

CHAPTER II

Distributions of statistical mechanics

The purpose of Statistical Mechanics is to explain the thermodynamic properties of macroscopic systems starting from underlying microscopic models—possibly based on “first principle” theories. In this chapter, we shall first argue that in practice such an approach unavoidably involves probabilities to become predictive (Sec. II.1). The specific modalities of the implementation of probabilities are then presented for both classical (Sec. II.2) and quantum mechanical (II.3) systems. We shall on purpose introduce descriptions that are very parallel to each other, which will in following chapters allow us to treat some problems only in the quantum mechanical framework and argue that the same reasoning could have been adopted in a classical setup, and reciprocally. Eventually, Sec. II.4 deals with the missing information arising from the probabilistic nature of the description, and its quantitative measure.

II.1 From the microscopic scale to the macroscopic world

II.1.1 Orders of magnitude and characteristic scales

The notion of a *macroscopic system* and the related variables are usually understood to apply to a system whose characteristic scales are close to (or much larger than) the scales relevant for a human observer, namely typical length scales of 1 mm – 1 m or larger, durations of 1 s, kinetic energies of 1 J, and so on.

In opposition, *microscopic scales*—at which the laws of the dynamics of point particles apply⁽²²⁾—are rather understood to refer to atomic or molecular scales. For instance, the typical distance between atoms or molecules—collectively called “particles” in the following for brevity⁽²²⁾—in a solid or a gas is about 10^{-10} to 10^{-8} m respectively, so that 1 cm³ of solid resp. gaseous phase consists of $N \approx 10^{24}$ resp. 10^{19} particles. The typical microscopic energy scale is the electron-volt, where 1 eV is 1.6×10^{-19} J, while the typical durations range from ca. 10^{-15} s [= $\hbar/(1 \text{ eV})$] to 10^{-9} s (typical time interval between two collisions of a particle in a gas under normal conditions).

Remarks:

* In the case of the usual application of the concepts of Statistical Physics discussed above, the interaction between microscopic degrees of freedom is mostly of electromagnetic nature. The distances between particles are too large for the strong or weak interactions to play a role, while the masses remain small enough, to ensure that gravitational effects are negligible with respect to electromagnetic ones—unless of course when investigating astrophysical objects.

* The methods of Statistical Physics are also sometimes applied in circumstances in which the particle number is much smaller than above. For instance, the 10^3 – 10^4 particles emitted in high-

⁽²²⁾This generic denomination “(point) particles” does not necessarily involve a description in terms of particles. It is rather a convenient shorthand expression for “elementary degrees of freedom”, which may possibly actually be described by some quantum field theory.

energy collisions of heavy nuclei are often considered as forming a statistical system—even worse, a system possibly in local thermodynamic equilibrium.

* At the other end of the length spectrum, some astrophysical and cosmological simulations treat stars or even galaxies as pointlike objects constituting the “microscopic” scale in the respective descriptions of galaxies or galaxy clusters.

II.1.2 Necessity of a probabilistic description

A typical macroscopic system consists in general of a large number of degrees of freedom, which obey microscopic laws. Theoretically, one could think of simulating the evolution of such a system—say for instance of 10^{23} particles—on a computer. The mere storage of the positions and momenta at a given instant t_0 of so many particles, coding them as real numbers in simple precision (4 Bytes), already requires about 10^{12} hard disks with a capacity of 1 TB! This represents only the first step, since one should still solve the equations of motion for the particles... As of late 2014, the most extensive simulations of molecular dynamics study the motion of $N \lesssim 10^7$ particles over a time duration $\lesssim 1 \mu\text{s}$, i.e. for about 10^6 – 10^8 time steps ⁽²³⁾⁽²⁴⁾

In addition, the dynamical equations of such a many-body system are often characterized by their sensitivity to the initial conditions, which grows with increasing particle number. Thus, two trajectories in the phase space of a classical system that correspond to initial conditions which only differ by an infinitesimal distance ϵ might after a duration Δt be about $\epsilon e^{\lambda \Delta t}$ away from each other, with some *Lyapunov* ^(t) exponent $\lambda > 0$; that is, an originally infinitesimal error grows exponentially. The system is then *chaotic* and any prediction regarding individual particles quickly become hazardous.

