

The Micro-Meso connection
also known as
Non-Equilibrium Statistical Mechanics
also known as
The Theory of Coarse-Graining

Pep Español

Lecture notes at Courant, November 2013

Contents

1	The Microscopic Dynamics	9
1.1	Classical Mechanics	9
1.2	Hamilton's equations	9
1.3	Liouville's theorem	13
1.4	Equilibrium at the microscopic level	15
1.5	The single, all encompassing problem of Non-Equilibrium Statistical Mechanics	17
2	The Mesoscopic Dynamics	19
2.1	Levels of description	19
2.2	Stochastic processes in Phase Space	20
2.3	Green's view of coarse-graining	23
2.4	Zwanzig view of coarse-graining	29
2.4.1	Exact equation for $P(a, t)$	29
2.4.2	The Markovian approximation and the Fokker-Planck Equation	30
3	Example: Diffusing interacting colloidal particles	35

1

The Microscopic Dynamics

The scope of these notes has some restrictions. In particular, we will consider only *isolated* systems. We will assume also that at the most microscopic level the system can be well described by *classical mechanics* or, to be more specific, by *Hamilton's equations*. No reference to quantum mechanics will be made. Another crucial assumption is that the Hamiltonian dynamics of the system has a well-defined *equilibrium state* that is reached by the system as the time proceeds. The assumption of isolated system implies that we will look at the *relaxational dynamics* of the system towards its equilibrium state. This might seem a strong restriction from an experimental point of view. Experiments often deal with situations in which a system is subject to the action of “external influences”, usually through the boundary of the system. Nevertheless, the theory for isolated systems already provides the basic model equations to which boundary conditions can be applied in a latter stage.

1.1 Classical Mechanics

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 that interact through forces. We will often idealize and refer to atoms and even molecules as point particles, even though they may be composite objects. The functional forms of the forces between the point particles is known only in an approximate way, and model force fields are usually required and adopted. In principle, these forces fields should be derived from the quantum mechanical origin of the atoms or molecules that are represented at a classical level with point particles. The formulation of accurate force fields is a vast subject in itself of great current interest.

1.2 Hamilton's equations

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 knowlegde of the initial conditions given by the positions and velocities in order to predict the future evolution of the system. As stressed originaly 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. In Hamilton's formulation, the microscopic state $z = \{\mathbf{q}_i, \mathbf{p}_i\}$ of a system of N point particles is given by the collection of all \mathbf{q}_i and momenta \mathbf{p}_i of

the particles. The microscopic state of the system evolves according to Hamilton's equations

$$\dot{\mathbf{q}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i}(z), \quad (1.1)$$

$$\dot{\mathbf{p}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_i}(z), \quad (1.2)$$

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

$$\mathcal{H}(z) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U(\mathbf{r}_1, \dots, \mathbf{r}_N) + \sum_{i=1}^N V(\mathbf{r}_i) \quad (1.3)$$

where the first sum is the kinetic energy of the system, the potential of interaction between particles is $U(\mathbf{r}_1, \dots, \mathbf{r}_N)$ 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 $\mathcal{H}(z) \geq 0$ for all microstates z . The condition of bounded Hamiltonian is a requisite for the existence of a proper equilibrium state. One important case for which this condition *is not* satisfied is when the system interacts with a gravitational potential. In this case there are microstates for which the energy is arbitrarily negative (as when two massive point particles keep approaching and decreasing without bound its potential energy). A collection of self-gravitating point objects does not have a well-defined equilibrium state.

Hamilton's equations are first order differential equations that require the knowledge of an initial condition $z_0 = \{\mathbf{q}_i(0), \mathbf{p}_i(0)\}$. The ulterior evolution of the microstate z_t is a trajectory in the $6N$ dimensional space of all microscopic states z known as *the phase space* Γ of the system. In fact, the trajectory is restricted to live in a submanifold of the full phase space, because of the existence of **dynamic invariants**. A dynamical invariant $\mathcal{I}(z)$ is any dynamical function that does not change in time, this is,

$$\frac{d}{dt}\mathcal{I}(z(t)) = 0. \quad (1.4)$$

The trajectory $z(t)$ is, therefore, restricted to be in the submanifold $\mathcal{I}(z) = I_0$ where $I_0 = \mathcal{I}(z_0)$ is the value of the dynamical invariants at the initial time. The dynamical invariants emerge as the result of symmetries of the Hamiltonian. According to a fundamental theorem due to Emmy Noether, to every symmetry of the Hamiltonian there is conserved property. Invariance under time translation ensures that the Hamiltonian itself is a conserved quantity. Invariance under space translations ensure that total momentum is conserved, while rotation invariance ensure that angular momentum is conserved. A time independent Hamiltonian like (1.3) conserves energy. If there is no external potential and the interaction potential depends only on the relative distances between particles, then total linear and angular momentum will also be conserved. For typical Hamiltonian systems describing molecular systems, the trajectories in phase space 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. 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 in fact what makes the statistical methods applicable and what, in the last instance, makes Statistical Mechanics a predictive theory.

Hamilton's equations may be written in compact form

$$\dot{z}_t = J_0 \frac{\partial \mathcal{H}}{\partial z}(z_t) \quad (1.5)$$

J_0 is the so called symplectic matrix having a block diagonal matrix form with the blocks given by

$$\begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}. \quad (1.6)$$

We may write also Hamilton's equations in the form

$$\dot{z} = iLz \quad (1.7)$$

where iL is the Liouville operator that has the explicit form

$$iL \equiv \sum_i^N \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \frac{\partial}{\partial \mathbf{q}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{q}_i} \frac{\partial}{\partial \mathbf{p}_i} \right) = -\frac{\partial \mathcal{H}}{\partial z} J_0 \frac{\partial}{\partial z}. \quad (1.8)$$

The Liouville operator acting on an arbitrary function $F(z)$ can be expressed in terms of the Poisson bracket,

$$iLF(z) = -\{\mathcal{H}, F\} \quad (1.9)$$

where the Poisson bracket of two arbitrary functions $F(z), G(z)$ is

$$\{F, G\} \equiv \frac{\partial F}{\partial \mathbf{q}_i} \frac{\partial G}{\partial \mathbf{p}_i} - \frac{\partial F}{\partial \mathbf{p}_i} \frac{\partial G}{\partial \mathbf{q}_i} \quad (1.10)$$

Given an initial microstate z , the solution of Hamilton's equations can be denoted by

$$z_t = T_t z \quad (1.11)$$

where T_t is a one-parameter operator acting on the initial condition z . This operator satisfies $T_0 = \mathbf{1}$ and $T_t T_{t'} = T_{t+t'}$. A formal expression for this operator can be obtained in terms of the Liouville operator. Indeed, the formal solution of (1.7) is

$$z_t = \exp\{iLt\}z \quad (1.12)$$

where the exponential operator is defined formally through the Taylor series

$$\exp\{iLt\} \equiv 1 + iLt + \frac{1}{2!}(iLt)^2 + \frac{1}{3!}(iLt)^3 + \dots \quad (1.13)$$

Note that by substitution of (1.13) into (1.12) we recover the usual Taylor series of $z(t) = z_t$ around $z(0) = z$. Any function in phase space $X(z)$ acquires a time dependence once evaluated on the time dependent microstate, this is $X(T_t z)$. Phase functions evolve because the microstates evolve in time. The time derivative of this function is

$$\frac{d}{dt} X(T_t z) = \frac{\partial X}{\partial z}(T_t z) \frac{d}{dt} T_t z = \frac{\partial X}{\partial z}(T_t z) J_0 \frac{\partial \mathcal{H}}{\partial z}(T_t z) = iLX(T_t z) \quad (1.14)$$

This differential equation has as formal solution

$$X(T_t z) = \exp\{iLt\}X(z) \quad (1.15)$$

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 defined as the operator ϵ that takes a microstate $z = \{\mathbf{r}_i, \mathbf{p}_i\}$ and produces the microstate $\epsilon z = \{\mathbf{r}_i, -\mathbf{p}_i\}$, i.e. it reverses the sign of the momentum. Obviously ϵ^2 is the identity operator. With this operator, it is straightforward to observe the following property of the Liouville operator

$$iL\epsilon z = \sum_i \left[\frac{\partial \mathcal{H}(z)}{\partial \mathbf{p}_i} \frac{\partial}{\partial \mathbf{r}_i} - \frac{\partial \mathcal{H}(z)}{\partial \mathbf{r}_i} \frac{\partial}{\partial \mathbf{p}_i} \right] \begin{pmatrix} \vdots \\ \mathbf{r}_j \\ -\mathbf{p}_j \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \frac{\mathbf{p}_i}{m_i} \\ -\mathbf{F}_i \\ \vdots \end{pmatrix} = -\epsilon \begin{pmatrix} \vdots \\ \frac{\mathbf{p}_i}{m_i} \\ \mathbf{F}_i \\ \vdots \end{pmatrix} = -\epsilon iLz \quad (1.16)$$

