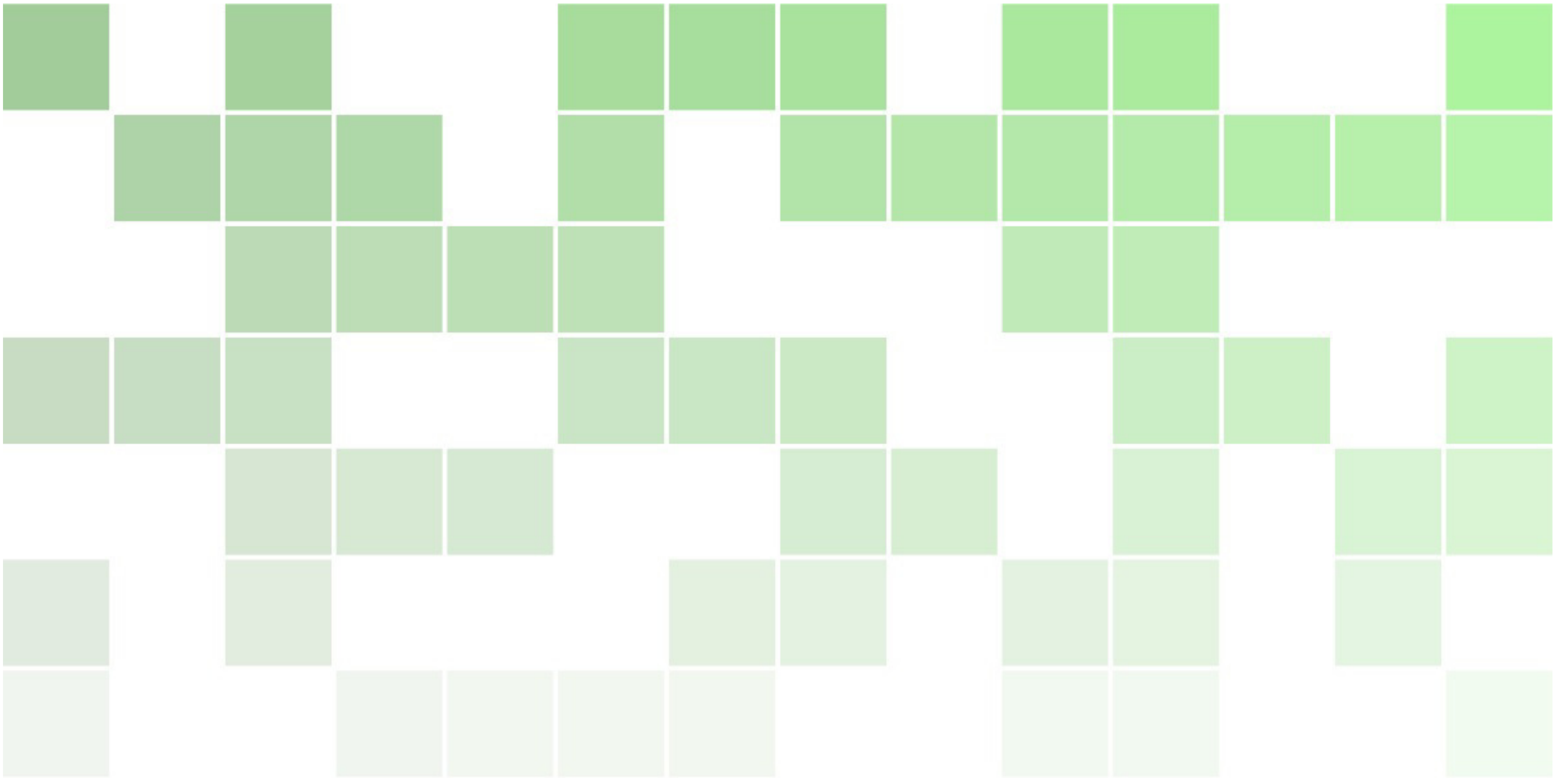

Topics in Nonequilibrium Physics

Nicolas Borghini

Version of September 8, 2016

Nicolas Borghini
Universität Bielefeld, Fakultät für Physik
Homepage: `http://www.physik.uni-bielefeld.de/~borghini/`
Email: borghini at physik.uni-bielefeld.de

Contents

Introduction

Notations, conventions, etc.

Bibliography

(in alphabetical order)

- de Groot & Mazur, *Nonequilibrium thermodynamics* [1]
- Kubo, Toda, & Hashitsume, *Statistical physics II: Nonequilibrium statistical physics* [2]
- Landau & Lifshitz, *Course of theoretical physics*:
  Vol. 5: Statistical physics, part 1 [3];
  Vol. 9: Statistical physics, part 2 [4];
  Vol. 10: Physical kinetics [5]
- Pottier, *Nonequilibrium statistical physics* [6]
- Zwanzig, *Nonequilibrium statistical mechanics* [7]

CHAPTER I

Thermodynamics of irreversible processes

Thermodynamics is a powerful generic formalism, which provides a description of physical systems involving many degrees of freedom in terms of only a small number of salient variables like the system's total energy, particle number or volume, irrespective of the actual underlying microscopic dynamics The strength of the approach is especially manifest for the so-called equilibrium states, which in a statistical-physical interpretation are the "most probable" macroscopic states—i.e. those which correspond to the largest number of microscopic states obeying given constraints—, and are characterized by only a handful of thermodynamic variables. Accordingly, first courses in thermodynamics chiefly focus on its equilibrium aspects. In that context, when considering physical transformations of a system across different macrostates, one mostly invokes "quasi-static processes", namely fictive continuous sequences of equilibrium states between the initial and final states.

An actual physical process in a macroscopic system is however not quasi-static, but rather involves intermediary macrostates that are not at thermodynamic equilibrium. As a result, the evolution is accompanied by an increase in the total entropy of the system, or more precisely, of the smallest whole which includes the system under study and its environment which is isolated from the rest of the universe. That is, such an out-of-equilibrium process is irreversible.

Similar departures from equilibrium also appear spontaneously in an equilibrated system, when at least one of its key thermodynamic variables is not exactly fixed—as is the case when the system cannot exchange the corresponding physical quantity with its exterior—, but only known "on average"—as happens when the system can exchange the relevant quantity with an external reservoir. In the latter case, the thermodynamic variable will possibly fluctuate around its expectation value,(1) which will again momentarily drive the system out of equilibrium.

In either case, it is necessary to consider also non-equilibrated thermodynamic systems, which constitute the topic of the present chapter. In a first step, the macroscopic variables necessary to describe such out-of-equilibrium systems as well as the processes which drive them to equilibrium are presented (Sec. I.1). Making physical assumptions on how far away the systems are from equilibrium and on the processes they undergo, one can postulate constitutive equations that relate the newly introduced variables with each other (Sec. I.2), irrespective of any microscopic picture. These relations, which actual encompass several known phenomenological laws, involve characteristic properties of the systems, namely their transport coefficients. The calculation of the latter, like that of thermodynamic coefficients, falls outside the realm of thermodynamics and necessitates more microscopical approaches, as will be presented in the following chapters.

For simplicity, the discussion is restricted to non-relativistic systems. ...but I am willing to change that, if I find the time.

---

(1) The standard deviation of these fluctuations is readily computed in statistical mechanics, by taking second derivatives of the logarithm of the relevant partition function, and involves thermodynamic coefficients like the compressibility or the specific heat.

I.1 Description of irreversible thermodynamic processes

This section is devoted to introducing the quantities needed to describe nonequilibrated systems irreversible processes at the macroscopic level. First, the laws of equilibrium thermodynamics, or *thermostatics*, are recalled (Sec. I.1.1), using Callen's approach [8, Chap. 1], which is closer to statistical mechanics than that starting from the traditional principles, thereby allowing one more easily to treat macroscopic physical quantities as effective concepts. After that, the novel variables that play a role in a situation of departure from (global) thermodynamic equilibrium in a system are presented, starting with the simpler case of discrete systems (Sec. I.1.2), then going on to the physically richer continuous media (Sec. I.1.3 and I.1.4).

I.1.1 Reminder: Postulates of equilibrium thermodynamics

Instead of using the traditional laws of thermodynamics—which for the sake of completeness will be quickly recalled at the end of this subsection—it is possible to give an alternative formulation, due to Herbert Callen(a) [8], which turns out to be totally equivalent and has the advantage of being more readily extended to out-of-equilibrium thermodynamics.

This approach takes as granted the existence of variables—namely the volume $\mathcal{V}$, the chemical composition N_1, $N_2, \ldots$, N_r and the internal energy U—to characterize properties of "simple" thermodynamic systems at rest, where "simple" means macroscopically homogeneous and isotropic, chemically inert and electrically neutral. All these variables, which will hereafter be collectively represented by $\{\mathcal{X}_a\}$, are *extensive*: for a system whose volume $\mathcal{V}$ become infinitely large, they all diverge in such a way that the ratio $\mathcal{X}_a/\mathcal{V}$ remains finite.

Remark: Interestingly enough, the extensive variables $\{\mathcal{X}_a\}$ are, with the exception of volume, all conserved quantities in isolated (and a fortiori closed) chemically inert systems, which somehow justifies the special role that they play.

Building upon the variables $\{\mathcal{X}_a\}$, thermostatics follows from four postulates:

- According to postulate I, there exist particular macroscopic states of simple systems at rest, the *equilibrium states*, that are *fully* characterized by the variables U, $\mathcal{V}$, $N_1, \ldots, N_r$.
- The three remaining postulates specify the characterization of the equilibrium state among all macrostates with the same values of the parameters $\{\mathcal{X}_a\}$:
  - Postulate II: For the equilibrium states of a composite system—defined as a collection of simple systems, hereafter labeled with a capital superscript (A)—, there exists a function of the extensive parameters $\{\mathcal{X}_a^{(A)}\}$ of the subsystems, the *entropy* S, which is *maximal* with respect to free variations of the variables $\{\mathcal{X}_a^{(A)}\}$.
  - Postulate III: The entropy of a composite system is the sum of the entropies of its subsystems. Additionally, S is continuous and differentiable, and is a monotonically increasing function of U.
  - Postulate IV: The entropy of any system vanishes in the state for which $\left(\dfrac{\partial U}{\partial S}\right)_{\mathcal{V},N_1,\ldots,N_r} = 0$.

Noting that any simple system can be in thought considered as a composite system of arbitrarily chosen subparts, the second postulate provides a variational principle for finding equilibrium states.

The generalization of these postulates to more complicated systems, e.g. magnetized systems or systems in which chemical reactions take place, is quite straightforward.

In this formulation, a special role is played by the entropy—which is actually only defined for equilibrium states. The functional relationship between S and the other characteristic variables,

---

(a) H. Callen, 1919–1993

$S = S(\{\mathcal{X}_a\})$, is referred to as the *fundamental equation*,(2) and contains every information upon the thermodynamic properties of the system at equilibrium. The differential form of the relation is *Gibbs'*(b) *fundamental equation*

$$\boxed{\mathrm{d}S = \sum_a \frac{\partial S}{\partial \mathcal{X}_a}\,\mathrm{d}\mathcal{X}_a = \sum_a \mathcal{Y}_a\,\mathrm{d}\mathcal{X}_a \quad \text{with} \quad \mathcal{Y}_a \equiv \left(\frac{\partial S}{\partial \mathcal{X}_a}\right)_{\{\mathcal{X}_b\}_{b\neq a}}.} \tag{I.1}$$

The partial derivatives $\mathcal{Y}_a$ are intensive parameters conjugate to the extensive variables. Mathematically these derivatives depend on the same set of variables $\{\mathcal{X}_b\}$ as the entropy; the functional relationships $\mathcal{Y}_a = \mathcal{Y}_a(\{\mathcal{X}_b\})$ are the *equations of state* of the system.

For instance, in the case of a simple multicomponent fluid, the fundamental equation reads in integral form $S = S(U, \mathcal{V}, N_1, \dots, N_r)$, and in differential form

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}U + \frac{\mathcal{P}}{T}\,\mathrm{d}\mathcal{V} - \sum_k \frac{\mu_k}{T}\,\mathrm{d}N_k, \tag{I.2a}$$

that is(3)

$$\mathcal{Y}_E = \frac{1}{T}, \qquad \mathcal{Y}_{\mathcal{V}} = \frac{\mathcal{P}}{T}, \qquad \mathcal{Y}_{N_k} = -\frac{\mu_k}{T}, \tag{I.2b}$$

with T the *temperature*, $\mathcal{P}$ the (thermodynamic) *pressure*, and μ_k the *chemical potential* for species k.