As a consequence, one has to abandon the idea of a purely deterministic microscopic description even in the case of a classical system, and adopt a new approach. Instead of attempting to describe the exact microscopic state, or more briefly *microstate*, of a system at a given instant, one must rather investigate the probability to find the system in a given microstate. In that approach, one must characterize the system through (a manageable amount of) macroscopic quantities, that altogether define a *macrostate*.

Such a macroscopic observable usually results from summing microscopic quantities over many particles, and it is often defined as the expectation value of the sum or of the arithmetic mean. Thanks to the central limit theorem (Appendix B.5) the fluctuations of the sum about this expectation value are of relative magnitude $1/\sqrt{N}$, i.e. very small when $N \gtrsim 10^{20}$, so that the observable is known with high accuracy.

In practice, the empirical laws relating such macroscopic observables have been obtained by repeating several measurements, the results of which have been averaged to get rid of experimental uncertainties. For instance, establishing the local form of Ohm’s law (I.43b) relies on performing many measurements of the electrostatic field in the conductor and of the resulting electric current density, and the obtained law should rather read $\langle \vec{J}_{\text{el.}} \rangle = \sigma_{\text{el.}} \langle \vec{\mathcal{E}} \rangle$, which describes the actual procedure better than the traditional form (I.43b), and again emphasizes the need for a statistical approach.

⁽²³⁾ An explicit example is the simulation of the motions of 6.5 million atoms for about 200 ns within a multi-time-step technique involving steps of 0.5, 2 and 4 fs for various “elementary processes” ^[23].

⁽²⁴⁾ The current tendency is not towards increasing these numbers by brute force, but rather to use “multiscale approaches” in which part of the large scale phenomena are no longer described microscopically, but using macroscopic variables and evolution equations.

^(t) A. LYAPUNOV, 1857–1918

II.2 Probabilistic description of classical many-body systems

In the previous section, it has been argued that the description of a macroscopic physical system should be based on a statistical approach. We now discuss the practical introduction of the corresponding probabilities in the microscopic formalism of classical mechanics.

After recalling the basics of the phase-space based formalism for exactly known classical systems (§ II.2.1), we introduce the notion of a probability density on phase space to account for incomplete knowledge of the exact microscopic state of a system with a given finite number of constituents (§ II.2.2). We next derive the evolution equations governing the dynamics of that density (§ II.2.3) and of observables (§ II.2.4). Eventually, we shortly mention the generalization to systems whose particle number is not fixed (§ II.2.5).

The purpose in the following is not only to present the necessary statistical concepts, but we also aim at developing a formalism that is close enough to that introduced in Sec. II.3 for quantum-mechanical systems—as will e.g. be reflected in formally identical evolution equations for the probability densities or the observables [see Eqs. (II.12) and (II.28) or (II.16) and (II.38)]. In this section this might first seem to entail unnecessary complications; yet it will later turn out to be useful.

II.2.1 Description of classical systems and their evolution

II.2.1 a State of a classical system

Consider an isolated classical system of N identical pointlike particles in three-dimensional Euclidean space. The microstate of the system at a time t is entirely characterized by the $3N$ spatial coordinates q^1, \dots, q^{3N} of the particles and their $3N$ conjugate momenta p_1, \dots, p_{3N} . Together, these positions and momenta constitute a point in a $6N$ -dimensional space, the *phase space* (or Γ -*space*) of the system. Reciprocally, each point in this phase space corresponds to a possible microstate of the system.

For such a system, a measurable quantity—or (classical) *observable*—is defined as a phase-space function $O_N(\{q^a\}, \{p_a\})$ of the $6N$ variables. We shall only consider observables without explicit time dependence, i.e. the mathematical function O_N remains constant over time.

Remarks:

* A better notation for the N -particle phase space might be Γ_{6N} , which specifies the phase-space dimension, yet the denomination Γ -space is traditional and thus will be kept.