This is, the time reversal operator and the Liouville operator anticommute when applied to microstates

$$iL\epsilon z = -\epsilon iLz \quad (1.17)$$

This is a reflection of the fact that the symplectic matrix J_0 and the time reversal matrix ϵ anticommute, this is

$$J_0\epsilon = \begin{pmatrix} \ddots & \vdots & \vdots & \dots \\ \dots & 0 & 1 & \dots \\ \dots & -1 & 0 & \dots \\ & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \ddots & \vdots & \vdots & \dots \\ \dots & 1 & 0 & \dots \\ \dots & 0 & -1 & \dots \\ & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \ddots & \vdots & \vdots & \dots \\ \dots & 0 & -1 & \dots \\ \dots & -1 & 0 & \dots \\ & \vdots & \vdots & \ddots \end{pmatrix} = -\epsilon J_0 \quad (1.18)$$

This property has the following consequence on the evolution operator applied to an initial condition z , $\exp\{iLt\}z$,

$$\begin{aligned} \epsilon \exp\{iLt\} &= \epsilon \left[1 + iLt + \frac{1}{2!}iLiLt^2 + \dots \right] z = \left[1\epsilon - iL\epsilon t - \frac{1}{2!}iL\epsilon iLt^2 + \dots \right] z \\ &= \left[1\epsilon - iL\epsilon t + \frac{1}{2!}iLiL\epsilon t^2 + \dots \right] z \\ &= \exp\{-iLt\}\epsilon z \end{aligned} \quad (1.19)$$

Changing the notation to $T_t = \exp\{iLt\}$ we have

$$\epsilon T_t z = T_{-t} \epsilon z \quad (1.20)$$

or, equivalently,

$$T_t \epsilon T_t z = \epsilon z \quad (1.21)$$

In words, this equation expresses the fact that if we take a microstate z and let it evolve according with Hamilton's equations to get $T_t z$, and then we reverse the velocities of the evolved microstate $\epsilon T_t z$, and then we evolve the resulting microstate for a time t , to get $T_t \epsilon T_t z$, we end up with the initial microstate with the velocities reversed ϵz .

1.3 Liouville's theorem

Hamilton's equation are first order differential equations that provide 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 identical 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 renders the evolution in phase space a stochastic process. The probability distribution function at a subsequent time is denoted by $\rho_t(z)$ and it obeys the Liouville equation which we will derive below for completeness. This equation is just an expression of a very important property of the solution of Hamilton's equation which is the conservation of volume in phase space.

We may think of (1.11) as a coordinate transformation from z to z_t . The Jacobian of this coordinate transformation is, by definition, the following determinant

$$\mathcal{J}_t(z_0) \equiv \det(\mathbf{J}_t(z_0)) \quad (1.22)$$

where the Jacobian matrix is

$$\mathbf{J}_t(z_0) \equiv \frac{\partial z_t(z_0)}{\partial z_0} \quad (1.23)$$

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

$$\mathcal{J}_t(z_0) = \det(\mathbf{J}_t(z_0)) = \exp\{\text{Tr}[\ln \mathbf{J}_t(z_0)]\} \quad (1.24)$$

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

$$\exp\{\text{Tr}[\ln \mathbf{J}_t(z_0)]\} = \exp\left\{\sum_k \ln \lambda_k\right\} = \prod_k \lambda_k \quad (1.25)$$

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

$$\begin{aligned} \frac{d}{dt} \mathcal{J}_t(z_0) &= \exp\{\text{Tr}[\ln \mathbf{J}_t(z_0)]\} \frac{d}{dt} \text{Tr}[\ln \mathbf{J}_t(z_0)] \\ &= \exp\{\text{Tr}[\ln \mathbf{J}_t(z_0)]\} \text{Tr} \left[\left(\frac{d}{dt} \mathbf{J}_t(z_0) \right) \mathbf{J}_t^{-1}(z_0) \right] \\ &= \mathcal{J}_t(z_0) \text{Tr} \left[\left(\frac{d}{dt} \mathbf{J}_t(z_0) \right) \mathbf{J}_t^{-1}(z_0) \right] \end{aligned} \quad (1.26)$$

Consider the time derivative of the Jacobian matrix

$$\begin{aligned}
\frac{d}{dt}\mathbf{J}_t(z_0) &= \frac{d}{dt}\frac{\partial z_t(z_0)}{\partial z_0} = \frac{\partial}{\partial z_0}\frac{d}{dt}z_t(z_0) \\
&= \frac{\partial}{\partial z_0}J_0\frac{\partial\mathcal{H}}{\partial z}(z_t(z_0)) \\
&= J_0\frac{\partial^2\mathcal{H}}{\partial z\partial z}(z_t(z_0))\frac{\partial z_t(z_0)}{\partial z_0} \\
&= J_0\frac{\partial^2\mathcal{H}}{\partial z\partial z}(z_t(z_0))\mathbf{J}_t(z_0)
\end{aligned} \tag{1.27}$$

Therefore, by inserting this result into (1.28) we obtain

$$\frac{d}{dt}\mathcal{J}_t(z_0) = \mathcal{J}_t(z_0)\text{Tr}\left[J_0\frac{\partial^2\mathcal{H}}{\partial z\partial z}(z_t(z_0))\right] = 0 \tag{1.28}$$

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 [?]. 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 T_tM 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 T_tM at $t = t$. For this reason,

$$\int_M \rho_0(z)dz = \int_{T_tM} \rho_t(z)dz. \tag{1.29}$$

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

$$\int_{T_tM} \rho_t(z)dz = \int_M \rho_t(T_tz)dz. \tag{1.30}$$

This is true for any region M and, therefore, the integrand of the right hand side of (1.29) and the left hand side of (1.30) should be equal, i.e.

$$\rho_0(z) = \rho_t(T_tz), \tag{1.31}$$

or, by a simple change of variables,

$$\rho_t(z) = \rho_0(T_{-t}z). \tag{1.32}$$

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 (1.32). This gives

$$\frac{d}{dt}\rho_t(T_tz) = \frac{d}{dt}\rho_0(z) = 0. \tag{1.33}$$

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 equation for the probability density in phase space,

$$\partial_t\rho(z, t) = -iL\rho(z, t), \tag{1.34}$$

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 (1.34) is given by (1.32).

An alternative derivation of the Liouville equation starts from the realization that the flow in phase space is incompressible. In order to prove this, note that from Hamilton's equations (1.2) we may define the flow velocity in phase space as

$$v(z) = J_0 \frac{\partial \mathcal{H}}{\partial z}(z) \quad (1.35)$$

Therefore, by taking the divergence of this velocity field we obtain

$$\frac{\partial}{\partial z} v(z) = J_0 \frac{\partial^2 \mathcal{H}}{\partial z \partial z}(z) = 0 \quad (1.36)$$

The zero comes from the fact that we are contracting a fully antisymmetric matrix J_0 with a fully symmetric one, the Hessian of the Hamiltonian. Eq. (1.36) 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. 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

$$\partial_t \rho_t(z) = - \frac{\partial}{\partial z} \cdot v(z) \rho_t(z) \quad (1.37)$$

By using the incompressibility condition (1.36), we obtain again the Liouville equation (1.34).

1.4 Equilibrium at the microscopic level

Let us consider now the final state predicted by the Liouville equation. A basic mathematical question is under which conditions the Liouville equation (1.34), 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 $\rho(z, t) = \delta(T_t z_0 - z)$, 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 broadening. To be more specific, if the dynamics of the system is of the *mixing* type, then the system reaches an *effective* stationary microscopic probability that it is a function of the microstate only through the *dynamical invariants* of the system [?]. Usually, the proof that a given system is of the mixing type is difficult but we will assume that our system is of the mixing type and has, therefore, a tendency to reach a well defined equilibrium state.

Any distribution function $\rho(z)$ which is a function $g(\mathcal{I}(z))$ will be, therefore, a stationary solution of the Liouville equation and, as stated, we will assume that any stationary solution is of this type. This stationary distribution is called the **equilibrium ensemble** $\rho^{\text{eq}}(z)$. Therefore,

$$\lim_{t \rightarrow \infty} \rho(z, t) = \rho^{\text{eq}}(z) = g(\mathcal{I}(z)). \quad (1.38)$$

Let us investigate the meaning of the function $g(I)$ by considering the probability distribution $P^{\text{eq}}(I)$ of dynamical invariants at equilibrium. By definition,

$$P^{\text{eq}}(I) = \int dz \rho^{\text{eq}}(z) \delta(\mathcal{I}(z) - I) = \int dz g(\mathcal{I}(z)) \delta(\mathcal{I}(z) - I) = g(I) \Omega^{\text{eq}}(I), \quad (1.39)$$

where we have introduced the measure $\Omega^{\text{eq}}(I)$ of the region of phase space compatible with a given set

of dynamical invariants

$$\Omega^{\text{eq}}(I) = \int dz \delta(\mathcal{I}(z) - I). \quad (1.40)$$

Equation (1.39) allows to identify $g(I)$ and (1.38) becomes

$$\rho^{\text{eq}}(z) = \frac{P^{\text{eq}}(\mathcal{I}(z))}{\Omega^{\text{eq}}(\mathcal{I}(z))}. \quad (1.41)$$

Therefore, the equilibrium ensemble is fully determined by the probability distribution of dynamical invariants at equilibrium.