Remark: Postulate I explicitly deals with systems at rest, that is, with a vanishing total linear momentum $\vec{P}$. Since $\vec{P}$ is also a conserved quantity in isolated systems, like internal energy or particle number, it is tempting to add it to the list of characteristic parameters $\{\mathcal{X}_a\}$.

Now, a system at thermodynamic equilibrium is a fortiori in mechanical equilibrium, that is, there is no macroscopic motion internal to the system: a finite linear momentum $\vec{P}$ is thus entirely due to some global motion of the system, with a velocity $\vec{\mathsf{v}} = \vec{P}/M$, where M denotes the mass of the system. Equivalently, the finite value of total momentum arises from our describing the system within a reference frame in motion with velocity $-\vec{\mathsf{v}}$ with respect to the system rest frame. The only interest in considering $\vec{P}$ among the basic parameters is that it allows us to find the conjugate intensive parameter, which will prove useful hereafter.(4)

Relying momentarily on the statistical mechanical interpretation of entropy as a measure of missing information, the entropy of a system of mass M does not change whether it is at rest [energy $E = U$, entropy $S(U, \vec{P}=\vec{0}) = S_0(U)$] or in collective motion with momentum $\vec{P}$, in which case its energy becomes $E = U + \vec{P}^2/2M$ and its entropy $S(E, \vec{P})$. One thus has

$$S(E, \vec{P}) = S_0\bigg(E - \frac{\vec{P}^{\,2}}{2M}\bigg).$$

Differentiating this identity with respect to one of the component P^i of momentum in a given coordinate system, there comes the conjugate variable

$$\mathcal{Y}_{P^i} \equiv \frac{\partial S}{\partial P^i} = \frac{\partial S_0}{\partial P^i} = -\frac{P_i}{M}\frac{\partial S_0}{\partial U} = -\frac{\mathsf{v}_i}{T}, \tag{I.3}$$

where v_i denotes the i-th component of the velocity $\vec{\mathsf{v}}$ of the system. One easily checks that the other intensive parameters $\mathcal{Y}_E$, $\mathcal{Y}_{\mathcal{V}}$, $\mathcal{Y}_{N_k}$ remain unchanged even if the system is in motion, which justifies a posteriori the notation convention mentioned in footnote 3.

---

(2) More precisely, it is the fundamental equation in "entropy representation".

(3) Throughout these notes, quantities related to the internal energy U—as here its conjugate variable or later below the corresponding affinity or the internal energy per unit volume—will be denoted with the letter E, instead of U.

(4) It is also more natural in order to allow the extension of the formalism to relativistic systems, since the energy alone is only a single component of a 4-vector.

(b) J. W. Gibbs, 1839–1903

For the sake of completeness, we recall here the "classical" laws of thermodynamics, as can be found in most textbooks:

– 0th law (often unstated): thermal equilibrium at a temperature T is a transitive property;

– 1st law: a system at equilibrium is characterized by its internal energy U; the changes of the latter are due to heat exchange Q with and/or macroscopic work W from the exterior, $\Delta U = Q + W$;

– 2nd law: a system in equilibrium is characterized by its entropy S; in an infinitesimal process between two equilibrium states, $\mathrm{d}S \geq \delta Q/T$, where the identity holds if and only if the process is quasi-static;

– 3rd law: for any system, $S \to S_0$ when $T \to 0^+$, where S_0 is independent of the system variables (which allows one to take $S_0 = 0$).

Callen discusses the equivalence between his postulates and these laws in Chapters 1, 2 & 4 of Ref. [8]. For instance, one sees at once that the third traditional law and the fourth postulate are totally equivalent.

I.1.2 Irreversible processes in discrete thermodynamic systems

The first of Callen's postulates of thermostatics has two implicit corollaries, namely the existence of other states of macroscopic systems than the equilibrium ones, and the necessity to introduce new quantities besides the extensive parameters $\{\mathcal{X}_a\}$ for the description of these out-of-equilibrium states. In this subsection, these extra variables are introduced for the case of discrete systems.

I.1.2 a Timescales

In the following, we shall consider composite systems made of several simple subsystems, each of which is characterized by a set of extensive parameters $\{\mathcal{X}_a^{(A)}\}$, where (A) labels the various subsystems. If the subsystems are isolated from each other, they can individually be in thermodynamic equilibrium, with definite values of the respective variables. Beginning with such a collection of equilibrium states and connecting the subsystems together, i.e. allowing them to interact with each other, the subsystems start evolving. This results in a time dependence of the extensive variables, $\{\mathcal{X}_a^{(A)}(t)\}$.

An essential assumption is that the interaction processes between macroscopic systems are slow compared to the microscopic ones within the individual subsystems, which drive each of them to its own thermodynamic equilibrium. In other terms, the characteristic timescales for macroscopic processes, i.e. for the evolutions of the $\{\mathcal{X}_a^{(A)}(t)\}$, are much larger than the typical timescale of microscopic interactions.

Under this assumption, one can consider that the composite system undergoes a transformation across macroscopic states such that one can meaningfully define an instantaneous entropy $S(\{\mathcal{X}_a(t)\})$, with the same functional form as in thermodynamic equilibrium.

I.1.2 b Affinities and fluxes in a discrete system

Consider an isolated composite system made of two simple systems A and B, with respective extensive variables $\{\mathcal{X}_a^{(A)}\}$, $\{\mathcal{X}_a^{(B)}\}$. Since the latter correspond to conserved quantities, the sum

$$\mathcal{X}_a^{(A)}(t) + \mathcal{X}_a^{(B)}(t) = \mathcal{X}_a^{\mathrm{tot}} \tag{I.4}$$

remains constant over time for every a, whether the subsystems A and B can actually exchange the quantity or not.

According to postulate III, the entropy of the composite system is

$$S^{\mathrm{tot}}(t) = S^{(A)}\big(\{\mathcal{X}_a^{(A)}(t)\}\big) + S^{(B)}\big(\{\mathcal{X}_a^{(B)}(t)\}\big). \tag{I.5}$$

Since for fixed $\mathcal{X}_a^{\mathrm{tot}}$, $\mathcal{X}_a^{(B)}(t)$ is entirely determined by the value of $\mathcal{X}_a^{(A)}(t)$ [Eq. (I.4)], the total entropy S^{tot} is actually only a function of the latter.

In thermodynamic equilibrium, $S^{\rm tot}(\{\mathcal{X}_a^{(A)}\})$ is maximal (second postulate), i.e. its derivative with respect to $\mathcal{X}_j^{(A)}$ should vanish:

$$\frac{\partial S^{\rm tot}}{\partial \mathcal{X}_a^{(A)}}\bigg|_{\mathcal{X}_a^{\rm tot}} = 0 = \frac{\partial S^{(A)}}{\partial \mathcal{X}_a^{(A)}} - \frac{\partial S^{(B)}}{\partial \mathcal{X}_a^{(B)}} = \mathcal{Y}_a^{(A)} - \mathcal{Y}_a^{(B)}. \tag{I.6}$$

Thus in thermodynamic equilibrium the so-called *affinity*

$$\boxed{\mathscr{F}_a \equiv \mathcal{Y}_a^{(A)} - \mathcal{Y}_a^{(B)} = \frac{\partial S^{\rm tot}}{\partial \mathcal{X}_a^{(A)}}\bigg|_{\mathcal{X}_a^{\rm tot}}} \tag{I.7}$$

conjugate to the extensive state variable $\mathcal{X}_a$ vanishes. Reciprocally, when $\mathscr{F}_a \neq 0$, the system is out of equilibrium. A process then takes places, that drives the system to equilibrium: $\mathscr{F}_a$ thus acts as a *generalized force* (and is sometimes referred to as such).

For instance, unequal temperatures result in a non-zero affinity $\mathscr{F}_E \equiv 1/T^{(A)} - 1/T^{(B)}$, and similarly one has for simple systems $\mathscr{F}_{\mathcal{V}} \equiv \mathcal{P}^{(A)}/T^{(A)} - \mathcal{P}^{(B)}/T^{(B)}$ and for every species k (note the signs!) $\mathscr{F}_{N_k} \equiv \mu_k^{(B)}/T^{(B)} - \mu_k^{(A)}/T^{(A)}$.

The response of a system to a non-vanishing affinity $\mathscr{F}_a$ is quite naturally a variation of the conjugate extensive quantity $\mathcal{X}_a^{(A)}$. This response is described by a *flux*, namely the rate of change

$$\boxed{\mathcal{J}_a \equiv \frac{{\rm d}\mathcal{X}_a^{(A)}}{{\rm d}t},} \tag{I.8}$$

which describes how much of quantity $\mathcal{X}_a$ is transferred from system B to system A per unit time. These fluxes are sometimes referred to as *generalized displacements*.

Remarks:

∗ The sign convention for the affinity is not universal: some authors define it as $\mathscr{F}_a \equiv \mathcal{Y}_j^{(B)} - \mathcal{Y}_a^{(A)}$, e.g. in Ref. [9], instead of Eq. (I.7). Accordingly, the conjugate flux is taken as the quantity transferred from system A to system B per unit time, that is the opposite of Eq. (I.8). All in all, the equation for the entropy production rate (I.9) below remains unchanged.

∗ In the case of discrete systems, all affinities and fluxes are scalar quantities.

I.1.2 c Entropy production

The affinities and fluxes introduced above allow one to rewrite the time derivative of the instantaneous entropy in a convenient way. Thus, differentiating $S^{\rm tot}$ with respect to time yields the *rate of entropy production*

$$\frac{{\rm d}S^{\rm tot}}{{\rm d}t} = \sum_a \frac{\partial S^{\rm tot}}{\partial \mathcal{X}_a^{(A)}} \frac{{\rm d}\mathcal{X}_a^{(A)}}{{\rm d}t}$$

that is, using definitions (I.7) and (I.8) of the affinities and fluxes,

$$\boxed{\frac{{\rm d}S^{\rm tot}}{{\rm d}t} = \sum_a \mathscr{F}_a \mathcal{J}_a.} \tag{I.9}$$

An important property of this rate is its bilinear structure, which will remain valid in the case of a continuous medium, and also allows one to identify the affinities and fluxes in the study of non-simple systems.

I.1.3 Local thermodynamic equilibrium of continuous systems

We now turn to the description of out-of-equilibrium macroscopic systems which can be viewed as continuous media, starting with the determination of the thermodynamic quantities suited to that case.

I.1.3 a Local thermodynamic variables

The starting point when dealing with an inhomogeneous macroscopic system is to divide it in thought in small cells of fixed—yet not necessarily universal—size fulfilling two conditions

- each cell can meaningfully be treated as a thermodynamic system, i.e. each cell must be large enough that the relative fluctuation of the usual thermodynamic quantities computed in the cell are negligible;
- the thermodynamic properties vary little over the cell scale, i.e. cells cannot be too large, so that (approximate) homogeneity is restored.

Under these assumptions, one can define *local thermodynamic variables*, corresponding to the values taken in each cell—labelled by its position $\vec{r}$—by the extensive parameters: $U(\vec{r})$, $N_k(\vec{r})$, … Since the size of each cell is physically irrelevant as long as it satisfies the above two conditions, there is no local variable corresponding to the volume $\mathcal{V}$, which only enters the game as the domain over which $\vec{r}$ takes its values.

On the other hand, since the separation between cells is immaterial, nothing prevents matter from actually flowing from a cell to its neighbours; one thus needs additional extensive parameters to describe this motion, namely the 3 components of the total momentum $\vec{P}(\vec{r})$ of the particles in each cell. For an isolated system, total momentum is a conserved quantity, as are the energy and (in the absence of chemical reactions) the particle numbers, thus $\vec{P}$ is on the same footing as U and the N_k.

Promoting $\vec{r}$ to a continuous variable, these local thermodynamic parameters become *fields*. To account for their possible time dependence, the latter will collectively be denoted as $\{X_a(t,\vec{r})\}$.