* If the particles are not pointlike, but possess internal degrees of freedom that can be described classically, then the formalism further applies taking into account these extra degrees of freedom. The q^a and p_a are then generalized positions and momenta.

II.2.1 b Evolution of the system

The time evolution of the system is represented in Γ -space by the trajectory $(\{q^a(t)\}, \{p_a(t)\})$ of the representative point, which describes a succession of microstates. The “velocity” tangent to this trajectory is the $6N$ -dimensional vector \mathbf{u} whose $6N$ components are the time derivatives $\{\dot{q}^a(t)\}, \{\dot{p}_a(t)\}$.

The dynamics of the system—or equivalently, of the representing point in Γ -space—is fully determined by specifying the time-independent Hamiltonian^(u) function $H_N(\{q^a\}, \{p_a\})$. More precisely, the trajectory $(\{q^a(t)\}, \{p_a(t)\})$ is governed by the *Hamilton equations*

$$\begin{aligned}\dot{q}^a(t) &\equiv \frac{dq^a(t)}{dt} = \frac{\partial H_N}{\partial p_a} = \{q^a, H_N\}, \\ \dot{p}_a(t) &\equiv \frac{dp_a(t)}{dt} = -\frac{\partial H_N}{\partial q^a} = \{p_a, H_N\}, \quad a = 1, \dots, 3N,\end{aligned}\tag{II.1}$$

^(u)W. R. HAMILTON, 1805–1865

where the derivatives of the Hamilton function, and accordingly the Poisson^(v) brackets, are to be computed at the point in Γ -space where the system sits at time t , i.e. at $(\{q^a = \mathbf{q}^a(t)\}, \{p_a = \mathbf{p}_a(t)\})$ for $1 \leq a \leq 3N$. The Poisson bracket of two functions f, g on phase space is itself a function on Γ defined by⁽²⁵⁾

$$\{f, g\} \equiv \sum_{a=1}^{3N} \left(\frac{\partial f}{\partial q^a} \frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a} \frac{\partial g}{\partial q^a} \right), \quad (\text{II.2})$$

where the arguments of the Poisson bracket and of the derivatives have been dropped for the sake of brevity.

It is important to realize that the Hamilton equations of motion are fully deterministic once the Hamilton function is fixed: given an initial condition $(\{q^a(t_0)\}, \{p_a(t_0)\})$, the microstate at any time t is uniquely determined by Eqs. (II.1)⁽²⁶⁾. Accordingly, there is only a single trajectory through each individual point of the Γ -space, so that the notation $\mathbf{u}(\{q^a\}, \{p_a\})$ is non-ambiguous.

II.2.2 Phase-space density

In a many-body system, the precise microstate corresponding at a given time to determined macroscopic properties—for example, given volume, particle number N and energy—is not exactly known. As a consequence, one introduces a probability distribution $\rho_N(t, \{q^a\}, \{p_a\})$ on the Γ -space, the N -particle *phase-space density*, which is as always non-negative and normalized to unity

$$\rho_N(t, \{q^a\}, \{p_a\}) \geq 0 \quad \forall \{q^a\}, \{p_a\} \quad \text{and} \quad \int_{\Gamma} \rho_N(t, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V} = 1, \quad (\text{II.3})$$

where the integral runs over the whole Γ -space. $\rho_N(t, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V}$ is the probability that the microstate of the system at time t lies in the infinitesimal volume element $d^{6N}\mathcal{V}$ around the point $(\{q^a\}, \{p_a\})$. The phase-space volume element $d^{6N}\mathcal{V}$ should represent a uniform measure on Γ , so that

$$d^{6N}\mathcal{V} = C_N \prod_{a=1}^{3N} dq^a dp_a, \quad (\text{II.4a})$$

with C_N a normalization factor. A possible choice for the latter—and admittedly the most natural—is simply $C_N = 1$. Another choice, which is less natural yet allows one to recover classical mechanics as a limiting case of quantum mechanics, is to adopt for a system of N *indistinguishable* particles the measure

$$d^{6N}\mathcal{V} = \frac{1}{N!} \prod_{a=1}^{3N} \frac{dq^a dp_a}{2\pi\hbar}, \quad (\text{II.4b})$$

with \hbar the reduced Planck^(w) constant. A further advantage of this choice is that $d^{6N}\mathcal{V}$ is dimensionless, and thus the probability density ρ_N as well.