It is obvious that the distribution of dynamical invariants at any time is itself invariant. The probability density of finding a value I of the dynamical invariants $\mathcal{I}(z)$ is given by

$$P(I, t) = \int dz \rho_t(z) \delta(\mathcal{I}(z) - I) \quad (1.42)$$

The time derivative of this probability is

$$\partial_t P(I, t) = \int dz (-iL) \rho_t(z) \delta(\mathcal{I}(z) - I) = \int dz \rho_t(z) iL \delta(\mathcal{I}(z) - I) = 0 \quad (1.43)$$

Therefore, the probability of dynamical invariants is itself an invariant. As a consequence the equilibrium distribution of dynamical invariants is just the same as the initial distribution of dynamical invariants $P^{\text{eq}}(I) = P(I, 0)$. We will denote $P_0(I) = P(I, t)$ and then the equilibrium ensemble is just

$$\rho^{\text{eq}}(z) = \rho_0 \frac{P_0(\mathcal{I}(z))}{\Omega_0(\mathcal{I}(z))}. \quad (1.44)$$

where we have introduced ρ_0 with dimensions of $(\text{action})^{-N}$ in order to have

$$\Omega_0(I) = \rho_0 \int dz \delta(\mathcal{I}(z) - I) \quad (1.45)$$

with the same physical dimensions as $P_0(I)$. The equilibrium ensemble $\rho^{\text{eq}}(z)$ is fully determined once the initial distribution $P_0(I)$ of dynamical invariants is known at the initial time. *Eq. (1.44) is a fundamental result of equilibrium Statistical Mechanics.* The intuitive meaning of (1.44) is very suggestive. $\Omega_0(I)$ is the measure of the submanifold of microstates corresponding to the invariants I and we may think that it is proportional to “the number” of microstates that have a value I for the dynamic invariants. Therefore, the probability density $\rho^{\text{eq}}(z)$ of a given microstate z is the “probability $P_0(\mathcal{I}(z))$ of being in the submanifold $\mathcal{I}(z) = I$ divided by the number of microstates of that submanifold”. We could say, then, that at equilibrium “all microstates with the same value of $\mathcal{I}(z)$ are equiprobable”. Of course, this is more of a mnemotechnical rule for (1.44) than a rigorous statement because, being a continuum submanifold, the “number of microstates” satisfying $\mathcal{I}(z) = I$ is infinite.

The reference value ρ_N^0

The function $\Omega(E)$ gives the overall measure of the submanifold $\mathcal{H}(z) = E$ in phase space. This submanifold is usually termed the **energy shell**. The function $\Omega(E)$ 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. It is also well-known that 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

$$\rho_N^0 = \frac{1}{h^{3N} N!} \quad (1.46)$$

Planck's constant h gives the appropriate dimensions whereas the factor $N!$ is due to the indistinguishability of the particles [?] and it is known as the correct Boltzmann counting. 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.

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

Non-Equilibrium Statistical Mechanics is based on the fundamental presupposition that all macroscopic processes can be ultimately understood in terms of the classical laws of motion of isolated systems. So far, we have seen how an isolated classical system made of particles 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 towards the equilibrium ensemble. This means that we restrict ourselves *to study 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.

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

$$\rho_0(z) \longrightarrow \rho^{\text{eq}}(z) \quad (1.47)$$

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 prepared¹ 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 which 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, 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

¹ Provided the distribution of dynamic invariants is the same in all preparations!

Statistical Mechanics, except that in the latter case not only the final equilibrium state is sought for, but the richer question of *how this state is reached in time* is answered.

2

The Mesoscopic Dynamics

2.1 Levels of description

The theory of coarse-graining is a formalization of the process of representing a given system with less information than that captured by the actual microstate of the system. One and the same system may be described at different **levels of description** depending on the amount of information which one retains macroscopically. The different levels of description of a system are characterized by the dynamical invariants of the system $\mathcal{I}(z)$ and a set of phase functions $A(z)$ *which are not dynamical invariants*. The phase functions $A(z)$ that characterize a given level of description will be referred to as **relevant variables** but they have received in the past a number of different names. Macrostates, macroscopic variables, gross variables, collective variables, coarse-grained variables, reaction coordinates, order parameters, internal variables, structural variables, etc. are all synonyms for relevant variables. With the symbol $A(z)$ we denote a collection of phase functions each one labeled with a discrete index like in, for example, $A(z) = \{A_\mu(z), \mu = 1, \dots, M\}$.

The identification of the relevant variables $A(z)$ is an art of the theory of coarse-graining and a crucial element 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 (1.47) 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 relevant 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 relevant 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. 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 relevant variables, we hope that the resulting description may be Markovian.

There are few guiding principles for the selection of relevant 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 Hamiltonian $\mathcal{H}(z)$ will be included in the description.* 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 the phase functions that best capture such feature in the collection of relevant variables. Ideally, one would like to develop tools for analyzing the flow in phase space and automatically produce the appropriate relevant variables for the problem at hand. This pattern recognition process is far from being addressed in the literature because the problems to face are enormous given the high dimensionality of phase space.

Let us turn back to the motion of the system in phase space. 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 relevant variables.

In the present chapter, we want to derive the governing equations for the probability distributions of the relevant variables starting from the microscopic dynamics of the system. The resulting exact dynamic equation for the probability distribution $P(a, t)$ takes, in the Markovian approximation, the form of a Fokker-Planck equation (FPE). This FPE was introduced by Green in a seminal paper in 1952 by using a line reasoning where he *assumed* that the stochastic process of relevant variables was a Markovian stochastic process. From this assumption he derived the governing equation for the transition probability of this process. Green’s paper is arguably a cornerstone in the theory of non-equilibrium statistical mechanics. The derivation by Zwanzig in 1961 of the same FPE with the help of a projection operator, showed how this equation emerges in the limit of clear separation of time scales from an exact non-Markovian equation.

2.2 Stochastic processes in Phase Space

We will consider relevant variables $A(z)$ with the property that an equation like $A(z) = a$ defines a proper submanifold of phase space. By this we mean that the measure $\Omega(a)$ of the submanifold $A(z) = a$ defined as

$$\Omega(a) = \int dz \rho_N^0 \delta(A(z) - a) \quad (2.1)$$

exists and it is well defined. Roughly, $\Omega(a)$ “counts” the number of microstates z that are compatible with a given macrostate a . When the relevant variables define proper submanifolds it is possible to describe the evolution of the relevant variables in terms of stochastic processes. A stochastic process is fully characterized [?] by giving the hierarchy of **joint probability distributions** $P(a_1, t_1, \dots, a_n, t_n)$ of having the value a_1 at time t_1 and the value a_2 at time t_2 etc. where $t_1 < t_2 < \dots < t_n$, for all n . Let us express this joint probability in microscopic terms. In this section we assume that the number of particles is known. Extension to the macrocanonical phase space is straightforward. The motion in phase space Γ of a microstate $z = (q_1, \dots, q_N, p_1, \dots, p_N)$, where q_i, p_i are the position and momentum of the i -th particle, can be viewed as a stochastic process itself. The corresponding joint probabilities for this process will be denoted by $\rho_n(z_1 t_1, \dots, z_n t_n)$. The combination of the two facts, that the randomness is given only at the initial time and that the later evolution is deterministic, makes the one-time probability density $\rho_1(z, t)$ the most relevant joint probability of the hierarchy. In fact, all the joint probabilities ρ_n with $n \geq 2$ can be expressed in terms of the one-time probability density $\rho_1(z, t)$. By denoting with $T_t z$

the solution of Hamilton's equations with initial conditions z we have

$$\rho_n(z_1 t_1, \dots, z_n t_n) = \rho_1(z_1 t_1) \delta(z_2 - T_{t_2 - t_1} z_1) \dots \delta(z_n - T_{t_n - t_{n-1}} z_{n-1}) \quad (2.2)$$

The one-time probability density or ensemble density satisfies the well known Liouville's equation with formal solution given by Eq. (1.32) All the stochasticity in the process given by $T_t z$ arises from the uncertainty in the initial conditions.

We next consider the evolution of the relevant variables $A(T_t z)$ as a consequence of the evolution of the microstate itself. The values that these dynamic variables take can be regarded as a stochastic process. The n -time joint probability densities $P(a_1 t_1, \dots, a_n t_n)$ which characterize the stochastic process of the relevant variables and the n -time probability densities $\rho(z_1 t_1, \dots, z_n t_n)$ of the corresponding microscopic process are related to each other through

$$P(a_1 t_1, \dots, a_n t_n) = \int \dots \int \rho(z_1 t_1, \dots, z_n t_n) \delta(A(z_1) - a_1) \dots \delta(A(z_n) - a_n) dz_1 \dots dz_n \quad (2.3)$$

which can be further simplified by using (2.2) and integrating over z_2, \dots, z_n

$$P(a_1 t_1, \dots, a_n t_n) = \int dz \rho(z, t_1) \delta(A(z) - a_1) \delta(A(T_{t_2 - t_1} z) - a_2) \dots \delta(A(T_{t_n - t_1} z) - a_n) \quad (2.4)$$

where it has been used that $T_t T_{t'} = T_{t+t'}$. By performing the change of coordinates $z = T_{-t_1} z_1$ (which has inverse $z_1 = T_{t_1} z$ and unit jacobian) equation (2.4) can be written as

$$P(a_1 t_1, \dots, a_n t_n) = \int dz \rho(z, 0) \delta(A(T_{t_1} z) - a_1) \delta(A(T_{t_2} z) - a_2) \dots \delta(A(T_{t_n} z) - a_n) \quad (2.5)$$

This is the final form for the macroscopic n -time joint probability density which appears as an integral over the initial ensemble $\rho(z, 0)$ of delta functions that "contract" the description from microstates z to macrostates a and which involves the microscopic dynamics T_t .