Remarks:

∗ The actual values of $\vec{P}(t,\vec{r})$ obviously depend on the reference frame chosen for describing the system.

∗ As always in field theory, one relies on a so-called *Eulerian*[(c)] description, in which one studies the changes in the thermodynamic variables with time at a given position, irrespective of the fact that the microscopic particles in a given cell do not remain the same over time, but constantly move from one cell to the other.

Rather than relying on the local thermodynamic variables, which depend on the arbitrary size of the cells, it is more meaningful to introduce their densities, i.e. the amounts of the quantities per unit volume: internal energy density $e(t,\vec{r})$, particle number densities $n_k(t,\vec{r})$, momentum density $\vec{p}(t,\vec{r})$… Except for the internal energy (see footnote 3), these densities will be denoted by the corresponding lowercase letter, and thus collectively referred to as $\{x_a(t,\vec{r})\}$.

Alternatively, one can also consider the quantities per unit mass,[(5)] which will be denoted in lowercase with a subscript m—$e_{\rm m}(t,\vec{r})$, $n_{k,{\rm m}}(t,\vec{r})$, …, and collectively $\{x_{a,{\rm m}}(t,\vec{r})\}$—, with the exception of the momentum per unit mass, which is called *flow velocity* and will be denoted by $\vec{\mathrm{v}}(t,\vec{r})$. Representing the mass density at time t at position $\vec{r}$ by $\rho(t,\vec{r})$, one trivially has the identity

$$x_a(t,\vec{r}) = \rho(t,\vec{r})\, x_{a,{\rm m}}(t,\vec{r}) \tag{I.10}$$

for every extensive parameter X_a.

---

(5) The use of these so-called specific quantities is for example favoured by Landau & Lifshitz.

(c) L. Euler, 1707–1783

I.1.3 b Local entropy

Since it is assumed that each small cell is at every instant t in a state of thermodynamic equilibrium, one can meaningfully associate to it a *local entropy* $S(t,\vec{r})$. From there, one can define the *local entropy density* $s(t,\vec{r})$ and the specific entropy $s_{\rm m}(t,\vec{r})$.

The important *local equilibrium assumption* amounts to postulating that the dependence of $s(t,\vec{r})$ on the thermodynamic densities $\chi_a(t,\vec{r})$—which momentarily include a "local volume density" $\chi_{\mathcal{V}}$ identically equal to 1—is given by the same fundamental equation as between the entropy S and its extensive parameters $\{\mathcal{X}_a\}$ in a system in thermodynamic equilibrium.

In differential form, this hypothesis leads to the total differential [cf. Eq. (I.1)]

$$\mathrm{d}s(t,\vec{r}) = {\sum_a}' \mathcal{Y}_a(t,\vec{r})\,\mathrm{d}\chi_a(t,\vec{r}). \tag{I.11}$$

Since the differential $\mathrm{d}\chi_{\mathcal{V}}$ is actually identically zero, the intensive parameter $\mathcal{Y}_{\mathcal{V}}$ conjugate to volume actually drops out from the sum, which is indicated by the primed sum sign.

Considering instead the integral form of the fundamental equation, and after summing over cells—i.e. technically integrating over the volume of the system—, the total entropy of the continuous medium reads

$$S^{\rm tot}(t) = \sum_j \int_{\mathcal{V}} \mathcal{Y}_a(t,\vec{r})\,\chi_a(t,\vec{r})\,\mathrm{d}^3\vec{r}. \tag{I.12}$$

Equations (I.11) and (I.12) yield for the local intensive variables

$$\mathcal{Y}_a(t,\vec{r}) = \frac{\partial s(t,\vec{r})}{\partial \chi_a(t,\vec{r})} = \frac{\delta S^{\rm tot}(t)}{\delta \chi_a(t,\vec{r})}, \tag{I.13}$$

i.e. $\mathcal{Y}_a(t,\vec{r})$ can be seen either as a partial derivative of the local entropy density, or as a functional derivative of the total entropy. The relations

$$\mathcal{Y}_a(t,\vec{r}) = \mathcal{Y}_a\big(\{\chi_b(t,\vec{r})\}\big) \tag{I.14}$$

for the various $\mathcal{Y}_a(t,\vec{r})$ are the *local equations of state* of the system. One easily checks that the local equilibrium hypothesis amounts to assuming that the local equations of state have the same form as the equations of state of a system in *global* thermodynamic equilibrium.

In the example of a simple system in mechanical equilibrium—so that $\vec{\mathrm{v}}(t,\vec{r})$ vanishes at each point—, Eq. (I.11) reads [cf. Eq. (I.2a)]

$$\mathrm{d}s(t,\vec{r}) = \frac{1}{T(t,\vec{r})}\,\mathrm{d}e(t,\vec{r}) - \sum_k \frac{\mu_k(t,\vec{r})}{T(t,\vec{r})}\,\mathrm{d}n_k(t,\vec{r}), \tag{I.15}$$

which defines the local temperature $T(t,\vec{r})$ and chemical potentials $\mu_k(t,\vec{r})$.

Remark: Throughout, "local" actually means "at each place at a given instant". A more accurate denomination would be to refer to "local and instantaneous" thermodynamic variables, which is however never done.

I.1.4 Affinities and fluxes in a continuous medium

We can now define the affinities and fluxes inside a continuous thermodynamic system. For that purpose, we first introduce the local formulation of balance equations in such a system, which relies on flux densities (§ I.1.4 a). Building on these, we consider the specific case of entropy balance (§ I.1.4 b), and deduce from it the form of the affinities (§ I.1.4 c). Throughout this section, the system is studied within its own rest frame, i.e. it is globally at rest.

I.1.4 a Balance equations

Consider a fixed geometrical volume $\mathcal{V}$ inside a continuous medium, delimited by an immaterial surface $\partial\mathcal{V}$. Let $G(t)$ be the amount of a given extensive thermodynamic quantity within this volume. For the sake of simplicity, we consider here only the case of a scalar quantity. The corresponding density and amount per unit mass are respectively denoted by $g(t,\vec{r})$ and $g_{\mathrm{m}}(t,\vec{r})$.

Introducing the local mass density $\rho(t,\vec{r})$, one has [see Eq. (I.10)]

$$G(t) = \int_{\mathcal{V}} g(t,\vec{r})\,\mathrm{d}^3\vec{r} = \int_{\mathcal{V}} \rho(t,\vec{r})\,g_{\mathrm{m}}(t,\vec{r})\,\mathrm{d}^3\vec{r}. \tag{I.16}$$

At each point $\vec{r}$ of the surface $\partial\mathcal{V}$, the amount of quantity G flowing through an infinitesimal surface element $\mathrm{d}^2\mathcal{S}$ in the time interval $[t, t+\mathrm{d}t]$ is given by

$$\vec{\mathcal{J}}_G(t,\vec{r})\cdot\vec{\mathrm{e}}_n(\vec{r})\,\mathrm{d}^2\mathcal{S}\,\mathrm{d}t,$$

where $\vec{\mathrm{e}}_n(\vec{r})$ denotes the unit normal vector to the surface, oriented towards the exterior of $\mathcal{V}$, while $\vec{\mathcal{J}}_G(t,\vec{r})$ is the *current density* or *flux density* (often referred to as *flux*) of G.

The integral *balance equation* for G reads

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t} + \int_{\partial\mathcal{V}} \vec{\mathcal{J}}_G(t,\vec{r})\cdot\vec{\mathrm{e}}_n(\vec{r})\,\mathrm{d}^2\mathcal{S} = \int_{\mathcal{V}} \sigma_G(t,\vec{r})\,\mathrm{d}^3\vec{r}. \tag{I.17}$$

$\sigma_G(t,\vec{r})$ is a *source density*,(6) which describes the rate at which the quantity G is created per unit volume—in the case of a conserved quantity, the corresponding source density σ_G vanishes. In words, Eq. (I.17) states that the net rate of change of G inside volume $\mathcal{V}$ and the flux of G exiting through the surface $\partial\mathcal{V}$ per unit time add up to the amount of G created per unit time in the volume.

In the first term of the balance equation, $G(t)$ can be replaced by a volume integral using Eq. (I.16), and the time derivation and volume integration can then be exchanged. In turn, the second term on the right-hand side of Eq (I.17) can be transformed with the help of the divergence theorem, leading to

$$\int_{\mathcal{V}} \frac{\partial g(t,\vec{r})}{\partial t}\,\mathrm{d}^3\vec{r} + \int_{\mathcal{V}} \vec{\nabla}\cdot\vec{\mathcal{J}}_G(t,\vec{r})\,\mathrm{d}^3\vec{r} = \int_{\mathcal{V}} \sigma_G(t,\vec{r})\,\mathrm{d}^3\vec{r}.$$

Since the equality should hold for arbitrary volume $\mathcal{V}$, one obtains the *local balance equation*

$$\boxed{\frac{\partial g(t,\vec{r})}{\partial t} + \vec{\nabla}\cdot\vec{\mathcal{J}}_G(t,\vec{r}) = \sigma_G(t,\vec{r}),} \tag{I.18a}$$

or equivalently

$$\frac{\partial}{\partial t}\big[\rho(t,\vec{r})g_{\mathrm{m}}(t,\vec{r})\big] + \vec{\nabla}\cdot\vec{\mathcal{J}}_G(t,\vec{r}) = \sigma_G(t,\vec{r}). \tag{I.18b}$$

When the source density vanishes, that is for conserved quantities G, these local balance equations reduce to so-called *continuity equations*.

I.1.4 b Entropy production

In the case of the entropy, which is not a conserved quantity, the general balance equation (I.17) becomes

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\int_{\partial\mathcal{V}} \vec{\mathcal{J}}_S(t,\vec{r})\cdot\vec{\mathrm{e}}_n(\vec{r})\,\mathrm{d}^2\mathcal{S} + \int_{\mathcal{V}} \sigma_S(t,\vec{r})\,\mathrm{d}^3\vec{r} \equiv \frac{\mathrm{d}S_{\text{ext.}}(t)}{\mathrm{d}t} + \frac{\mathrm{d}S_{\text{int.}}(t)}{\mathrm{d}t}, \tag{I.19}$$

with $\vec{\mathcal{J}}_S$ the entropy flux density and σ_S the entropy source density, which is necessarily nonnegative. The first term in the right member of the equation arises from the exchanges with the exterior of the volume $\mathcal{V}$ under consideration. If $\mathcal{V}$ corresponds to the whole volume of an isolated system, then this term vanishes, since by definition the system does not exchange anything with its environment.

---

(6) ... or "sink density", in case the quantity G is destroyed.

The second term on the right-hand side of the balance equation (I.19) corresponds to the creation of entropy due to internal changes in the bulk of $\mathcal{V}$, and can be non-vanishing even for isolated systems. This contribution is called *entropy production rate*, or often more briefly *entropy production* or even *dissipation*.

The corresponding local balance equation for entropy reads [cf. Eq. (I.18a)]

$$\frac{\partial s(t,\vec{r})}{\partial t} + \vec{\nabla}\cdot\vec{\mathcal{J}}_S(t,\vec{r}) = \sigma_S(t,\vec{r}). \tag{I.20}$$

This equation will now be exploited to define, in analogy with the case of a discrete system of Sec. I.1.2, the affinities conjugate to the extensive thermodynamic variables.