Remark: To interpret probabilities as counting the number of favorable cases among all possible outcomes, Gibbs introduced the idea of mentally considering many copies of a system—which altogether constitute a *statistical ensemble*—, where the copied systems all have the same macroscopic properties, although the corresponding microstates differ.

After having introduced the phase-space density ρ_n , the position $(\{q^a(t)\}, \{p_a(t)\})$ in Γ -space at time t may be viewed as a $6N$ -dimensional random variable. Again, the microstate at t is *not* random if $(\{q^a(t=t_0)\}, \{p_a(t=t_0)\})$ is known at some time t_0 , but randomness enters due to our knowledge of the initial condition only on a statistical basis.

⁽²⁵⁾The sign convention for Poisson brackets is not universal... The choice taken here is the same as in Goldstein [24] or Arnold [25], while Landau & Lifshitz adopt the opposite convention [26].

⁽²⁶⁾This even holds irrespective of whether t lies in the future or in the past of t_0 .

^(v)S. POISSON, 1781–1840 ^(w)M. PLANCK, 1858–1947

In turn, the value taken at time t by an observable $O_N(\{q^a\}, \{p_a\})$, namely

$$O_N(t) \equiv O_N(\{q^a = \mathbf{q}^a(t)\}, \{p_a = \mathbf{p}_a(t)\}), \quad (\text{II.5})$$

is also a random variable, whose successive moments are given by the usual formulae (see App. [B](#)). For instance, the average value at time t is

$$\langle O_N(t) \rangle_t = \int_{\Gamma} O_N(\{q^a\}, \{p_a\}) \rho_N(t, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V} \quad (\text{II.6a})$$

and the variance reads

$$\sigma_{O_N}^2(t) = \langle O_N(t)^2 \rangle_t - \langle O_N(t) \rangle_t^2, \quad (\text{II.6b})$$

where the subscript t emphasizes the use of the phase-space density at time t in computing the expectation value.

Remark: Even though it has been assumed that the observable O_n has no explicit time dependence, the moments of the observable do depend on the instant at which they are computed.

II.2.3 Time evolution of the phase-space density

Consider a fixed volume \mathcal{V} of the N -particle phase space Γ , and let $\mathcal{N}(t)$ denote the number of particles inside that volume at time t . We can write the rate of change of this number in two alternative ways.

Expressing first the number of particles with the help of the phase-space density

$$\mathcal{N}(t) = N \int_{\mathcal{V}} \rho_N(t, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V}, \quad (\text{II.7})$$

one finds that $\mathcal{N}(t)$ changes because the phase-space density is evolving in time:

$$\frac{d\mathcal{N}(t)}{dt} = N \int_{\mathcal{V}} \frac{\partial \rho_N(t, \{q^a\}, \{p_a\})}{\partial t} d^{6N}\mathcal{V}. \quad (\text{II.8})$$

Alternatively, one can view the change in $\mathcal{N}(t)$ as due to the flow of particles through the surface $\partial\mathcal{V}$ enclosing the volume \mathcal{V} . Let $\mathbf{e}_n(\{q^a\}, \{p_a\})$ denote the unit vector normal to $\partial\mathcal{V}$ at a given point, oriented towards the exterior of \mathcal{V} , and $\mathbf{u}(\{q^a\}, \{p_a\})$ be the velocity-vector tangent to the trajectory passing through the point $(\{q^a\}, \{p_a\})$. One then has

$$\frac{d\mathcal{N}(t)}{dt} = -N \int_{\partial\mathcal{V}} \rho_N(t, \{q^a\}, \{p_a\}) \mathbf{u}(\{q^a\}, \{p_a\}) \cdot \mathbf{e}_n(\{q^a\}, \{p_a\}) d^{6N-1}\mathcal{S}.$$

The divergence theorem transforms the surface integral over $\partial\mathcal{V}$ into a volume integral over \mathcal{V} :

$$\frac{d\mathcal{N}(t)}{dt} = -N \int_{\mathcal{V}} \nabla \cdot [\rho_N(t, \{q^a\}, \{p_a\}) \mathbf{u}(\{q^a\}, \{p_a\})] d^{6N}\mathcal{V}, \quad (\text{II.9})$$

where ∇ denotes the ($6N$ -dimensional) gradient in phase space.