Initial ensemble

A basic question that arises now is, what is the actual functional form of $\rho(z, 0)$? As mentioned, in principle we cannot measure the initial microscopic state z exactly. If we are going to describe a system at a given coarse-grained level, we must assume that we have access to the measurement of the coarse-grained variables $A(z)$. In general, all the information we have about our system at the initial time is a particular distribution $P(a, 0)$, which is the outcome of a repeated set of measurements of the functions $A(z)$ with numerical outcomes a over the system prepared in an identical manner at the initial time. Therefore, we have to determine the distribution function $\rho(z, 0)$ with the sole information that it should provide precisely the distribution $P(a, 0)$. Both distribution functions are related through

$$P(a, 0) = \int dz \delta(A(z) - a) \rho(z, 0). \quad (2.6)$$

However, there are many different $\rho(z, 0)$ that can produce the same $P(a, 0)$. Which is the correct one? According to information theory [?], the *least biased* distribution which is compatible with the macroscopic information $P(a, 0)$ is the one that maximizes the entropy functional

$$S[\rho_0] \equiv -k_B \int_{\Gamma} \rho(z, 0) \ln \frac{\rho(z, 0)}{\rho_N^0} dz. \quad (2.7)$$

conditioned to the restriction (2.6). We encounter here a problem of Lagrange multipliers. By introducing the multipliers $\lambda(a)$ for the continuum set of restrictions (2.6) (one for each a), we maximize the functional $I[\rho_0] = S[\rho_0] + \int dx \lambda(a) P(a, 0) + \mu \int dx P(a, 0)$, where the μ Lagrange multiplier stands for the normalization to unity restriction of $\rho(z, 0)$. The Lagrange multipliers are obtained by substituting the maximum value into the restriction (2.6). The following final result is obtained [?]

$$\rho(z, 0) = \frac{P(A(z), 0)}{\Omega(A(z))}, \quad (2.8)$$

where $\Omega(a)$ is given in (2.1).

By substituting this initial ensemble into equation (2.5) and by choosing t_1 equal to 0 for simplicity, one obtains

$$P(a_1 0, \dots, a_n t_n) = \frac{P_1(a_1, 0)}{\Omega(a_1)} \int dz \delta(A(z) - a_1) \delta(A(T_{t_2} z) - a_2) \dots \delta(A(T_{t_n} z) - a_n) \quad (2.9)$$

We now introduce the conditional probability density $P(a_1 0 | a_2 t_2, \dots, a_n t_n)$ of finding the system in a_2 at time t_2 and in a_3 at time t_3 and so on till a_n, t_n , *provided* it was in a_1 at the initial time $t_1 = 0$. It is given by

$$\begin{aligned} P(a_1 0 | a_2 t_2, \dots, a_n t_n) &\equiv \frac{P(a_1 0, a_2 t_2, \dots, a_n t_n)}{P(a_1, 0)} = \\ &= \frac{1}{\Omega(a_1)} \int dz \delta(A(z) - a_1) \delta(A(T_{t_2} z) - a_2) \dots \delta(A(T_{t_n} z) - a_n) \end{aligned} \quad (2.10)$$

For further reference it is convenient to consider $n = 2$ in (2.10), that receives the name of **transition probability**

$$P(a_1 0 | a_2 t_2) = \frac{1}{\Omega(a_1)} \int dz \delta(A(z) - a_1) \delta(A(T_{t_2} z) - a_2) \quad (2.11)$$

This is a fundamental equation that relates the transition probability of the CG variables with the microscopic dynamics. It can be given an heuristic interpretation as follows. The numerator $\Omega(a_1)$ is “the number of microstates compatible with a_1 ” while the denominator in (2.11) is “the number of microstates compatible with a_1 that after a time are at a_2 ”. Therefore, the transition probability is just the fraction of microstates compatible with a_1 that after a given time are at a_2 .

2.3 Green's view of coarse-graining

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 relevant variables is a continuum Markov process*. As we will see in this section, this single hypothesis is sufficient to obtain a dynamical equations for the relevant variables, the Fokker-Planck equation, with all the objects appearing in the dynamics defined in terms of microscopic expressions. But first we have to introduce the Markov stochastic process and some of its properties.

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 fact, any other n -time probability is written as

$$P(a_1 t_1, \dots, a_n t_n) = P(a_1, t_1) P(a_1, t_1 | a_2, t_2) \cdots P(a_{n-1}, t_{n-1} | a_n, t_n) \quad (2.12)$$

For a Markov process the full stochastic process is characterized by the one time probability and the transition probability alone. We have already encountered a Markov process, the one corresponding to the deterministic Hamiltonian dynamics described in Eq. (2.2). 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

$$P(a_1, t_1 | a_2, t_2, a_3, t_3) = P(a_1, t_1 | a_2, t_2) P(a_2, t_2 | a_3, t_3) \quad (2.13)$$

We can interpret the Markov condition in geometrical terms as we did when discussing the mixing property. The Markov property says that the fraction of points of the submanifold a_1 that happen to be at a_2 at t_2 and then at a_3 at t_3 equals the fraction of points of a_1 that will be at a_2 at t_2 (irrespective of where they will go afterwards) times the fraction of points of a_2 that will be in a_3 at time t_3 .

If we integrate Eq. (2.13) over a_2 we need to have the following consistency condition

$$P_2(a_1, t_1 | a_3, t_3) = \int da_2 P(a_1, t_1 | a_2, t_2) P(a_2, t_2 | a_3, t_3) \quad (2.14)$$

This consistency condition is known as the **Chapman-Kolmogorov equation** for the transition probabilities. The intuitive idea 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 .

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 which is named as 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

$$\frac{\partial}{\partial t} P(a, t) = -\frac{\partial}{\partial a} D^{(1)}(a) P(a, t) + \frac{1}{2} \frac{\partial^2}{\partial a \partial a} D^{(2)}(a) P(a, t) \quad (2.15)$$

The two time joint probability and one time probability are related by

$$P(a_1, t_1) = \int da_0 P(a_0, t_0, a_1, t_1) \quad (2.16)$$

because both sides of this equation are the probability of finding a_1 at time t_1 irrespective of the value of a_0 at time t_0 . This equation gives the following integral equation relating the one time probability and

the transition probability

$$P(a_1, t_1) = \int da_0 P(a_0, t_0) P(a_0, t_0 | a_1, t_1) \quad (2.17)$$

Note that, from Eq. (2.17) if we take as initial condition $P(a_0, t_0) = \delta(a_0 - \hat{a})$ then $P(a, t) = P(\hat{a}, t_0 | a, t)$, 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 (2.15)

$$\frac{\partial}{\partial t} P(a_0 t_0 | a, t) = -\frac{\partial}{\partial a} D^{(1)}(a) P(a_0 t_0 | a, t) + \frac{1}{2} \frac{\partial^2}{\partial a \partial a} D^{(2)}(a) P(a_0 t_0 | a, t) \quad (2.18)$$

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

$$P(a_0, t_0 | a, t_0) = \delta(a - a_0) \quad (2.19)$$

Therefore, Green's basic assumption that the stochastic process of the relevant variables is a Markov process is equivalent to the hypothesis that the one-time and transition probabilities of the relevant variables obey the Fokker-Planck equation.

Molecular expresion of drift and diffusion

The drift vector $D^{(1)}(a)$ and the diffusion tensor $D^{(2)}(a)$ introduced in Eq. (2.15) are given in terms of moments of the transition probability at short times. Its particular form is, actually, specified as conditions in the derivation of the Fokker-Planck equation from the Chapman-Kolmogorov equation [?]. However, we may also obtain the specific form of these objects by just the requirement that the transition probability obeys the FPE (2.18). Let us see how this arise. The solution of the Fokker-Planck equation (2.18) is difficult to obtain in general due to the fact that the objects $D^{(1)}(a), D^{(2)}(a)$ may depend in general on the state a in a non-linear way. Nevertheless, it is possible to obtain an explicit solution *for short times*. Because the initial condition (2.19) of the transition probability is a Dirac delta, we expect that for sufficiently short times $t = t_0 + \Delta t$ with $\Delta t \approx 0$, the transition probability will remain highly peaked. In this case, we may approximate in (2.18) $D^{(1)}(a) \approx D^{(1)}(a_0)$ and $D^{(2)}(a) \approx D^{(2)}(a_0)$. This results in a Fokker-Planck equation with constant coefficients which is easy to solve

$$\frac{\partial}{\partial t} P(a_0 t_0 | a, t) = -D^{(1)}(a_0) \frac{\partial}{\partial a} P(a_0 t_0 | a, t) + D^{(2)}(a_0) \frac{1}{2} \frac{\partial^2}{\partial a \partial a} P(a_0 t_0 | a, t) \quad (2.20)$$