I.1.4 c Affinities and fluxes

The local equilibrium assumption, according to which the local entropy $S(t,\vec{r})$ has the same functional dependence on the local thermodynamic variables $\mathcal{X}_a(t,\vec{r})$ as given by Gibbs' fundamental equation in equilibrium, leads on the one hand to Eq. (I.11), from which follows

$$\frac{\partial s(t,\vec{r})}{\partial t} = {\sum_a}' \mathcal{Y}_a(t,\vec{r})\,\frac{\partial \chi_a(t,\vec{r})}{\partial t}. \tag{I.21}$$

On the other hand, the same hypothesis suggests for the entropy flux density $\vec{\mathcal{J}}_S$ the expression

$$\vec{\mathcal{J}}_S(t,\vec{r}) = {\sum_a}' \mathcal{Y}_a(t,\vec{r})\,\vec{\mathcal{J}}_a(t,\vec{r}), \tag{I.22}$$

with $\vec{\mathcal{J}}_a$ the flux density for the quantity $\mathcal{X}_a$. Taking the divergence of this identity yields

$$\vec{\nabla}\cdot\vec{\mathcal{J}}_S(t,\vec{r}) = {\sum_a}' \Big[\vec{\nabla}\mathcal{Y}_a(t,\vec{r})\cdot\vec{\mathcal{J}}_a(t,\vec{r}) + \mathcal{Y}_a(t,\vec{r})\,\vec{\nabla}\cdot\vec{\mathcal{J}}_a(t,\vec{r})\Big].$$

Inserting this divergence together with the time derivative (I.21) in the local balance equation for the entropy (I.20) gives the entropy production rate

$$\sigma_S(t,\vec{r}) = {\sum_a}' \bigg\{\vec{\nabla}\mathcal{Y}_a(t,\vec{r})\cdot\vec{\mathcal{J}}_a(t,\vec{r}) + \mathcal{Y}_a(t,\vec{r})\bigg[\vec{\nabla}\cdot\vec{\mathcal{J}}_a(t,\vec{r}) + \frac{\partial \chi_a(t,\vec{r})}{\partial t}\bigg]\bigg\},$$

i.e., after taking into account the continuity equations for the conserved thermodynamic quantities,

$$\sigma_S(t,\vec{r}) = {\sum_a}' \vec{\nabla}\mathcal{Y}_a(t,\vec{r})\cdot\vec{\mathcal{J}}_a(t,\vec{r}). \tag{I.23}$$

Defining now *affinities* as

$$\boxed{\mathscr{F}_a(t,\vec{r}) \equiv \vec{\nabla}\mathcal{Y}_a(t,\vec{r})} \tag{I.24}$$

the entropy production rate (I.23) can be rewritten as

$$\boxed{\sigma_S(t,\vec{r}) = {\sum_a}' \vec{\mathscr{F}}_a(t,\vec{r})\cdot\vec{\mathcal{J}}_a(t,\vec{r}).} \tag{I.25}$$

The entropy production rate in a continuous medium thus has the same bilinear structure in the affinities and fluxes as in a discrete thermodynamic system. This remains true when one considers not only the exchange of scalar quantities (like energy or particle number), but also when exchanging vector quantities (like momentum) or when allowing for chemical reactions.

One should however note several differences:

- σ_S is the rate of entropy production per unit volume, while $\mathrm{d}S^{\mathrm{tot}}/\mathrm{d}t$ is for the whole volume of the system;

- the "fluxes" $\vec{\mathcal{J}}_a$ are actually flux densities, in contrast to the discrete fluxes $\mathcal{J}_a$, which are rates of change;

- the affinities conjugate to scalar extensive quantities in a continuous medium are the gradients of the intensive parameters, while in the discrete case they are differences.

  Since the intensive variable conjugate to a vectorial extensive parameter is itself a vector, as exemplified by Eq. (I.3) for momentum, one easily finds that the corresponding affinity is a tensor of rank 2. In that case, the flux density is also a tensor of rank 2.

  Eventually, chemical reactions in a continuous medium can be accounted for by splitting it in thought in a discrete set of continuous media, corresponding to the various chemical components. The affinities and fluxes describing the exchanges between these discrete systems then follows the discussion in Sec. I.1.2.

> The transport of scalar quantities like internal energy or particle number is thus a vectorial process (the $\vec{\mathcal{J}}_a$ are vectors), while the transport of momentum is a tensorial process (of rank 2), and chemical reactions are scalar processes.

As an example of the considerations in this paragraph, consider a chemically inert simple continuous medium in local mechanical equilibrium—i.e. $\vec{\mathsf{v}}(\vec{r}) = \vec{0}$ everywhere. The entropy production rate (I.23) reads (for the sake of brevity the dependences on time and position are omitted)

$$\sigma_S = \vec{\nabla}\left(\frac{1}{T}\right) \cdot \vec{\mathcal{J}}_E - \sum_k \vec{\nabla}\left(\frac{\mu_k}{T}\right) \cdot \vec{\mathcal{J}}_{N_k}, \tag{I.26}$$

with $\vec{\mathcal{J}}_E$ the flux density of internal energy and $\vec{\mathcal{J}}_{N_k}$ the particle flux density for species k. This entropy production rate can be rewritten using the entropy flux density, which according to formula (I.22) is given by

$$\vec{\mathcal{J}}_S = \frac{1}{T}\vec{\mathcal{J}}_E - \sum_k \frac{\mu_k}{T}\vec{\mathcal{J}}_{N_k}, \tag{I.27a}$$

so that

$$\sigma_S = -\frac{1}{T}\vec{\mathcal{J}}_S \cdot \vec{\nabla}T - \sum_k \frac{1}{T}\vec{\mathcal{J}}_{N_k} \cdot \vec{\nabla}\mu_k. \tag{I.27b}$$

According to this expression, the affinities conjugate to the flux densities $\vec{\mathcal{J}}_S$ and $\vec{\mathcal{J}}_{N_k}$—which are the "natural" fluxes in the *energy representation*, where the variables are S and the $\{N_k\}$, rather than U and the $\{N_k\}$ in the entropy representation—are respectively $-(1/T)\vec{\nabla}T$ and $-(1/T)\vec{\nabla}\mu_k$.

I.2 Linear irreversible thermodynamic processes

The affinities and fluxes introduced in the previous section to describe out-of-equilibrium thermodynamic systems remain useless as long as they are not supplemented with relations that specify how the fluxes are related to the other thermodynamic parameters. In the framework of thermodynamics, these are phenomenological laws, involving coefficients, characteristic of each system, which have to be taken from experimental measurements.

In Sec. I.2.1, we introduce a few physical assumptions that lead to simplifications of the functional form of these relations. The various coefficients entering the laws cannot be totally arbitrary, but are restricted by symmetry considerations as well as by relations, due to Lars Onsager,(d) which within a macroscopic approach can be considered as an additional fundamental principle (Sec. I.2.2). Several long known phenomenological laws describing the transport of various quantities are presented and recast within the general framework of irreversible thermodynamics (Secs. I.2.3 & I.2.4).

(d) L. Onsager, 1903–1976

I.2.1 Linear processes in Markovian thermodynamic systems

For a given thermodynamic system, the various local intensive parameters $\mathcal{Y}_a$, affinities $\mathscr{F}_a$, and fluxes $\mathcal{J}_a$—where for the sake of brevity the tensorial nature of the quantities has been omitted—represent a set of variables that are not fully constrained by the assumption of local thermodynamic equilibrium, that is through the knowledge of the local equations of state alone. To close the system of equations for these variables, one need further relations, and more precisely between the fluxes and the other parameters.

Remark: As implied here, the customary approach is to use the parameters $\{\mathcal{Y}_a\}$ instead of the corresponding conjugate extensive variables $\{\mathcal{X}_a\}$ (resp. the densities $\{x_a\}$ in continuous systems). Both choices are however equivalent. Again, the intensive parameter conjugate to volume $\mathcal{Y}_{\mathcal{V}}$ drops out from the list of relevant parameters.

Most generally, a given flux $\mathcal{J}_a(t,\vec{r})$ might conceivably depend on the values of the intensive parameters $\mathcal{Y}_b$ and affinities $\mathscr{F}_b$ at every instant and position allowed by causality, i.e. any time $t' \leq t$ and position $\vec{r}\,'$ satisfying $|\vec{r}-\vec{r}\,'| \leq c(t-t')$, with c the velocity of light in vacuum.

In many systems, one can however assume that the fluxes at a given time only depend on the values of the parameters $\{\mathcal{Y}_b\}$, $\{\mathscr{F}_b\}$ at the same instant—that is, automatically, at the same point. For these memoryless, "Markovian"(7) systems, one thus has

$$\mathcal{J}_a(t,\vec{r}) = \mathcal{J}_a\big(\{\mathscr{F}_b(t,\vec{r})\},\{\mathcal{Y}_b(t,\vec{r})\}\big). \tag{I.28}$$

In the remainder of this section, we shall drop the t and $\vec{r}$ dependence of the various fields.

Remark: The assumption of instantaneous relationship between cause and effect automatically leaves aside hysteresis phenomena, in which the past history of the system plays an essential role, as for instance in ferromagnets.

Viewing the flux as a function of the affinities, a Taylor expansion gives

$$\mathcal{J}_a = \mathcal{J}_a^{\text{eq.}} + \sum_b{}' L_{ab}\mathscr{F}_b + \frac{1}{2!}\sum_{b,c}{}' L_{abc}\mathscr{F}_b\mathscr{F}_c + \ldots, \tag{I.29a}$$

where the *kinetic coefficients* L_{ab}, L_{abc}, … are functions of the intensive parameters

$$L_{ab} = L_{ab}\big(\{\mathcal{Y}_d\}\big), \qquad L_{abc} = L_{abc}\big(\{\mathcal{Y}_d\}\big), \ldots \tag{I.29b}$$

The expansion (I.29a) also includes an equilibrium current $\mathcal{J}_a^{\text{eq.}}$, which however does not contribute to entropy production, to account for the possible motion of the system with respect to the reference frame in which it is studied. In the presence of such a current, the relation between the entropy production rate and the affinities and fluxes becomes

$$\sigma_S = \sum_a{}' \mathscr{F}_a\big(\mathcal{J}_a - \mathcal{J}_a^{\text{eq.}}\big) \tag{I.30}$$

instead of Eq. (I.25).

If the affinities and fluxes are vectors or more generally tensors of rank 2 or above, the kinetic coefficients are themselves tensors. For instance, in the case of vectorial transport, the first order coefficients are tensors L_{ab} of rank 2, with components L_{ab}^{ij} where i, $j = 1,2,3$.

When the affinities are small, one may approximate the flux (I.29a) by the constant and first order terms in the expansion only, while the higher order terms can be neglected. Such a *linear process* thus obeys the general relationship

$$\mathcal{J}_a = \mathcal{J}_a^{\text{eq.}} + \sum_b{}' L_{ab}\mathscr{F}_b. \tag{I.31}$$

---

(7) A. A. Markov, 1859–1922

Remarks:

$*$ The kinetic coefficients L_{ab}, as well as the various related transport coefficients (κ, D, $\sigma_{\text{el.}}$, ϵ_S, Π, η, $\zeta\dots$) introduced in Secs. I.2.3–I.2.4 below, are conventionally defined for *stationary* flux densities. As we shall see in chapter VI, these coefficients are in fact the low-frequency, long-wavelength limits of respective response functions relating time- and position-dependent affinities and fluxes.

$*$ The relations (I.31)—or more generally (I.29a)—between fluxes and affinities are sometimes called *constitutive equations*.

$*$ Restricting the discussion to that of linear processes, as we shall from now on do, amounts to restricting the class of out-of-equilibrium states under consideration: among the vast number of possible macroscopic states of a system, we actually only consider those that are relatively close to equilibrium, i.e. in which the affinities are "small". This smallness of the gradients $\vec{\nabla}\mathcal{Y}_a$ means that the typical associated length scale $(|\vec{\nabla}\mathcal{Y}_a|/\mathcal{Y}_a)^{-1}$ should be "large" compared to the size of the mesoscopic scale on which the medium can be subdivided into small cells.

In the case of linear processes, the rate of entropy production (I.30) becomes

$$\sigma_S = \sum_{a,b}{}' L_{ab}\mathscr{F}_a\mathscr{F}_b. \qquad \text{(I.32)}$$

Since the product $\mathscr{F}_a\mathscr{F}_b$ is symmetric in the exchange of quantities a and b, only the symmetric part $\frac{1}{2}(L_{ab}+L_{ba})$ contributes to the entropy production (I.32), while the antisymmetric part does not contribute.

The requirement that $\sigma_S \geq 0$ implies $L_{aa} \geq 0$ for every a, as well as $L_{aa}L_{bb} - \frac{1}{4}(L_{ab}+L_{ba})^2 \geq 0$ for every a and b.(8)

Specific case of discrete systems

In a discrete system, the fluxes are the rates of change of the basic extensive quantities $\{\mathcal{X}_a(t)\}$, see Eq. (I.8). By working in the system rest frame, one can ensure the absence of equilibrium fluxes.