Equating Eqs. [\(II.8\)](#) and [\(II.9\)](#) and arguing that they hold for an arbitrary volume \mathcal{V} , one obtains the local conservation equation in Γ -space (for the sake of brevity, the variables will from now on be omitted)

$$\frac{\partial \rho_N}{\partial t} + \nabla \cdot (\rho_N \mathbf{u}) = 0. \quad (\text{II.10a})$$

The divergence term can be rewritten as

$$\begin{aligned} \nabla \cdot (\rho_N \mathbf{u}) &= \sum_{a=1}^{3N} \frac{\partial}{\partial q^a} (\rho_N \dot{q}^a) + \sum_{a=1}^{3N} \frac{\partial}{\partial p_a} (\rho_N \dot{p}_a) \\ &= \sum_{a=1}^{3N} \left(\frac{\partial \rho_N}{\partial q^a} \dot{q}^a + \frac{\partial \rho_N}{\partial p_a} \dot{p}_a \right) + \sum_{a=1}^{3N} \left(\frac{\partial \dot{q}^a}{\partial q^a} + \frac{\partial \dot{p}_a}{\partial p_a} \right) \rho_N. \end{aligned}$$

With the help of the Hamilton equations (II.1) this gives the *Liouville*^(x) equation⁽²⁷⁾

$$\frac{\partial \rho_N}{\partial t} + \sum_{a=1}^{3N} \left(\frac{\partial \rho_N}{\partial q^a} \frac{\partial H_N}{\partial p_a} - \frac{\partial \rho_N}{\partial p_a} \frac{\partial H_N}{\partial q^a} \right) = \frac{\partial \rho_N}{\partial t} + \{\rho_N, H_N\} = 0. \quad (\text{II.10b})$$

Introducing the *Liouville operator* (or *Liouvillian*) \mathcal{L} defined by⁽²⁸⁾

$$i\mathcal{L} \equiv \sum_{a=1}^{3N} \left(\frac{\partial H_N}{\partial p_a} \frac{\partial}{\partial q^a} - \frac{\partial H_N}{\partial q^a} \frac{\partial}{\partial p_a} \right) = \{ \cdot, H_N \}, \quad (\text{II.11})$$

where the dot stands for the phase-space function on which the operator acts, Eq. (II.10b) can be recast in the form

$$\frac{\partial \rho_N}{\partial t} + i\mathcal{L}\rho_N = 0. \quad (\text{II.12})$$

Like the Hamilton function, the Liouville operator for an isolated system is time-independent, which allows one to formally integrate the Liouville equation as

$$\rho_N(t, \{q^a\}, \{p_a\}) = e^{-i\mathcal{L}t} \rho_N(t=0, \{q^a\}, \{p_a\}), \quad (\text{II.13})$$

with $\rho_N(t=0, \{q^a\}, \{p_a\})$ the initial phase-space density at $t=0$. To account for this result, $e^{-i\mathcal{L}t}$ is sometimes called *time propagation operator*.

Remarks:

* An equivalent formulation of the Liouville equation, which follows from Eq. (II.10a) under consideration of the Hamilton equations, which yield $\nabla \cdot \mathbf{u} = 0$ (see the identity in footnote 27), is

$$\frac{\partial \rho_N}{\partial t} + \mathbf{u} \cdot \nabla \rho_N = \frac{D\rho_N}{Dt} = 0, \quad (\text{II.14})$$

with $\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla$ the *material* (or *convective*, *substantial*, *hydrodynamic*) derivative.

* Equations (II.10b), (II.12) or (II.14) represent the Eulerian^(y) viewpoint on the Liouville equation. Alternatively, one can adopt the Lagrangian^(z) viewpoint and follow individual trajectories in Γ -space in their motion.