The exact solution of this equation with initial condition (2.19) has a Gaussian form

$$P(a_0, t_0 | a_1, t_0 + \Delta t) = \exp \left\{ -\frac{1}{2\Delta t} \left(a_1 - a_0 - \Delta t D^{(1)}(a_0) \right) D_{(2)}^{-1}(a_0) \left(a_1 - a_0 - \Delta t D^{(1)}(a_0) \right) \right\} \\ \times \frac{1}{(2\pi\Delta t)^{M/2} \det(D^{(2)}(a_0))^{1/2}} \quad (2.21)$$

provided that the inverse of the diffusion matrix $D^{(2)}(a)$ 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 delta functions [?]. The transition probability (2.21) has the following moments

$$\int da_1 (a_1 - a_0) P(a_0, t_0 | a_1, t_0 + \Delta t) = D^{(1)}(a_0) \Delta t \\ \int da_1 (a_1 - a_0)(a_1 - a_0) P(a_1, t_0 + \Delta t | a_0, t_0) = D^{(2)}(a_0) \Delta t + D^{(1)}(a_0) D^{(1)}(a_0) \Delta t^2 \quad (2.22)$$

The second moment can also be expressed as

$$\int da_1 (a_1 - a_0 - D^{(1)}(a_0)\Delta t)(a_1 - a_0 - D^{(1)}(a_0)\Delta t)P(a_1, t_0 + \Delta t | a_0, t_0) = D^{(2)}(a_0)\Delta t \quad (2.23)$$

Note that these expressions give the drift $D^{(1)}(a)$ and diffusion $D^{(2)}(a)$ in terms of the first and second moments of the short time form of the transition probability. But note now that we have also a microscopic expression for the transition probability, given in (2.11)! Therefore, by substituting (2.11) into (2.22) and (2.23) we may obtain explicit molecular expressions for $D^{(1)}(a)$ and $D^{(2)}(a)$. Consider the first term $D^{(1)}(a_0)$ and use Eq. (2.11) in Eq. (2.22)

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int da (a - a_0) \frac{1}{\Omega(a_1)} \int \delta(A(z) - a_0) \delta(A(T_{\Delta t}z) - a) dz \quad (2.24)$$

We perform the integral over the variable a of the Dirac delta function and obtain

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int dz \frac{\delta(A(z) - a_0)}{\Omega(a_0)} (A(T_{\Delta t}z) - A(z)) = \left\langle \frac{A(\Delta t) - A(0)}{\Delta t} \right\rangle^{a_0} \quad (2.25)$$

where we have introduced the generalized microcanonical average or conditional average

$$\langle \dots \rangle^{a_0} \equiv \int dz \frac{\delta(A(z) - a_0)}{\Omega(a_0)} \dots \quad (2.26)$$

Eq. (2.25) is the desired microscopic expression for the drift $D^{(1)}(a_0)$. In a similar way, by substitution of (2.11) into (2.23) we obtain a microscopic expression for the diffusion tensor

$$D^{(2)}(a_0) = \frac{1}{\Delta t} \left\langle [A(\Delta t) - a_0 - D^{(1)}(a_0)\Delta t][A(\Delta t) - a_0 - D^{(1)}(a_0)\Delta t] \right\rangle^{a_0} \quad (2.27)$$

This expression known as the **Einstein-Helfand formula** for the diffusion coefficient and it basically says that the square displacement of the relevant variables increases proportional to Δt , with the proportionality factor given by the diffusion tensor.

Now, while both (2.25) and (2.27) are microscopic expressions for the drift and diffusion tensor, they depend on the time Δt . Of course we would like to eliminate somehow the dependence on Δt . We want to expand the above expressions in terms of Δt and keep only the zero order term which is independent of Δt . However, the fact that Δt , although small, is not vanishing small leads to some subtleties. Let us consider the drift term first. As a first approximation, we would use a simple Taylor expansion of the relevant variable

$$A(T_{\Delta t}z) = A(z) + iLA(z)\Delta t + \mathcal{O}(\Delta t)^2 \quad (2.28)$$

and neglect terms of high order, to obtain

$$D^{(1)}(a_0) = \langle iLA \rangle^{a_0} + \mathcal{O}(\Delta t) \quad (2.29)$$

The drift would be given as the conditional average of the “velocity” of the relevant variables. However, this is not quite correct. Even though we have assumed that Δt is “small”, the time interval Δt cannot go to zero strictly because then the basic Markovian assumption with no memory of the past would not hold. We need to consider the time interval Δt as necessarily finite and the microscopic dynamics may

have time to do funny things. It is, therefore, preferable to use the following identity

$$A(T_{\Delta t}z) - A(z) = \int_0^{\Delta t} dt \frac{d}{dt} A(T_t z) = \int_0^{\Delta t} dt iLA(T_t z) \quad (2.30)$$

By using this identity in (2.25), we obtain an mathematically identical expression for $D^{(1)}(a_0)$, this is

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \int dz \frac{\delta(A(z) - a_0)}{\Omega(a_0)} iLA(T_t z) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \langle iLA(t) \rangle^{a_0} \quad (2.31)$$

In this way, we see that it the drift is the *time average* over the time interval Δt of the average “velocity” of the relevant variables. We now perform the change of variables $T_t z \rightarrow z$, with unit Jacobian, and obtain

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \int dz \frac{\delta(A(T_{-t}z) - a_0)}{\Omega(a_0)} iLA(z) \quad (2.32)$$

Now consider the identity, similar to (2.30)

$$A(T_{-t}z) = A(z) + \int_{-t}^0 dt' iLA(T_{t'}z) \quad (2.33)$$

$$D^{(1)}(a_0) = \frac{1}{\Delta t} \int_0^{\Delta t} dt \frac{1}{\Omega(a_0)} \int dz \delta \left(A(z) + \int_{-t}^0 dt' iLA(T_{t'}z) - a_0 \right) iLA(z) \quad (2.34)$$

We can now formally expand the Dirac delta function around $A(z) - a_0$

$$\delta \left(A(z) + \int_{-t}^0 dt' iLA(T_{t'}z) - a_0 \right) = \delta(A(z) - a_0) - \frac{\partial}{\partial a_0} \delta(A(z) - a_0) \int_{-t}^0 dt' iLA(T_{t'}z) + \dots \quad (2.35)$$

By inserting this formal expansion into (2.34) we have

$$\begin{aligned} D^{(1)}(a_0) &= \langle iLA \rangle^{a_0} \\ &+ \frac{1}{\Delta t} \int_0^{\Delta t} dt \frac{1}{\Omega(a_0)} \frac{\partial}{\partial a_0} \int_{-t}^0 dt' \int dz \delta(A(z) - a_0) iLA(T_{t'}z) iLA(z) \\ &+ \frac{1}{\Delta t} \int_0^{\Delta t} dt \frac{1}{\Omega(a_0)} \frac{\partial^2}{\partial a_0 \partial a_0} \int_{-t}^0 dt' \int_{-t}^0 dt'' \int dz \delta(A(z) - a_0) iLA(T_{t''}z) iLA(T_{t'}z) iLA(z) \\ &+ \dots \end{aligned} \quad (2.36)$$

This expression can be written in a more compact form as

$$\begin{aligned} D^{(1)}(a_0) &= \langle iLA \rangle^{a_0} \\ &+ \frac{1}{\Omega(a_0)} \frac{\partial}{\partial a_0} \Omega(a_0) \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_{-t}^0 dt' \langle iLA(t') iLA \rangle^{a_0} \\ &+ \frac{1}{\Omega(a_0)} \frac{\partial^2}{\partial a_0 \partial a_0} \Omega(a_0) \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_{-t}^0 dt' \int_{-t}^0 dt'' \langle iLA(t'') iLA(t') iLA \rangle^{a_0} \\ &+ \dots \end{aligned} \quad (2.37)$$

By retaining all the terms in this expression we have an expression that is mathematical equivalent to (2.25). We observe, therefore, that, in addition to the naive term $\langle iLA \rangle^{a_0}$ we have additional contributions that are time integrals of correlation functions. The task now is to evaluate whether these contributions scale as $(\Delta t)^0$ or with a higher power of Δt . Let us consider the first correlation matrix

$$\begin{aligned} \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_{-t}^0 dt' \langle iLA(t')iLA \rangle^{a_0} &= \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_0^t dt'' \langle iLA(-t'')iLA \rangle^{a_0} \\ &= \frac{1}{\Delta t} \int_0^{\Delta t} dt'' \int_{t''}^{\Delta t} dt \langle iLA(-t'')iLA \rangle^{a_0} \\ &= \frac{1}{\Delta t} \int_0^{\Delta t} dt'' (\Delta t - t'') \langle iLA(-t'')iLA \rangle^{a_0} \end{aligned} \quad (2.38)$$

where we have performed a change of variables. Next we add and subtract the term $\langle iLA \rangle^{a_0} \langle iLA \rangle^{a_0}$ in order to have a proper correlation that decays to zero

$$\begin{aligned} \frac{1}{\Delta t} \int_0^{\Delta t} dt \int_{-t}^0 dt' \langle iLA(t')iLA \rangle^{a_0} &= \int_0^{\Delta t} dt'' \langle (iLA(-t'') - \langle iLA \rangle^{a_0})(iLA - \langle iLA \rangle^{a_0}) \rangle^{a_0} \\ &\quad - \frac{1}{\Delta t} \int_0^{\Delta t} dt'' t'' \langle (iLA(-t'') - \langle iLA \rangle^{a_0})(iLA - \langle iLA \rangle^{a_0}) \rangle^{a_0} \\ &\quad + \frac{\Delta t}{2} \langle iLA \rangle^{a_0} \langle iLA \rangle^{a_0} \end{aligned} \quad (2.39)$$