Instead of the parameters $\{\mathcal{X}_a(t)\}$, let us consider their departures $\{\Delta\mathcal{X}_a(t)\}$ from their respective equilibrium values, $\Delta\mathcal{X}_a(t) \equiv \mathcal{X}_a(t) - \mathcal{X}_a^{\text{eq.}}$. Obviously the flux $\mathcal{J}_a(t)$ is also the rate of change of $\Delta\mathcal{X}_a(t)$:

$$\mathcal{J}_a(t) = \frac{\mathrm{d}\Delta\mathcal{X}_a(t)}{\mathrm{d}t}.$$

The entropy $S(t)$ is a function of the variables $\{\mathcal{X}_a(t)\}$, or equivalently of the $\{\Delta\mathcal{X}_a(t)\}$. In turn, each affinity $\mathscr{F}_b(t)$, which is a derivative of the entropy, is a function of the $\{\Delta\mathcal{X}_a(t)\}$. For small departures from equilibrium, i.e. small values of the $\{\Delta\mathcal{X}_a(t)\}$, this dependence can be linearized:(9)

$$\mathscr{F}_b = -\sum_c \beta_{bc}\Delta\mathcal{X}_c.$$

Defining then $\lambda_{ac} \equiv \sum_b L_{ab}\beta_{bc}$, and using the expressions for the flux $\mathcal{J}_a$ resp. the affinities $\mathscr{F}_b$ as given by the previous two equations, the constitutive linear relation (I.31) becomes

$$\frac{\mathrm{d}\Delta\mathcal{X}_a(t)}{\mathrm{d}t} = -\sum_c \lambda_{ac}\Delta\mathcal{X}_c(t). \qquad \text{(I.33)}$$

That is, we find coupled first-order differential equations for the departures from equilibrium $\Delta\mathcal{X}_a(t)$. These equations should describe the relaxation of each individual $\Delta\mathcal{X}_a(t)$ to 0 at equilibrium—which amounts to the relaxation of $\mathcal{X}_a(t)$ to its equilibrium value $\mathcal{X}_a^{\text{eq.}}$: the eigenvalues of the matrix with coefficients λ_{ac} should thus all be positive.

---

(8) More generally, every minor of the symmetric matrix with elements $\frac{1}{2}(L_{ab}+L_{ba})$ is nonnegative.

(9) We denote the coefficients as $-\beta_{bc}$ to parallel the notation in Landau & Lifshitz [3, § 120], who use the opposite sign convention for affinities and fluxes as adopted in these notes, see the remark following Eq. (I.8).

I.2.2 Curie principle and Onsager relations

In the relation (I.31) [or more generally Eq. (I.29a)] between flux and affinities, it is assumed that a given flux $\mathcal{J}_a$ depends not only on the conjugate affinity $\mathscr{F}_a$, but also on the other affinities $\mathscr{F}_b$ with $b \neq a$. We now discuss general principles that restrict the possible values of kinetic coefficients L_{ab} (and more generally $L_{abc\ldots}$), that go beyond the already mentioned positivity of the symmetric matrix with elements $\frac{1}{2}(L_{ab}+L_{ba})$.

I.2.2 a Curie symmetry principle

A first principle is that, going back to Pierre Curie(e) (1894), according to which the effects—here, the fluxes—should have the same symmetry elements as their causes—here, the affinities.

Remark: Strictly speaking, this principle holds when considering *all* possible effects of a given cause, i.e. when all realizations of some possible spontaneous symmetry breaking—which does not occur here—are taken into account.

Restricting ourselves to locally isotropic continuous media, which are symmetric under arbitrary space rotations and under space parity, two consequences of this principle can be listed:

- In the transport of scalar quantities, for which fluxes and affinities are vectors, the tensors $\mathbf{L}_{ab}$ are actually proportional to the identity, i.e. involve a single number: $\mathbf{L}_{ab} = L_{ab}\mathbb{1}_3$, with $\mathbb{1}_3$ the unit rank-two tensor on three-dimensional space; in terms of (Cartesian) components $L^{ij}_{ab} = L_{ab}\,\delta^{ij}$, with δ^{ij} the Kronecker symbol.
- Fluxes and affinities whose tensorial ranks are of different parities cannot be coupled together. Such a pair, for instance a vector (rank 1) and a tensor of rank 2, would involve a tensorial kinetic coefficient of odd rank, in the example of rank 1 or 3, which does not stay invariant under rotations or space parity.

I.2.2 b Onsager reciprocal relations

Another symmetry principle—which was first found experimentally in various systems and then formalized in 1931 by Lars Onsager [10, 11] within statistical mechanics—regards the cross-coefficients L_{ab} with $a \neq b$.(10)

The latter describe “indirect” transport, as e.g. when energy is transported not only because of a temperature gradient [or more accurately, a non-vanishing $\mathscr{F}_E = \vec{\nabla}(1/T)$]—which amounts to transfer through *conduction*—, but also due to a gradient in particle density (within the formalism, a gradient in $\mathcal{Y}_N = -\mu/T$) or in velocity—which is energy transfer due to *convection*.

In the simplest case where both extensive quantities $\mathcal{X}_a$ and $\mathcal{X}_b$ behave similarly under time reversal—as is for instance the case of internal energy U and any particle number N, which all remain unchanged when t is changed to $-t$ in the equations of motion—then the associated cross-coefficients are equal

$$\boxed{L_{ab} = L_{ba}.} \tag{I.34}$$

Thus when a gradient in $\mathcal{Y}_b$ causes a change in $\mathcal{X}_a$, then a gradient in $\mathcal{Y}_a$ induces a change in $\mathcal{X}_b$ of the same relative size.

These relations were generalized by Casimir(f) [13] to relate the kinetic coefficients for thermodynamic parameters that behave differently under time reversal. Let $\epsilon_a = \pm 1$ denote the parity (or *signature*) of $\mathcal{X}_a$, or equivalently the density χ_a, under the substitution $t \to -t$. Internal energy U, particle numbers N_k, position $\vec{r}$ have parity $+1$, while momentum $\vec{P}$ or velocity $\vec{v}$ have parity -1.

---

(10) A review of (older) experimental results supporting the Onsager reciprocal relations can be found in Ref. [12].

(e) P. Curie, 1859–1906 (f) H. Casimir, 1909–2000

Under consideration of these signatures, the Onsager–Casimir relations in the absence of external magnetic field and of global rotation read

$$\boxed{L_{ab} = \epsilon_a \epsilon_b L_{ba}.} \tag{I.35}$$

The relations express the symmetry or antisymmetry of the kinetic coefficients.

Taking now into account the possible presence of an external magnetic field $\vec{\mathscr{B}}$ and/or of a global rotation of the system with angular velocity $\vec{\Omega}$, the generalized Onsager–Casimir relations become

$$\boxed{L_{ab}(\vec{\mathscr{B}}, \vec{\Omega}) = \epsilon_a \epsilon_b L_{ba}(-\vec{\mathscr{B}}, -\vec{\Omega}).} \tag{I.36}$$

Note that the latter relations actually relate *different* systems, with opposite values of the parameters $\vec{\mathscr{B}}$, $\vec{\Omega}$.

Remarks:

∗ The Onsager(–Casimir) relations are sometimes considered as the "4th law of thermodynamics", which supplements the three classical laws recalled at the end of Sec. I.1.1.

∗ The Onsager relations will be derived from general principles in chapter VI.

I.2.3 First examples of linear transport phenomena

Following the example set by Onsager in his original articles [10, 11], we now enumerate several phenomenological linear transport laws formulated in the 19th century and re-express them in terms of relations between fluxes and affinities as formalised in section I.2.1.

We shall begin with a few "direct" transport phenomena—for heat, particle number, or electric charges. Next, we turn to a case in which indirect transport plays a role, namely that of thermoelectric effects. These first examples will be studied in the respective rest frames of the systems under study, so that the equilibrium fluxes $\mathcal{J}_j^{\text{eq.}}$ will vanish. Eventually, we describe the various transport phenomena in a simple fluid, which will allow us to derive the classical laws of hydrodynamics.

In most of this section, we shall for the sake of brevity drop the $(t, \vec{r})$-dependence of the various physical quantities under consideration.

I.2.3 a Heat transport

In an insulating solid with a temperature gradient, heat is transported through the vibrations of the underlying crystalline structure—whose quantum mechanical description relies on phonons—rather than through particle transport.

Traditionally, this transport of energy is expressed in the form of *Fourier's*[(g)] *law* (1822)

$$\boxed{\vec{\mathcal{J}}_E = -\kappa\, \vec{\nabla} T,} \tag{I.37a}$$

with κ the *heat conductivity* of the insulator.

Using the general formalism of linear irreversible thermodynamic processes, the relationship between the energy flux density and the conjugate affinity, in the case when there is no gradient of the ratio μ/T,(11) reads in the linear regime

$$\vec{\mathcal{J}}_E = \mathbf{L}_{EE} \cdot \vec{\nabla}\left(\frac{1}{T}\right), \tag{I.37b}$$

with $\mathbf{L}_{EE}$ a tensor of rank 2 of (first-order) kinetic coefficients. If the insulating medium under

---

(11) Phonons are massless and carry no conserved quantum number, so that their chemical potential vanishes everywhere.

(g) J. Fourier, 1768–1830

consideration is isotropic,(12) this tensor is for symmetry reasons proportional to the identity

$$\mathbf{L}_{EE} = L_{EE}\,\mathbb{1}_3.$$

The comparison between Eqs. (I.37a) and (I.37b) then gives the identification

$$\kappa = \frac{1}{T^2}\,L_{EE}. \tag{I.37c}$$

Since $L_{EE} \geq 0$ to ensure the positivity of the entropy production rate, κ is also nonnegative. The flux (I.37a) thus transports energy from the regions of higher temperatures to the colder ones.

Combining Fourier's law (I.37a) with the continuity equation (I.18a) applied to the energy density e yields

$$\frac{\partial e}{\partial t} = -\vec{\nabla}\cdot\vec{\mathcal{J}}_E = \vec{\nabla}\cdot\big(\kappa\,\vec{\nabla}T\big)$$

Assuming that the heat conductivity is uniform in the medium under study, κ can be taken out of the divergence, so that the right-hand side becomes $\kappa\triangle T$, with $\triangle$ the Laplacian. According to a well known thermodynamic relation, at fixed volume the change in the internal energy equals the heat capacity at constant volume multiplied by the change in the temperature, which results in $\mathrm{d}e = c_{\mathcal{V}}\,\mathrm{d}T$ with $c_{\mathcal{V}}$ the heat capacity per unit volume. If the latter is independent of temperature, one readily obtains the evolution equation

$$\frac{\partial T}{\partial t} = \frac{\kappa}{c_{\mathcal{V}}}\,\triangle T. \tag{I.38}$$

This is the generic form of a diffusion equation [see Eq. (I.40) below], with diffusion coefficient $\kappa/c_{\mathcal{V}}$.

I.2.3 b Particle diffusion

Consider now "particles" immersed in a motionless and homogeneous medium, in which they can move around—microscopically, through scatterings on the medium constituents—without affecting the medium characteristics. Examples are the motion of dust in the air, of micrometer-scale bodies in liquids, but also of impurities in a solid or of neutrons in the core of a nuclear reactor.

Let n denote the number density of the particles. The transport of particles can be described by *Fick's*[h] *law* (1855) [14]

$$\boxed{\vec{\mathcal{J}}_N = -D\,\vec{\nabla}n,} \tag{I.39a}$$

with $\vec{\mathcal{J}}_N$ the flux density of particle number and D the *diffusion coefficient*.

Remark: Relation (I.39a) is sometimes referred to as Fick's first law, the second one being actually the diffusion equation (I.40).

In the absence of temperature gradient and of collective motion of the medium, the general relation (I.31) yields for the particle number flux density

$$\vec{\mathcal{J}}_N = L_{NN}\vec{\nabla}\bigg(-\frac{\mu}{T}\bigg) \tag{I.39b}$$

with $L_{NN} \geq 0$. Relating the differential of chemical potential to that of number density with

$$\mathrm{d}\mu = \bigg(\frac{\partial\mu}{\partial n}\bigg)_T \mathrm{d}n,$$

---

[12] ...which is strictly speaking never the case at the microscopic level in a crystal, since the lattice structure is incompatible with local invariance under the whole set of three-dimensional rotations. Nevertheless, for lattices with a cubic elementary mesh, isotropy holds, yet at the mesoscopic level.