More precisely one should consider a continuous distribution of microstates, to ensure that the phase-space volume they occupy at a given instant has a non-zero measure. This collection of Γ -space points is then sometimes referred to as a *phase(-space) fluid*, and its motion—along the corresponding trajectories—as the *flow* of that fluid.

The corresponding statement of the Liouville equation, which is then known as *Liouville theorem*, is the following:

$$\textit{The volume in phase space occupied by a collection of microstates for a system obeying the Hamilton equations of motion remains constant in time.} \quad (\text{II.15})$$

That is, the volume of the phase-space fluid is an integral constant of the motion. Accordingly, one often states that the flow of trajectories in phase space is incompressible.

⁽²⁷⁾Here it is implicitly assumed that the Hamilton function H_N is sufficiently regular—namely that the second partial derivatives are continuous—so as to have the identity

$$\frac{\partial \dot{q}^a}{\partial q^a} = \frac{\partial}{\partial q^a} \frac{\partial H_N}{\partial p_a} = \frac{\partial}{\partial p_a} \frac{\partial H_N}{\partial q^a} = -\frac{\partial \dot{p}_a}{\partial p_a}.$$

⁽²⁸⁾The conventional—and not universally adopted—factor i has the advantage that it leads to results that are easily compared to the quantum-mechanical ones.

^(x)J. LIOUVILLE, 1809–1882 ^(y)L. EULER, 1707–1783 ^(z)J.-L. LAGRANGE, 1736–1813

* The Liouville equation (II.10b) or (II.12) shows that if the phase-space density ρ_N is a function of the Hamilton function H_N only, then it is stationary. This is for instance the case at thermodynamic equilibrium⁽²⁹⁾, but not only.

II.2.4 Time evolution of macroscopic observables

As mentioned above, classical observables are defined as time-independent functions on phase space, which however acquire an implicit time dependence when computed along the Γ -space trajectory of a system, see Eq. (II.5). Differentiating $O_N(t)$ with the chain rule gives

$$\frac{dO_N(t)}{dt} = \sum_{a=1}^{3N} \left[\frac{\partial O_N}{\partial q^a} \dot{q}^a(t) + \frac{\partial O_N}{\partial p_a} \dot{p}_a(t) \right],$$

where the derivatives with respect to the Γ -space coordinates are taken at the point $(\{q^a(t)\}, \{p_a(t)\})$ along the system phase-space trajectory. Under consideration of the Hamilton equations (II.1), this derivative becomes

$$\frac{dO_N}{dt} = \sum_{a=1}^{3N} \left(\frac{\partial O_N}{\partial q^a} \frac{\partial H_N}{\partial p_a} - \frac{\partial O_N}{\partial p_a} \frac{\partial H_N}{\partial q^a} \right) = \{O_N, H_N\} = i\mathcal{L}O_N, \quad (\text{II.16})$$

where we have used the definitions of the Poisson bracket (II.2) and of the Liouville operator (II.11).

Invoking again the time-independence of the Liouville operator for an isolated system, the differential equation (II.16) can formally be integrated as

$$O_N(t) = e^{i\mathcal{L}t} O_N(t=0), \quad (\text{II.17})$$

with $O_N(t=0)$ the initial value of the observable at time $t = 0$.

The expectation value of the observable, obtained by averaging over possible initial conditions at $t = 0$ then reads

$$\begin{aligned} \langle O_N(t) \rangle_0 &= \int_{\Gamma} \rho_N(t=0, \{q^a\}, \{p_a\}) O_N(t) d^{6N}\mathcal{V} \\ &= \int_{\Gamma} \rho_N(t=0, \{q^a\}, \{p_a\}) e^{i\mathcal{L}t} O_N(t=0) d^{6N}\mathcal{V}. \end{aligned} \quad (\text{II.18})$$

Alternatively, one can directly average $O_N(\{q^a\}, \{p_a\})$ with the phase-space density at time t , as done in Eq. (II.6a):

$$\langle O_N(t) \rangle_t = \int_{\Gamma} O_N(\{q^a\}, \{p_a\}) \rho_N(t, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V}.$$