We expect that the matrix of correlations $\langle (iLA(-t'') - \langle iLA \rangle^{a_0})(iLA - \langle iLA \rangle^{a_0}) \rangle^{a_0}$ decays to zero as $|t''|$ increases and that the integral then converges to a term which is independent of Δt , for sufficiently large Δt . We denote with $D(a_0)$ the resulting matrix

$$D(a_0) \equiv \int_0^{\Delta t} dt'' \langle (iLA(-t'') - \langle iLA \rangle^{a_0})(iLA - \langle iLA \rangle^{a_0}) \rangle^{a_0} \quad (2.40)$$

On the other hand, we also expect that the second integral in Eq. (2.39) also converges, but due to the prefactor $1/\Delta t$, this second integral will *decay* as $1/\Delta t^1$. Therefore, if Δt is slightly larger than the correlation time in which the velocity fluctuations of the relevant variables decay, we may neglect the second integral. Finally, the last term in Eq. (2.39) is of order Δt .

The analysis of the third and higher order terms in Eq. (2.37) becomes readily very complicated. We will assume, though, that these terms are “small” and that can be neglected. This approximation is not rigorous and one has to judge the resulting expression a posteriori. Therefore, we obtain the following form for the drift term

$$D^{(1)}(a_0) = \langle iLA \rangle^{a_0} + \frac{1}{\Omega(a_0)} \frac{\partial}{\partial a_0} \Omega(a_0) D(a_0) + \mathcal{O}(\Delta t) \quad (2.41)$$

where the explicitly displayed terms in this equation are independent of Δt . Note that, as compared with the naive result (2.29), we have an additional term in the drift. This term arises due to the fact that the time Δt is *large* compared with the correlation of the fluctuations of the velocity of the relevant variables. At the same time, Δt needs to be short in comparison with the evolution of the relevant variables themselves, otherwise, we are not allowed to use the Gaussian form for the transition probability on which the whole procedure is based. The existence of such time Δt is based on the fact that there must be a separation of time scales between the relevant variables and the fluctuations of their velocities.

¹One may consider a simple example in which the correlation decay as exponentials to get a feeling for this behaviour.

We may obtain an expression for the diffusion tensor (2.27) by expanding again in the small time Δt . In order to have the proper orders in Δt , substitute just one of the terms $(A(T_{\Delta t}z) - a_0)$ by using (2.28), this is

$$D^{(2)}(a_0) = \left\langle [iLA - D^{(1)}(a_0)][A(\Delta t) - A(0)] \right\rangle^{a_0} + \mathcal{O}(\Delta t) \quad (2.42)$$

This expression is equivalent to (2.27) up to small terms of order Δt and can also be used to obtain the diffusion tensor by microscopic means. This second form of the diffusion tensor has no specific name associated to it.

Still, a third way to express the diffusion tensor is by using the identity (2.30) in (2.42). Then we have

$$D^{(2)}(a_0) = \int_0^{\Delta t} dt \left\langle [iLA - D^{(1)}(a_0)]iLA(t) \right\rangle^{a_0} + \mathcal{O}(\Delta t) \quad (2.43)$$

We may symmetrize this expression by subtracting $D^{(1)}(a_0)$ to the last $iLA(T_t z)$ term

$$D^{(2)}(a_0) = \int_0^{\Delta t} dt \left\langle [iLA - D^{(1)}(a_0)][iLA(t) - D^{(1)}(a_0)] \right\rangle^{a_0} + \mathcal{O}(\Delta t) \quad (2.44)$$

This is the **Green-Kubo formula** for the diffusion tensor. The three expressions (2.27), (2.42) and (2.44) are equivalent forms for the diffusion tensor in microscopic terms.

We may now ask about the value of Δt . We have used the form (2.21) for the transition probability under the assumption that Δt is “small”. However, we cannot take the mathematical limit $\Delta t \rightarrow 0$ because the above microscopic expressions give the result $D^{(2)}(a) = 0$. Clearly Δt has to be a time which is large enough for the correlations of the fluctuations $iLA - D^{(1)}$ of the “velocity” to have decayed or, equivalently, for the square of the “displacement” $A(\Delta t) - a_0 - D^{(1)}(a_0)\Delta t$ in the Einstein-Helfand expression to have reached a behaviour linear in time. At the same time, it has to be short enough for being able to use the Gaussian approximation for the transition probability. As this Gaussian behaviour is related to the fact that we could make the substitution $D^{(1)}(a), D^{(2)}(a)$ with $D^{(1)}(a_0), D^{(2)}(a_0)$, we expect that the time Δt has to be short in front of the typical scale of evolution of the relevant variables. In summary, one obvious requirement for the validity of the above expressions is that there exists a *clear separation of time scales* between the relevant variables and the the fluctuations of the velocity of the relevant variables. In other words, the relevant variables need to be *slow variables* (in the time scale of its velocity fluctuations). It is obvious that one needs to include in the set of relevant variables a the dynamic invariants, in particular the Hamiltonian, of the system as they are the slowest possible variables. This is, the first component of the vector $A(z)$ needs to be the Hamiltonian. In this way the conditional expectations (2.26) do, in fact, contain a Dirac delta function over the Hamiltonian which is nothing else than the equilibrium microcanonical ensemble.

In this 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 relevant 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 relevant 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 an uncomfortable feeling.

In order to get some more insight into the problem, in the next sections, we will not postulate the Fokker-Planck equation but rather *will derive it from the microscopic Hamilton's equations*. We hope that in this way, we will have some more light into the problem.

2.4 Zwanzig view of coarse-graining

2.4.1 Exact equation for $P(a, t)$

The evolution of the microscopic ensemble $\rho(z, t)$ according to the Liouville equation induces an evolution of the mesoscopic distribution $P(a, t)$, because both are related according to

$$P(a, t) = \int \delta(A(z) - a) \rho(z, t) dz. \quad (2.45)$$

Of course, we would like to have a *closed* dynamical equation for $P(a, t)$ that makes no reference to the underlying dynamics given by $\rho(z, t)$. This closed equation can be obtained with the help of a projection operator technique. Following Zwanzig [?], we introduce a projection operator \mathcal{P} that applies to any function $F(z)$ of phase space Γ

$$\mathcal{P}F(z) = \langle F \rangle^{A(z)}, \quad (2.46)$$

where we have introduced the *conditional average* $\langle F \rangle^a$ by

$$\langle F \rangle^a = \frac{1}{\Omega(a)} \int dz \rho_N^0 \delta(A(z) - a) F(z), \quad (2.47)$$

Note that the effect of the operator \mathcal{P} on an arbitrary function of phase space is to transform it into a function of the relevant variables $A(z)$. The operator \mathcal{P} satisfies the projection property $\mathcal{P}^2 = \mathcal{P}$. We introduce also the complementary projection operator $\mathcal{Q} = 1 - \mathcal{P}$ which satisfies $\mathcal{P}\mathcal{Q} = 0$ and $\mathcal{Q}^2 = \mathcal{Q}$. The operators \mathcal{P}, \mathcal{Q} satisfy

$$\int dz \rho_N^0 A(z) \mathcal{P}B(z) = \int dz \rho_N^0 B(z) \mathcal{P}A(z), \quad (2.48)$$

for arbitrary functions $A(z), B(z)$. It is convenient to introduce the following notation

$$\Psi_a(z) = \delta(A(z) - a), \quad (2.49)$$

and consider the Dirac's delta function as an ordinary phase function with a continuum index a . According to the formal solution (1.15) this phase function will evolve according to

$$\Psi_a(T_t z) = \exp\{iLt\} \Psi_a(z), \quad (2.50)$$

and, therefore,

$$\partial_t \Psi_a(T_t z) = \exp\{iLt\} iL \Psi_a(z). \quad (2.51)$$

Now we introduce a *mathematical identity* between operators

$$\exp\{iLt\} = \exp\{iLt\} \mathcal{P} + \int_0^t dt' \exp\{iLt'\} \mathcal{P} iL \mathcal{Q} \exp\{iL\mathcal{Q}(t-t')\} + \mathcal{Q} \exp\{iL\mathcal{Q}t\}. \quad (2.52)$$

This identity can be proved by taking the time derivative on both sides. If two operators have the same derivative and coincide at $t = 0$ then they are the same operator. We now apply this identity (2.52) to the left hand side of (2.51). After some algebra, which uses the explicit form of the operators \mathcal{P}, \mathcal{Q} , the properties (2.48), and the chain rule in the form

$$iL \Psi_a(z) = iL A_\mu \frac{\partial}{\partial a_\mu} \Psi_a(z), \quad (2.53)$$

where summation over repeated indices is implied, one obtains,

$$\begin{aligned}\partial_t \Psi_a(T_t z) &= -\frac{\partial}{\partial a_\mu} \cdot v_\mu(a) \Psi_a(T_t z) \\ &+ \int_0^t dt' \int da' \Omega(a') \frac{\partial}{\partial a_\mu} \cdot D_{\mu\nu}(a, a', t-t') \cdot \frac{\partial}{\partial a'_j} \frac{\Psi_{a'}(T_{t'} z)}{\Omega(a')} \\ &+ \mathcal{Q} \exp\{iL\mathcal{Q}t\} \mathcal{Q} iL \Psi_a(T_t z).\end{aligned}\quad (2.54)$$