[h] A. Fick, 1829–1901

the identification of Fick's law (I.39a) with formula (I.39b) yields

$$D = \frac{1}{T}\left(\frac{\partial\mu}{\partial n}\right)_T L_{NN}, \tag{I.39c}$$

where the precise form of the partial derivative depends on the system under study.

Diffusion equation

If the number of diffusing particles is conserved—which is for instance not the case for neutrons in a nuclear reactor(13)—and if the diffusion coefficient is independent of position, the associated continuity equation (I.18a) leads to the *diffusion equation*

$$\boxed{\frac{\partial n(t,\vec{r})}{\partial t} = D\,\triangle n(t,\vec{r}).} \tag{I.40}$$

To tackle this partial differential equation (considered on $\mathbb{R}^3$), one can introduce the Fourier transform with respect to space coordinates

$$\tilde{n}(t,\vec{k}) \equiv \int n(t,\vec{r})\,\mathrm{e}^{-\mathrm{i}\vec{k}\cdot\vec{r}}\,\mathrm{d}^3\vec{r}$$

of the number density. This transform then satisfies for each $\vec{k}$ the ordinary differential equation

$$\frac{\partial\tilde{n}(t,\vec{k})}{\partial t} = -D\,\vec{k}^{\,2}\tilde{n}(t,\vec{k}),$$

where we assumed that the number density and its spatial derivatives vanish at infinity at every instant. These assumptions respectively guarantee the finiteness of the overall particle number and the absence of particle flux at infinity.

The solution to the differential equation reads $\tilde{n}(t,\vec{k}) = \mathrm{e}^{-D\vec{k}^2t}\,\tilde{n}(0,\vec{k})$, with $\tilde{n}(0,\vec{k})$ the initial condition at $t = 0$ in Fourier space. An inverse Fourier transform then yields

$$n(t,\vec{r}) = \int \mathrm{e}^{-D\vec{k}^2t}\,\tilde{n}(0,\vec{k})\,\mathrm{e}^{\mathrm{i}\vec{k}\cdot\vec{r}}\,\frac{\mathrm{d}^3\vec{k}}{(2\pi)^3}.$$

If the initial condition is $n(0,\vec{r}) = n_0\,\delta^{(3)}(\vec{r})$—which physically amounts to introducing a particle density n_0 at the point $\vec{r} = \vec{0}$ at time $t = 0$—then the Fourier transform is trivially $\tilde{n}(0,\vec{k}) = n_0$, so that the inverse Fourier transform above is simply that of a Gaussian, which gives

$$n(t,\vec{r}) = \frac{n_0}{(4\pi Dt)^{3/2}}\,\mathrm{e}^{-\vec{r}^{\,2}/4Dt}.$$

The typical width of the particle number density increases with $\sqrt{t}$.

I.2.3 c Electrical conduction

Another example of particle transport is that of the moving charges in an electrical conductor in the presence of an electric field $\vec{\mathscr{E}} = -\vec{\nabla}\Phi$, with Φ the electrostatic potential. The latter is assumed to vary very slowly at the mesoscopic scale, so as not to spoil the local equilibrium assumption. If q denotes the electric charge of the carriers, then the electric charge flux density, traditionally referred to as *current density*, is simply related to the number flux density of the moving charges through

$$\vec{J}_{\text{el.}} = q\vec{\mathcal{J}}_N. \tag{I.41a}$$

---

(13) There, one should also include various source and loss terms, to account for the production of neutrons through fission reactions, or their "destruction" through reactions with the nuclear fuel, with the nuclear waste present in the reactor, or with the absorber bars that moderate the chain reaction, or their natural decay.

The relation between electric field and current density in a microscopically isotropic conductor at constant temperature in the absence of magnetic field is (the microscopic version of) *Ohm's*(i) *law*

$$\boxed{\vec{J}_{\rm el.} = \sigma_{\rm el.}\vec{\mathscr{E}}} \tag{I.41b}$$

with $\sigma_{\rm el.}$ the (isothermal) *electrical conductivity*.

To relate the electrical conductivity to the kinetic coefficients of Sec. I.2.1, and more specifically to L_{NN} since $\vec{J}_{\rm el.}$ is proportional to $\vec{\mathcal{J}}_N$, one needs to determine the intensive variable conjugate to particle number—or equivalently, thanks to the local equilibrium assumption, conjugate to particle number density. Now, if e denotes the (internal) energy density in the absence of electrostatic potential, then the energy density in presence of Φ becomes $e + nq\Phi$: meanwhile, the particle number density n remains unchanged. The entropy per unit volume then satisfies—as can most easily be checked within the grand-canonical ensemble of statistical mechanics—the identity

$$s(e, n, \Phi) = s(e - nq\Phi, n, 0),$$

which yields

$$\frac{\partial s}{\partial n} = -\frac{\mu + q\Phi}{T} \equiv -\frac{\mu_\Phi}{T}, \tag{I.42}$$

with μ the chemical potential at vanishing electric potential. μ_Φ is referred to as *electrochemical potential*.

Assuming a uniform temperature in the conductor, the linear relation (I.31) for the flux of particle number then reads

$$\vec{\mathcal{J}}_N = L_{NN}\vec{\nabla}\bigg(-\frac{\mu + q\Phi}{T}\bigg). \tag{I.43}$$

Using again the uniformity of temperature, this gives

$$\vec{\mathcal{J}}_N = -\frac{1}{T}\bigg(\frac{\partial\mu}{\partial n}\bigg)_T L_{NN}\vec{\nabla}n - \frac{q}{T}\,L_{NN}\vec{\nabla}\Phi,$$

where the first term is the same as in Sec. I.2.3 b, while the second can be rewritten with the help of the electric field. If the density is uniform, the first term vanishes, and the identification with Eqs. (I.41a) and (I.41b) yields the electrical conductivity

$$\sigma_{\rm el.} = \frac{q^2}{T}\,L_{NN}. \tag{I.44}$$

Einstein relation

Equations (I.39c) and (I.44) show that the diffusion coefficient D and the electrical conductivity σ are both related to the same kinetic coefficient L_{NN}, so that they are related with each other:

$$D = \frac{\sigma_{\rm el.}}{q^2}\bigg(\frac{\partial\mu}{\partial n}\bigg)_T. \tag{I.45}$$

Let $\mu_{\rm el.}$ denote the *electrical mobility* of the charge carriers, which is the proportionality factor between the mean velocity $\vec{v}_{\rm av.}$ they acquire in an electric field $\vec{\mathscr{E}}$ and this field

$$\vec{v}_{\rm av.} = \mu_{\rm el.}\vec{\mathscr{E}}. \tag{I.46}$$

Obviously, the determination of $\mu_{\rm el.}$ requires a microscopic model for the motion of the charges.

At the macroscopic level, the electric current density is simply the product of the mean velocity of charges times the charge density, i.e.

$$\vec{J}_{\rm el.} = nq\vec{v}_{\rm av.} = nq\mu_{\rm el.}\vec{\mathscr{E}},$$

(i) G. S. Ohm, 1789–1854

which after identification with Ohm's law (I.41b) gives $\sigma_{\text{el.}} = nq\mu_{\text{el.}}$. Together with Eq. (I.45), one obtains

$$D = \frac{\mu_{\text{el.}}}{q}\, n\left(\frac{\partial\mu}{\partial n}\right)_T. \tag{I.47}$$

For a classical ideal gas, one has $\left(\frac{\partial\mu}{\partial n}\right)_T = \frac{k_BT}{n}$, which gives

$$D = \frac{\mu_{\text{el.}}}{q}\, k_BT, \tag{I.48}$$

which is a special case of a general relation derived by A. Einstein(j) in his 1905 paper on Brownian motion [15].

I.2.3 d Thermoelectric effects

We now turn to a first example of systems in which several quantities can be transported at the same time, namely that of isotropic electrical conductors, in which both heat and particles—corresponding to the charge carriers—can be transferred simultaneously from one region to the other.

For the sake of simplicity, we consider a single type of moving particles, with electric charge q. Throughout the section it will be assumed that their number density is uniform, i.e. $\vec{\nabla}n$ vanishes. On the other hand, these charges are able to move collectively, resulting in an electric current density $\vec{J}_{\text{el.}} = q\vec{\mathcal{J}}_N$. In the system reigns a slowly spatially varying electrostatic potential Φ, which results as seen in Sec. I.2.3 c in the replacement of the chemical potential μ by the electrochemical potential μ_Φ defined by Eq. (I.42).

In the linear regime, the transports of particles and energy are governed by the constitutive equations [Eq. (I.31)]

$$\begin{aligned} \vec{\mathcal{J}}_N &= L_{NN}\vec{\nabla}\left(-\frac{\mu_\Phi}{T}\right) + L_{NE}\vec{\nabla}\left(\frac{1}{T}\right),\\ \vec{\mathcal{J}}_E &= L_{EN}\vec{\nabla}\left(-\frac{\mu_\Phi}{T}\right) + L_{EE}\vec{\nabla}\left(\frac{1}{T}\right), \end{aligned} \tag{I.49}$$

where Curie's symmetry principle has already been accounted for, while Onsager's reciprocal relation reads $L_{NE} = L_{NE}$ since both particle number and energy are invariant under time reversal.

Instead of $\vec{\mathcal{J}}_E$, it is customary to consider the *heat flux* (density) $\vec{\mathcal{J}}_Q$ defined as

$$\vec{\mathcal{J}}_Q = T\vec{\mathcal{J}}_S = \vec{\mathcal{J}}_E - \mu_\Phi\vec{\mathcal{J}}_N, \tag{I.50}$$

where the second identity follows from Eq. (I.27a).
Inspecting the entropy production rate (I.27b)

$$\sigma_S = \vec{\mathcal{J}}_Q\cdot\vec{\nabla}\left(\frac{1}{T}\right) - \frac{1}{T}\,\vec{\mathcal{J}}_N\cdot\vec{\nabla}\mu_\Phi, \tag{I.51}$$

one finds that the affinities conjugate to $\vec{\mathcal{J}}_Q$ and $\vec{\mathcal{J}}_N$ are respectively $\vec{\nabla}(1/T)$ and $-(1/T)\vec{\nabla}\mu_\Phi$. Using these new fluxes and affinities as variables, we can introduce alternative linear relations

$$\vec{\mathcal{J}}_N = -L_{11}\frac{1}{T}\,\vec{\nabla}\mu_\Phi + L_{12}\vec{\nabla}\left(\frac{1}{T}\right), \tag{I.52a}$$

$$\vec{\mathcal{J}}_Q = -L_{21}\frac{1}{T}\,\vec{\nabla}\mu_\Phi + L_{22}\vec{\nabla}\left(\frac{1}{T}\right), \tag{I.52b}$$

(j) A. Einstein, 1879–1955

with new kinetic coefficients L_{jk}, which are related to the original ones by

$$\begin{aligned}
&L_{11} = L_{NN},\\
&L_{12} = L_{NE} - \mu_\Phi L_{NN}, \quad L_{21} = L_{EN} - \mu_\Phi L_{NN},\\
&L_{22} = L_{EE} - \mu_\Phi\big(L_{EN} + L_{NE}\big) + \mu_\Phi^2 L_{NN}.
\end{aligned} \tag{I.52c}$$

Again a reciprocal relation $L_{21} = L_{12}$ holds when $L_{EN} = L_{NE}$.

