}^\beta = 0 \tag{28.7.25}$$

This corresponds to the identity

$$-\rho\nabla\left(\frac{\mathbf{v}^2}{2}\right) + \mathbf{g}\nabla\mathbf{v} = 0 \tag{28.7.26}$$