We have defined the drift $v_\mu(a)$ and the diffusion tensor $D_{\mu\nu}(a, a', t)$ through

$$\begin{aligned}v_\mu(a) &= \langle iLA_\mu \rangle^a, \\ D_{\mu\nu}(a, a', t) &= \langle (iLA_\nu - \langle iLA_\nu \rangle^{a'}) \exp\{iL\mathcal{Q}t\} \Psi_a(iLA_\mu - \langle iLA_\mu \rangle^a) \rangle^{a'}.\end{aligned}\quad (2.55)$$

If we multiply (2.54) by $\rho(z, 0)$, integrate over z , and use (??), we obtain a final exact and closed equation for $P(a, t)$

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} \cdot v_\mu(a) P(a, t) + \int_0^t dt' \int da' \Omega(a') \frac{\partial}{\partial a_\mu} \cdot D_{\mu\nu}(a, a', t-t') \cdot \frac{\partial}{\partial a'_\nu} \frac{P(a', t')}{\Omega(a')}, \quad (2.56)$$

where we have used that the initial ensemble (2.8) is a function of $A(z)$ and, therefore,

$$\int dz \rho(z, 0) \mathcal{Q} \exp\{iL\mathcal{Q}t\} \mathcal{Q} iL \Psi_a(T_t z) = 0, \quad (2.57)$$

where we have used the hermiticity (2.48) as well as the projection property $\mathcal{Q}f(A) = 0$.

2.4.2 The Markovian approximation and the Fokker-Planck Equation

Equation (2.56) is an *exact an rigorous* closed equation governing the distribution function $P(a, t)$. No approximations have been made and, essentially, it is another way of rewriting the Liouville equation. In principle, it is as difficult to solve as the original Liouville equation. However, as it happens often in Physics, just by rewriting the same thing in a different form, it is possible to perform suitable approximations that allow for an advance in the understanding of the problem. In the case of (2.56), the approximation is called the *Markovian* approximation and transforms the integro-differential equation into a simple Fokker-Planck equation.

The Markovian assumption is one about *separation of time scales* between the time scale of evolution of the phase function $A(z)$ and the rest of variables of the system. If this separation of time scales exists then, in the time scale in which the tensor $D_{\mu\nu}(a, a', t-t')$ decays, the probability $P(a, t')$ has not changed appreciably. Schematically, we write the memory term in (2.56) as

$$\int_0^t dt' D(t-t') P(t') \approx P(t) \int_0^{\Delta t} D(t') dt'. \quad (2.58)$$

This approximation is depicted in Fig. 2.1. We have extended in (2.58) the upper limit of integration to a time Δt sufficiently large for the memory kernel $D(t)$ to have decayed. Note that the tensor $D_{\mu\nu}(a, a', t-t')$ is a quantity of order $(iLA)^2$, i.e. second order in the time derivatives of the relevant variables. The time scale of evolution of $P(a, t)$ is the same as the time scale of the variables $A(z)$. The approximation (2.58) amounts, therefore, to neglect third order time derivatives of the relevant variables in front of second order terms. We, therefore, consistently perform a formal expansion of the tensor

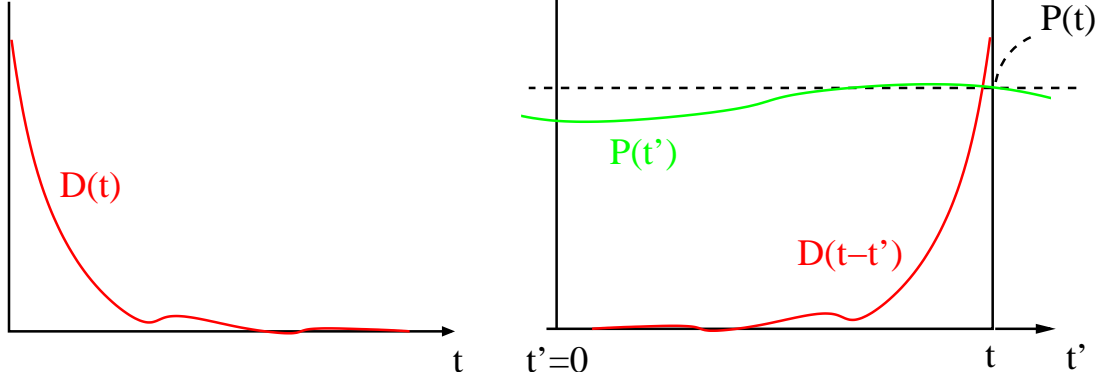


Figure 2.1: The Markovian approximation.

$D_{\mu\nu}(a, a', t - t')$ in (2.55) in terms of iLA and keep only second order terms. Then,

$$\exp\{iLQt\}\Psi_a Q iLA = \Psi_a \exp\{iLQt\} Q iLA + \mathcal{O}(iLA)^2. \quad (2.59)$$

Therefore, up to terms of order $\mathcal{O}(iLA^3)$ we have

$$D_{\mu\nu}(a, a', t) = \delta(a - a') \langle (iLA_\nu - \langle iLA_\nu \rangle^a) \exp\{iLQt\} (iLA_\mu - \langle iLA_\mu \rangle^a) \rangle^a, \quad (2.60)$$

and the tensor becomes diagonal in a, a' . By substitution of the approximate form (2.60) into the exact equation (2.56) and using (2.58) we obtain

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} v_\mu(a) P(a, t) + \frac{\partial}{\partial a_\mu} \Omega(a) D_{\mu\nu}(a) \frac{\partial}{\partial a_\nu} \frac{P(a, t)}{\Omega(a)}, \quad (2.61)$$

where we have defined the *diffusion tensor*

$$D_{\mu\nu}(a) = \int_0^{\Delta t} dt' \langle (iLA_\nu - \langle iLA_\nu \rangle^a) \exp\{iLQt'\} (iLA_\mu - \langle iLA_\mu \rangle^a) \rangle^a, \quad (2.62)$$

Note that within the neglect of third order terms, we can also substitute the projected dynamics with the real dynamics, this is

$$D_{\mu\nu}(a) = \int_0^{\Delta t} d\tau \langle (iLA_\nu - V_\nu(a)) \exp\{-iL\tau\} (iLA_\mu - V_\mu(a)) \rangle^{Ea} \quad (2.63)$$

We summarize for completeness the rest of quantities appearing in (2.61)

$$\begin{aligned} v_\mu(a) &= \langle iLA_\mu \rangle^a, \\ \langle \dots \rangle^a &= \frac{1}{\Omega(a)} \int dz \rho_N^0 \delta(A(z) - a) \dots, \\ \Omega(a) &= \int dz \rho_N^0 \delta(A(z) - a). \end{aligned} \quad (2.64)$$

As we include in the set $A(z)$ the total energy as a relevant variable, it is convenient to single out its

effect. Let us write explicitly

$$\begin{aligned} v_\mu(a, E) &= \langle iLA_\mu \rangle^{a, E}, \\ \langle \dots \rangle^a &= \frac{1}{\Omega(a, E)} \int dz \rho_N^0 \delta(H(z) - E) \delta(A(z) - a) \dots, \\ \Omega(a, E) &= \int dz \rho_N^0 \delta(H(z) - E) \delta(A(z) - a). \end{aligned} \quad (2.65)$$

It is convenient to multiply the numerator and denominator of the conditional averages with

$$\Omega(E) \equiv \int dz \delta(H(z) - E) \quad (2.66)$$

which is the measure of the number of microstates of a given energy. In this way, we obtain the conditional averages as

$$\langle \cdot \rangle^{a, E} = \frac{1}{P^{\text{mic}}(a)} \int dz \rho^{\text{mic}}(z) \delta(A(z) - a) \dots \quad (2.67)$$

where we have introduced the usual microcanonical ensemble $\rho^{\text{mic}}(z)$ and the equilibrium probability of the relevant variables as

$$\begin{aligned} \rho^{\text{mic}}(z) &= \frac{1}{\Omega(E)} \delta(H(z) - E) \\ P^{\text{eq}}(a) &= \int dz \rho^{\text{mic}}(z) \delta(A(z) - a) \end{aligned} \quad (2.68)$$

at the same time we have

$$\Omega(a, E) = \Omega(E) P^{\text{eq}}(a) \quad (2.69)$$

Because $iLH = 0$, many of the terms in the FPE vanish, precisely those involving derivatives with respect to the total energy. For this reason, we may write (2.61) as