Heat conduction

Let us first investigate the transport of heat in a situation where the particle number flux vanishes, $\vec{\mathcal{J}}_N = \vec{0}$, i.e. for an open electric circuit $\vec{\mathcal{J}}_{\rm el.} = \vec{0}$. Equation (I.52a) gives

$$\vec{\nabla}\mu_\Phi = -\frac{1}{T}\frac{L_{12}}{L_{11}}\,\vec{\nabla}T. \tag{I.53}$$

Inserting this identity in the expression of the heat flux (I.52b) then yields

$$\vec{\mathcal{J}}_Q = -\frac{L_{11}L_{22} - L_{12}L_{21}}{T^2\,L_{11}}\,\vec{\nabla}T. \tag{I.54}$$

Since $L_{11} = L_{NN} \geq 0$ and $L_{11}L_{22} - L_{12}L_{21} = L_{NN}L_{EE} - L_{NE}L_{EN} \geq 0$, the ratio is a nonnegative number. Now, as the particle flux vanishes, $\vec{\mathcal{J}}_Q = \vec{\mathcal{J}}_E$. The comparison of Eq. (I.54) with Fourier's law (I.37a) allows us to interpret the prefactor of $\vec{\nabla}T$ in relation (I.53) as the *heat conductivity* κ. With the help of the relations (I.52c) it can be rewritten as

$$\kappa = \frac{L_{11}L_{22} - L_{12}L_{21}}{T^2\,L_{11}} = \frac{L_{NN}L_{EE} - L_{NE}L_{EN}}{T^2\,L_{NN}}. \tag{I.55}$$

This result differs from the expression (I.37c) of the heat conductivity in an insulator, which is not unexpected since in the case of a electric conductor, both phonons and moving charges contribute to the transport of heat.

Seebeck effect

Consider again the case of an open circuit, $\vec{\mathcal{J}}_N = \vec{0}$. In such a circuit, a temperature gradient induces a gradient of the electrochemical potential, see Eq. (I.53). This constitutes the *Seebeck*(k) *effect* (1821). The relationship is traditionally written in the form

$$\frac{1}{q}\,\vec{\nabla}\mu_\Phi = -\epsilon_S\,\vec{\nabla}T, \tag{I.56a}$$

which defines the *Seebeck coefficient* ϵ_S of the conductor.(14) Since the number density of moving charges is assumed to be uniform, $\vec{\nabla}\mu_\Phi = q\vec{\nabla}\Phi = -q\vec{\mathscr{E}}$, so that Eq. (I.56a) can be recast as

$$\vec{\mathscr{E}} = \epsilon_S\,\vec{\nabla}T, \tag{I.56b}$$

with $\vec{\mathscr{E}}$ the electric field.

Comparing Eqs. (I.53) and (I.56a), one finds at once

$$\epsilon_S = \frac{1}{qT}\frac{L_{12}}{L_{11}}. \tag{I.57}$$

The Seebeck effect is an instance of indirect transport, since its magnitude, measured by ϵ_S, is proportional to the cross-coefficient L_{12}.

To evidence the Seebeck effect, one can use a circuit consisting of two conductors A and B made of different materials, a "thermocouple", whose junctions are at different temperatures T_2 and T_3,

---

(14) This coefficient, characteristic of the conducting material, is often denoted by S, which we wanted to avoid here.

(k) T. J. Seebeck, 1770–1831

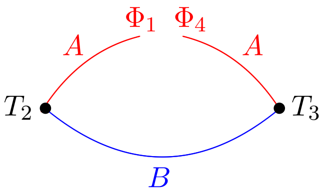

Figure I.1 – Schema of a thermocouple to evidence the Seebeck effect.

as illustrated in Fig. I.1. There appears then between the points 1 and 4 a voltage

$$\Phi_4 - \Phi_1 = \frac{1}{q}\int_1^4 \vec{\nabla}\mu_\Phi \cdot \mathrm{d}\vec{\ell} = \int_{T_2}^{T_3} \big(\epsilon_S^{(A)} - \epsilon_S^{(B)}\big)\,\mathrm{d}T.$$

This voltage can be measured with a high-resistance (so as not to close the circuit) voltmeter,(15) so that the Seebeck coefficient of one of the materials, say B, can be assessed when all other quantities (T_2, T_3, $\epsilon_S^{(A)}$) are known. Conversely, when using materials whose coefficients are known, this thermocouple allows the measurement of temperature differences.

Remark: Similar phenomena were recently discovered in magnetic materials, either conductors or insulators. Thus, in ferromagnetic materials, a so-called *spin Seebeck effect* was discovered [16], in which a temperature gradient induces a gradient in the "spin voltage" $\mu^\uparrow - \mu^\downarrow$, where $\mu^\uparrow$ resp. $\mu^\downarrow$ denotes the electrochemical potential of spin up resp. down electrons.(16)
An exact analogue of the "usual" spin-independent electric Seebeck effect is the *magnetic Seebeck effect* theorized in Ref. [19], in which a temperature gradient induces a magnetic field in a material with magnetic dipoles [20].

Peltier effect

The *Peltier*(1) *effect* (1834) consists in the fact that in a conductor at uniform temperature, a current density $\vec{J}_{\text{el.}}$ is accompanied by a flux heat. This is usually written as

$$\vec{\mathcal{J}}_Q = \Pi\,\vec{J}_{\text{el.}}, \tag{I.58}$$

which defines the *Peltier coefficient* Π of the conducting material.

Setting $\vec{\nabla}T = \vec{0}$ in Eqs. (I.52a)–(I.52b) and eliminating $(1/T)\vec{\nabla}\mu_\Phi$ between the two equations leads at once to

$$\Pi = \frac{1}{q}\frac{L_{21}}{L_{11}}. \tag{I.59}$$

This is again an indirect transport phenomenon, somehow "reverse" to the Seebeck effect since it involves the reciprocal kinetic coefficient L_{21} instead of L_{12}.

Consider the junction between two different conducting materials A and B at the same temperature depicted in Fig. I.2. An electric current $\vec{J}_{\text{el.}}$ crosses the junction without change, as dictated by local charge conservation. In each conductor, this current is accompanied by heat fluxes $\vec{\mathcal{J}}_Q^{(A)}$

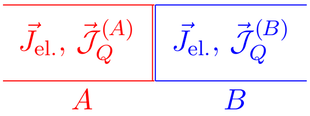

Figure I.2 – Schema of an isothermal junction to evidence the Peltier effect.

---

(15) One can easily convince oneself that the temperature of the voltmeter is irrelevant.

(16) For a review on "spin caloritronics"—the interplay of heat and spin transport—see Ref. [17]. The theory of the spin Seebeck effect is reviewed in Ref. [18].

---

(1) J. Peltier, 1785–1845

and $\vec{\mathcal{J}}_Q^{(B)}$, which differ since $\Pi^{(A)} \neq \Pi^{(B)}$. To ensure energy conservation, a measurable amount of heat

$$\frac{\mathrm{d}Q}{\mathrm{d}t}\bigg|_{\text{Peltier}} = \left(\Pi^{(A)} - \Pi^{(B)}\right) J_{\text{el.}}$$

is released per unit time and unit cross-sectional area at the junction, with $J_{\text{el.}} \equiv \left|\vec{J}_{\text{el.}}\right|$. Note that $\mathrm{d}Q/\mathrm{d}t$ can be negative, which means that heat is actually absorbed from the environment at the junction.

Comparing now the Seebeck coefficient (I.57) with the Peltier coefficient (I.59), one sees that the relation $L_{21} = L_{12}$—which follows from Onsager's reciprocal relation between L_{EN} and L_{NE}—leads to

$$\Pi = \epsilon_S T, \tag{I.60}$$

which is known as *second Kelvin*(m) *relation* (or sometimes *second Thomson relation*(17)).

Coming back to the fluxes (I.52a)–(I.52b) in the most general case of non-vanishing gradients in temperature as well as in electrochemical potential, and eliminating $(1/T)\vec{\nabla}\mu_\Phi$ between the two relations, one find with the help of Eqs. (I.55), (I.59) and (I.60) the heat flux

$$\vec{\mathcal{J}}_Q = -\kappa\,\vec{\nabla}T + \epsilon_S\,T\,q\,\vec{\mathcal{J}}_N. \tag{I.61}$$

The first term corresponds to thermal conduction, the second to convection.

I.2.4 Linear transport phenomena in simple fluids

As last example of application of the formalism of linear Markovian thermodynamic processes, let us investigate transport phenomena in isotropic "simple" non-relativistic fluids, i.e. fluids made of a single electrically neutral constituent, whose only species of (spherically symmetric) particles have a mass m.

In such a system, the subsystems that coincide with the cells at the level of which local thermodynamic equilibrium and local extensive quantities are defined—the so-called *fluid particles*—can move with respect to each other. The local momenta $\vec{P}(t,\vec{r})$ of the various cells thus differ from each other, so that there is no global rest frame in which every local momentum would vanish. This constitutes a new feature compared to the previous examples, and will require our determining the proper variables, as defined in some global reference frame, for the description of the system, before we can apply the generic ideas of linear Markovian processes.

Hereafter, the dependence of fields on time t and position $\vec{r}$ will generally not be written. We shall use Cartesian coordinates labeled by indices i, $j\dots$ running from 1 to 3, whose position will have no meaning. The components of $\vec{r}$ will be denoted as x_i.

I.2.4 a Extensive parameters, intensive quantities and fluxes

Each fluid cell is characterized by a set of local extensive variables, namely its total energy, particle number, and momentum. Both energy and momentum clearly depend on the reference frame. On the one hand, we shall consider a fixed inertial frame $\mathcal{R}_0$—corresponding to the frame in which the observer who describes the fluid is at rest. In that frame, energy and momentum will be denoted by E and $\vec{P}$.

Alternatively, we shall also use an inertial frame $\mathcal{R}_{\vec{v}}$ which at time t moves with respect to $\mathcal{R}_0$ with the same velocity $\vec{v} = \vec{v}(t,\vec{r})$ as the fluid particle located at position $\vec{r}$. In that *comoving* frame, the fluid particle is momentarily at rest, so that $\mathcal{R}_{\vec{v}}$ will be referred to as *comoving (local) rest frame*,

---

(17)... which is historically more accurate, since William Thomson had not yet been ennobled as Lord Kelvin when he empirically found this relation in 1854.

(m)W. Thomson, Lord Kelvin, 1824–1907

where "local" conveniently emphasizes that at a given instant, the velocity takes different values at different points, resulting in the existence of different comoving rest frames. In $\mathcal{R}_{\vec{\mathsf{v}}}$, the energy of the fluid cell reduces to its internal energy U while its momentum vanishes. Our first task is to find what are the conjugate intensive variables and fluxes in the fixed frame $\mathcal{R}_0$.

Let $M(t,\vec{r}) \equiv N(t,\vec{r})m$ denotes the mass of fluid contained in the cell at position $\vec{r}$ at time t. $E(t,\vec{r})$ is then simply equal to the sum of the internal energy $U(t,\vec{r})$ and the kinetic energy $\vec{P}(t,\vec{r})^2/2M(t,\vec{r})$ of the fluid particle. Thus, the characteristic extensive parameters of a cell in the fixed frame $\mathcal{R}_0$ read

$$E = U + \frac{\vec{P}^2}{2M}, \quad N, \quad P_i,\ i \in \{1,2,3\}, \tag{I.62a}$$

with respective densities (amount per unit volume)

$$e + \frac{1}{2}\rho\,\vec{\mathsf{v}}^2, \quad n, \quad \rho\,\mathsf{v}_i, \tag{I.62b}$$

where $\rho(t,\vec{r}) = mn(t,\vec{r})$ is the mass density of the fluid.

Writing that the entropy does not depend on the choice of the inertial frame in which it is measured and equating its values in $\mathcal{R}_0$ and $\mathcal{R}_{\vec{\mathsf{v}}}$, one finds the intensive variables conjugate to the extensive parameters (I.62), namely (see the derivation of $\mathcal{Y}_{P_i}$ in Sec. I.1.3 a)

$$\mathcal{Y}_E = \frac{1}{T}, \quad \mathcal{Y}_N = -\frac{\mu_{\vec{\mathsf{v}}}}{T} = -\frac{\mu + \frac{1}{2}m\,\vec{\mathsf{v}}^2}{T}, \quad \mathcal{Y}_{P_i} = -\frac{\mathsf{v}_i}{T} \tag{I.63}$$

respectively, where μ denotes the chemical potential in the comoving rest frame. Note that the temperature does not depend on the reference frame.

The traditionally adopted variables in non-relativistic fluid dynamics are the mass density ρ (instead of n), the *flow velocity* $\vec{\mathsf{v}}$ (instead of the momentum density) and the temperature T or equivalently the pressure $\mathcal{P}$ (instead of the energy density). The latter is given by the Gibbs–Duhem(n) relation

$$\mathcal{P} = Ts - e + \mu n. \tag{I.64}$$

Since the right-hand side also reads $Ts - \left(e + \frac{1}{2}\rho\,\vec{\mathsf{v}}^2\right) + \left(\mu + \frac{1}{2}m\,\vec{\mathsf{v}}^2\right)n$, the pressure $\mathcal{P}$ keeps the same value in every inertial frame.