Using Eq. (II.13), this becomes

$$\langle O_N(t) \rangle_t = \int_{\Gamma} O_N(\{q^a\}, \{p_a\}) e^{-i\mathcal{L}t} \rho_N(t=0, \{q^a\}, \{p_a\}) d^{6N}\mathcal{V}. \quad (\text{II.19})$$

Both points of view actually yield the same result—i.e. $\langle O_N(t) \rangle_0 = \langle O_N(t) \rangle_t$ —, which means that one can attach the time dependence either to the observables or to the phase-space density, which stands for the macrostate of the system.

The equivalence can be seen by computing the time derivatives of Eqs. (II.18)—which is the average of Eq. (II.16) over initial positions—and (II.19). Replacing the Liouville operator by its expression in terms of the Hamilton function, and performing a few partial integrations to handle the Poisson brackets, one finds that these time derivatives coincide at any time t . Since the “initial” conditions at $t = 0$ also coincide, the identity of $\langle O_N(t) \rangle_0$ and $\langle O_N(t) \rangle_t$ follows.

⁽²⁹⁾... in which case $\rho_N \propto e^{-\beta H_N}$ with $\beta \equiv 1/k_B T$.

Remark: More generally, the identity

$$\int_{\Gamma} g^*({q^a}, {p_a}) \mathcal{L}h({q^a}, {p_a}) d^{6N}\mathcal{V} = \int_{\Gamma} (\mathcal{L}g)^*({q^a}, {p_a}) h({q^a}, {p_a}) d^{6N}\mathcal{V},$$

holds for every pair of phase-space functions $g({q^a}, {p_a})$ and $h({q^a}, {p_a})$ that vanish sufficiently rapidly at infinity, where f^* denotes the complex conjugate function to f . Recognizing in the phase-space integral of g^*h an inner product $\langle g, h \rangle$, the above identity can be recast as $\langle g, \mathcal{L}h \rangle = \langle \mathcal{L}g, h \rangle$, which expresses the fact that the Liouville operator is *Hermitian* for the inner product. In turn, the operators $e^{\pm it\mathcal{L}}$, which govern the evolution of the density ρ_N or of observables $O_N(t)$, are *unitary* for this product, i.e. $\langle g, e^{-it\mathcal{L}}h \rangle = \langle e^{it\mathcal{L}}g, h \rangle$ or equivalently:

$$\int_{\Gamma} g({q^a}, {p_a}) e^{-it\mathcal{L}}h({q^a}, {p_a}) d^{6N}\mathcal{V} = \int_{\Gamma} \left[e^{it\mathcal{L}}g({q^a}, {p_a}) \right] h({q^a}, {p_a}) d^{6N}\mathcal{V}. \quad (\text{II.20})$$

II.2.5 Fluctuating number of particles

Until now, we have assumed that the particle number N is exactly known. It is however often not the case, so that N also becomes a random variable, with a discrete probability distribution.

The formalism can easily be generalized to accommodate for this possibility. The new phase space Γ is the union—to be precise, the direct sum—of the individual N -particle phase spaces Γ_{6N} (see first Remark in § II.2.1 a) for every acceptable value of N , i.e. for $N \in \mathbb{N}$.⁽³⁰⁾ The probability density ρ on this phase space consists of (the tensor product of) densities $\tilde{\rho}_N$ proportional to the respective N -particle densities ρ_N , yet normalized so that

$$\pi_N = \int \tilde{\rho}_N(t, {q^a}, {p_a}) d^{6N}\mathcal{V}$$

represents the probability to have N particles in the system at time t .

An observable O is also defined as a tensor product of functions O_N on each N -particle phase space, with the expectation value

$$\langle O(t) \rangle_t = \sum_{N=0}^{\infty} \int O_N({q^a}, {p_a}) \tilde{\rho}_N(t, {q^a}, {p_a}) d^{6N}\mathcal{V}.$$

⁽³⁰⁾The case $N = 0$ has to be considered as well, corresponding here to a 0-dimensional phase space reduced to a single point.