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} V_\mu(a) P(a, t) + \frac{\partial}{\partial a_\mu} P^{\text{eq}}(a) D_{\mu\nu}(a) \frac{\partial}{\partial a_\nu} \frac{P(a, t)}{P^{\text{eq}}(a)}, \quad (2.70)$$

which may also be written in the equivalent form

$$\partial_t P(a, t) = -\frac{\partial}{\partial a_\mu} \left[V_\mu(a) + D_{\mu\nu}(a) \frac{\partial S}{\partial a}(a) \right] P(a, t) + \frac{\partial}{\partial a_\mu} D_{\mu\nu}(a) \frac{\partial}{\partial a_\nu} P(a, t) \quad (2.71)$$

where we have introduced the entropy function $S(a)$ through the celebrated Einstein's formula for fluctuations

$$P^{\text{eq}}(a) = \exp\{S(a)/k_B\} \quad (2.72)$$

We can check that $P^{\text{eq}}(a)$ is the equilibrium solution. By substituting $P^{\text{eq}}(a)$ into (2.61) we have that $P^{\text{eq}}(a)$ will be a stationary solution of the FPE (2.61) if and only if

$$\sum_\mu \frac{\partial}{\partial a_\mu} v_\mu(a) P^{\text{eq}}(a) = 0. \quad (2.73)$$

By using the definition (2.55), the constrained average (2.47), and the chain rule we have

$$\frac{\partial}{\partial a_\mu} v_\mu(a) P^{\text{eq}}(a) = \int dz \rho^{\text{mic}}(z) \frac{\partial}{\partial a_\mu} \delta(A(z) - a) iL A_\mu(z) = - \int dz \rho^{\text{mic}}(z) iL \delta(A(z) - a) = 0, \quad (2.74)$$

where we have integrated by parts the Liouville operator and used that $iLH = 0$.

Equation (2.61) is the desired Fokker-Planck equation for the dynamics at the mesoscopic coarse-grained level of description. This equation is one of the cornerstones of non-equilibrium statistical mechanics and was obtained by Zwanzig in 1961 [?] following the path pioneered by Green in [?]. In this equation, all the objects $v_\mu(a)$, $D_{\mu\nu}(a, t)$ and $\Omega(a)$ have a definite microscopic definition. In particular, (2.62) is a general form of the well-known *Green-Kubo formulae* that relates the *transport coefficients* $D_{\mu\nu}(a)$ with a time integral of a correlation function of microscopic variables. The FPE (2.61) is valid whenever there is a clear separation of time scales such that the Markovian approximation is valid. Clearly, the FPE will describe correctly the evolution of $P(a, t)$ only for times which are larger than the typical time scales of decay of the correlation involved in $D_{\mu\nu}(a)$. We cannot investigate with this equation shorter time scales. For this short time scales, the transport coefficients start to depend on time.

3

Example: Diffusing interacting colloidal particles

At the most microscopic level, we can model the colloidal suspension by assuming that the solid suspended objects are spherical and we need only 6 degrees of freedom for describing the state of the object, the position \mathbf{Q}_i and the momentum \mathbf{P}_i of its center of mass. For irregular objects we would need also to consider orientation, angular velocities, etc. The fluid in which these solid colloidal particles are suspended will be described at the most microscopic level by the positions \mathbf{q}_i and momenta \mathbf{p}_i of the center of mass of the molecules that constitute the fluid. Again, we assume spherical molecules for simplicity. The microscopic state will be denoted by $z = \{\mathbf{q}_i, \mathbf{p}_i, \mathbf{Q}_i, \mathbf{P}_i\}$. The evolution of the microstate is governed by Hamilton's equations,

$$\begin{aligned}\dot{\mathbf{q}}_i &= \frac{\partial H(z)}{\partial \mathbf{p}_i}, & \dot{\mathbf{Q}}_i &= \frac{\partial H(z)}{\partial \mathbf{P}_i}, \\ \dot{\mathbf{p}}_i &= -\frac{\partial H(z)}{\partial \mathbf{q}_i}, & \dot{\mathbf{P}}_i &= -\frac{\partial H(z)}{\partial \mathbf{Q}_i},\end{aligned}\tag{3.1}$$

where the Hamiltonian is given by

$$H(z) = \sum_i \left(\frac{p_i^2}{2m_i} + \frac{P_i^2}{2M_i} \right) + \frac{1}{2} \sum_{ij} (V_{ij}^{SS}(q) + V_{ij}^{SC}(q, Q) + V_{ij}^{CC}(Q)).\tag{3.2}$$

Here, m_i is the mass of a solvent molecule, M_i the mass of a colloidal particle, and V^{SS} , V^{SC} , V^{CC} are the potential of the forces between solvent molecules, solvent and colloidal particles, and colloidal particles, respectively.

In principle, the differential equations (3.1) can be solved numerically with a computer. The technique is known as *molecular dynamics* and allows us to keep track of all the microscopic dynamics of the system [?]. The smallest typical time scale is a collision time in the range of picoseconds and, consistently, we will need to use a time step for the numerical solution which is much smaller than this 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.

We illustrate now how the general formalism developed in the previous section can be applied to the case of a colloidal suspension in order to derive the FPE. The idea is simply to translate to our system the different objects defined in (2.62), (2.65) that appear in the FPE (2.61). The microscopic Hamiltonian is given in (3.2). We select as relevant variables $A(z) = x$ the positions of the colloidal particles \mathbf{Q}_i which take numerical values $\bar{\mathbf{Q}}_i$. Let us consider the equilibrium probability for these variables. It is given by

$$P^{\text{mic}}(\bar{Q}) = \int dz \rho^{\text{mic}}(z) \prod_i \delta(\mathbf{Q}_i - \bar{\mathbf{Q}}_i) \quad (3.3)$$

As we know from equilibrium statistical mechanics, we may use the canonical ensemble instead of the microcanonical ensemble in order to compute averages. The canonical ensemble is given

$$\rho^{\text{eq}}(z) = \frac{1}{Z} \exp\{-\beta H(z)\}, \quad (3.4)$$

where $\beta = (k_B T)^{-1}$ is proportional to the inverse of the temperature T . The particular value of T is fixed in order to give an average energy given by E . By using the canonical ensemble, we may write the equilibrium distribution of the positions of the colloidal particles as

$$\begin{aligned} P^{\text{mic}}(\bar{Q}) &\propto \exp\{-\beta V^{CC}(\bar{Q})\} \int dq \exp\{-\beta (V^{CS}(\bar{Q}, q) + V^{SS}(q))\} \\ &\equiv \exp\{-\beta \bar{V}^{\text{eff}}(\bar{Q})\}, \end{aligned} \quad (3.5)$$

where $\int dq$ is a condensed notation for the integral over solvent positions. We have introduced the effective potential as

$$\bar{V}^{\text{eff}}(\bar{Q}) = V^{CC}(\bar{Q}) - k_B T \ln \int dq \exp\{-\beta (V^{CS}(\bar{Q}, q) + V^{SS}(q))\}. \quad (3.6)$$

The effective potential has a contribution $V^{CC}(\bar{Q})$ coming from the direct interaction potential and an additional contribution that represents the effect of the static and equilibrium averaged solvent mediated interaction between colloidal particles.

After performing the integrals over the Dirac delta functions, the constrained average in (2.65) now takes the form

$$\langle \dots \rangle_{\bar{Q}} = \frac{1}{\Omega(\bar{Q})} \int dq \rho^{\text{eq}}(q, \bar{Q}) \dots \quad (3.7)$$

Note that this constrained average is simply an equilibrium average over the solvent degrees of freedom, in which the colloidal particles are assumed to be fixed at the values \bar{Q} . It is, therefore, an equilibrium ensemble average in which the colloidal particles act as external static force fields.

Because the time derivatives iLA are simply P_i/M_i , the drift term $v(a) = \langle iLA \rangle^a$ defined in (2.65) is now the constrained equilibrium average of the momentum of the colloidal particles, which is zero by isotropy of the equilibrium ensemble. The diffusion tensor (2.62) becomes

$$\mathbf{D}_{ij}(\bar{Q}) = \int_0^{\Delta t} dt' \langle \mathbf{V}_j \exp\{iLt'\} \mathbf{V}_i \rangle_{\bar{Q}}. \quad (3.8)$$

The final FPE (2.61) takes now the form

$$\partial_t P(Q, t) = \frac{\partial}{\partial \mathbf{Q}_i} \left[\beta \mathbf{D}_{ij}(Q) \frac{\partial \bar{V}^{\text{eff}}(Q)}{\partial \mathbf{Q}_j} \right] P(Q, t) + \frac{\partial}{\partial \mathbf{Q}_i} \mathbf{D}_{ij}(Q) \cdot \frac{\partial}{\partial \mathbf{Q}_j} P(Q, t), \quad (3.9)$$

In this example, we observe how the general Fokker-Planck description can be applied to a specific level

of description of a given system. The essential benefits of this approach are that it is very simple to get the *structure* of the coarse-grained equation. Also, we obtain explicit microscopic expressions for all the objects in the FPE. In particular, the diffusion tensor which describes the mutual, solvent-mediated influence of the colloidal particles is given in terms of the auto and cross-correlations of the velocities of the colloidal particles, where the averages are taken over the solvent degrees of freedom which are distributed according to an equilibrium ensemble in the presence of the external fields due to the static colloidal particles. The FPE that governs now the probability density $P(\mathbf{Q}, t)$ is called the Smoluchowsky equation.