Besides the densities (I.62b) and intensive variables (I.63), whose respective gradients are the various affinities, we still have to introduce the flux densities of the extensive parameters. Following the generic expansion (I.29a), these fluxes generally consist of a contribution depending on the affinities, which describes the response of the system to a departure from equilibrium, and of affinity-independent equilibrium fluxes, which we shall now determine.

For that purpose, it is convenient to first establish the form of the equilibrium fluxes in the comoving local rest frame $\mathcal{R}_{\vec{\mathsf{v}}}$ before performing a Galilean transformation with velocity $-\vec{\mathsf{v}}$ to obtain the expressions in the fixed reference frame $\mathcal{R}_0$.

In an equilibrated isotropic fluid at rest, invariance under rotations implies that the vectorial fluxes of internal energy and of particle number should vanish. Moreover, the momentum flux density, which is a tensor of rank 2, must be proportional to the identity tensor, again to fulfill rotational symmetry, as argued in § I.2.2 a. To interpret the proportionality coefficient, one should realize that the flux of the i-th component of linear momentum through a surface perpendicular to the i-axis represents a normal force on that surface. In mechanical equilibrium, this force is balanced by the i-component of the force exerted by the remainder of the fluid on the surface element, i.e. in a fluid at rest through the *hydrostatic pressure*. All in all, one thus finds that the equilibrium fluxes in the comoving local rest frame are

$$\vec{\mathcal{J}}_E^{\text{eq.}}\Big|_{\mathcal{R}_{\vec{\mathsf{v}}}} = \vec{0}, \quad \vec{\mathcal{J}}_N^{\text{eq.}}\Big|_{\mathcal{R}_{\vec{\mathsf{v}}}} = \vec{0}, \quad \mathbf{J}_{\vec{P}}^{\text{eq.}}\Big|_{\mathcal{R}_{\vec{\mathsf{v}}}} = \mathcal{P}\,\mathbb{1}, \tag{I.65}$$

(n) P. Duhem, 1861–1916

where the latter identity can be expressed in term of components as $(J^{\text{eq.}}_{\vec{P}})_{ij} = \mathcal{P}\,\delta_{ij}$.

When performing the Galilean transformation with velocity $-\vec{\mathsf{v}}$ to the fixed reference frame $\mathcal{R}_0$, two effects have to be taken into account. First, the transported quantities may be modified, as is the case of energy density or momentum density. Secondly, a flux is defined as the quantity flowing per unit time through a motionless surface, and the motionless surfaces in both frames differ.

To consistently account for both these effects, one should first consider an infinitesimal Galilean transformation from a frame $\mathcal{R}_{\vec{\mathsf{v}}'}$ with velocity $\vec{\mathsf{v}}'$ (with respect to $\mathcal{R}_0$) to a frame $\mathcal{R}_{\vec{\mathsf{v}}'+\mathrm{d}\vec{\mathsf{v}}'}$ with velocity $\vec{\mathsf{v}}' + \mathrm{d}\vec{\mathsf{v}}'$. In $\mathcal{R}_{\vec{\mathsf{v}}'}$, the characteristic densities at a point moving with velocity $\vec{\mathsf{v}}$ in $\mathcal{R}_0$ take the values [cf. Eq. (I.62b)]

$$e + \frac{1}{2}\rho(\vec{\mathsf{v}} - \vec{\mathsf{v}}')^2, \quad n, \quad \rho(\mathsf{v}_i - \mathsf{v}'_i).$$

Observed from $\mathcal{R}_{\vec{\mathsf{v}}'+\mathrm{d}\vec{\mathsf{v}}'}$, the densities at the same point become

$$e + \frac{1}{2}\rho(\vec{\mathsf{v}} - \vec{\mathsf{v}}' - \mathrm{d}\vec{\mathsf{v}}')^2 = e + \frac{1}{2}\rho(\vec{\mathsf{v}} - \vec{\mathsf{v}}')^2 - \rho(\vec{\mathsf{v}} - \vec{\mathsf{v}}')\cdot\mathrm{d}\vec{\mathsf{v}}', \quad n, \quad \rho(\mathsf{v}_i - \mathsf{v}'_i - \mathrm{d}\mathsf{v}'_i) = \rho(\mathsf{v}_i - \mathsf{v}'_i) - n\,m\,\mathrm{d}\mathsf{v}'_i.$$

From the latter formulae, one deduces the variations of the densities in the infinitesimal Galilean transformation, namely respectively

$$\mathrm{d}\bigg(e + \frac{1}{2}\rho\,\vec{\mathsf{w}}^{\,2}\bigg) = \rho\,\vec{\mathsf{w}}\cdot\mathrm{d}\vec{\mathsf{w}}, \quad \mathrm{d}n = 0, \quad \mathrm{d}\big(\rho\,\vec{\mathsf{w}}\big) = n\,m\,\mathrm{d}\vec{\mathsf{w}},$$

where we have set $\vec{\mathsf{w}} = \vec{\mathsf{v}} - \vec{\mathsf{v}}'$ and accordingly $\mathrm{d}\vec{\mathsf{w}} = -\mathrm{d}\vec{\mathsf{v}}'$.

Consider now the fluxes of the extensive quantities at the same point. Recognizing in the infinitesimal variation of the energy density above the momentum density $\rho\,\vec{\mathsf{w}}$ multiplied by the velocity increment, one deduces that the change in energy flux due to the variation of the transported energy density will involve the momentum flux, again multiplied by the velocity increment. Similarly, the variation of momentum density involves the particle number density, so that the change in momentum flux will involve the flux of particle number.

On the other hand, the motion with velocity $-\mathrm{d}\vec{\mathsf{v}}'$ in $\mathcal{R}_{\vec{\mathsf{v}}'+\mathrm{d}\vec{\mathsf{v}}'}$ of a surface which is motionless in $\mathcal{R}_{\vec{\mathsf{v}}'}$ contributes to any flux an amount given to first order in $\mathrm{d}\vec{\mathsf{w}}$ by the product of the corresponding density as measured in $\mathcal{R}_{\vec{\mathsf{v}}'}$ with $\mathrm{d}\vec{\mathsf{w}}$

All in all, the differences between the values of the equilibrium fluxes as measured in $\mathcal{R}_{\vec{\mathsf{v}}'+\mathrm{d}\vec{\mathsf{v}}'}$ and $\mathcal{R}_{\vec{\mathsf{v}}'}$ read

$$\mathrm{d}\mathcal{J}^{\text{eq.}}_{E,i} = \sum_{j=1}^{3}(\mathsf{J}^{\text{eq.}}_{\vec{P}})_{ij}\,\mathrm{d}\mathsf{w}_j + \bigg(e + \frac{1}{2}\rho\,\vec{\mathsf{w}}^{\,2}\bigg)\mathrm{d}\mathsf{w}_i, \quad \mathrm{d}\vec{\mathcal{J}}^{\text{eq.}}_N = n\,\mathrm{d}\vec{\mathsf{w}}, \quad \mathrm{d}(\mathsf{J}^{\text{eq.}}_{\vec{P}})_{ij} = \mathcal{J}^{\text{eq.}}_{N,j}\,m\,\mathrm{d}\mathsf{w}_i + \rho\,\mathsf{w}_i\,\mathrm{d}\mathsf{w}_j.$$

These equations can be viewed as defining partial differential equations, which can be integrated from $\vec{\mathsf{w}} = \vec{0}$ (i.e. from the local rest frame comoving with the fluid at velocity $\vec{\mathsf{v}}$) to $\vec{\mathsf{w}} = \vec{\mathsf{v}}$ (the fixed frame), starting with known initial conditions at velocity $\vec{\mathsf{w}} = \vec{0}$—namely the equilibrium fluxes (I.65). One first finds $\vec{\mathcal{J}}^{\text{eq.}}_N$, which is then injected in the equations for $(\mathsf{J}^{\text{eq.}}_{\vec{P}})_{ij}$. Solving the latter, the result can be used in the equations for $\mathcal{J}^{\text{eq.}}_{E,i}$.

In the fixed reference frame $\mathcal{R}_0$, the equilibrium fluxes read

$$\vec{\mathcal{J}}^{\text{eq.}}_E = \bigg(e + \frac{1}{2}\rho\vec{\mathsf{v}}^2 + \mathcal{P}\bigg)\vec{\mathsf{v}}, \quad \vec{\mathcal{J}}^{\text{eq.}}_N = n\vec{\mathsf{v}}, \quad \big(J^{\text{eq.}}_{\vec{P}}\big)_{ij} = \mathcal{P}\,\delta_{ij} + \rho\,\mathsf{v}_i\mathsf{v}_j. \tag{I.66}$$

These equilibrium fluxes have to be complemented with affinity-dependent terms, to yield the total fluxes. For the sake of simplicity, we shall restrict ourselves to Markovian, memoryless transport processes.

By definition of the non-relativistic flow velocity as the average velocity of particles in the fluid cell under consideration, the flux of particle number does not receive any such extra term, and will always remain equal to its equilibrium value, $\vec{\mathcal{J}}_N = \vec{\mathcal{J}}^{\text{eq.}}_N$.

As a consequence (cf. the derivation above), the difference between the values of the momentum flux components $\big(J_{\vec{P}}\big)_{ij}$ taken in the frames $\mathcal{R}_0$ and $\mathcal{R}_{\vec{\mathsf{v}}}$ is the same as for the corresponding equilibrium fluxes, namely $\rho\,\mathsf{v}_i\mathsf{v}_j$. The only dependence on any affinity in $\big(\mathsf{J}_{\vec{P}}\big)_{ij}$ thus already affects its value in the local rest frame of the fluid, in which the diagonal tensor $\mathcal{P}\,\delta_{ij}$ becomes an affinity-dependent tensor $\boldsymbol{\pi}$ with components π_{ij} called *stress tensor*.

Eventually, this stress tensor plays a role in the change of frame for the energy flux, with $\boldsymbol{\pi}\cdot\vec{\mathsf{v}}$ replacing $\mathcal{P}\,\vec{\mathsf{v}}$. Additionally, one also has to allow for an affinity-dependent contribution $\vec{\mathcal{J}}_U$ to the energy flux in the local rest frame. Altogether, one obtains the fluxes

$$\vec{\mathcal{J}}_E = \left(e+\frac{1}{2}\rho\vec{\mathsf{v}}^2\right)\vec{\mathsf{v}} + \vec{\mathcal{J}}_U + \boldsymbol{\pi}\cdot\vec{\mathsf{v}}, \quad \vec{\mathcal{J}}_N = n\vec{\mathsf{v}}, \quad \big(\mathcal{J}_{\vec{P}}\big)_{ij} = \pi_{ij} + \rho\,\mathsf{v}_i\mathsf{v}_j, \tag{I.67}$$

where the functional dependences of $\vec{\mathcal{J}}_U$ and the stress tensor $\boldsymbol{\pi}$ on the various affinities depend on the model under consideration, i.e. concretely on the specific fluid under study.

Remark: One can also show that angular momentum conservation implies that the stress tensor is symmetric.

I.2.4 b Conservation laws

Having obtained the fluxes in the fixed inertial frame $\mathcal{R}_0$, we can now insert their expressions and those of the densities (I.62b) in the general local balance equation (I.18a), where the source/sink term will vanish since energy, particle number and momentum are conserved.

Starting with particle number density, or equivalently—to respect the tradition—with the mass density $\rho = nm$, the simple flux $m\vec{\mathcal{J}}_N = \rho\,\vec{\mathsf{v}}$ leads to the continuity equation

$$\boxed{\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\big(\rho\,\vec{\mathsf{v}}\big) = 0,} \tag{I.68}$$

which locally expresses mass conservation.

In a second step, we can consider the balance equation for momentum, for the sake of simplicity component by component. The momentum density $\rho\,\mathsf{v}_i$ and flux $\mathsf{J}_{\vec{P}}^{ij}$ give for every $i\in\{1,2,3\}$

$$\frac{\partial(\rho\,\mathsf{v}_i)}{\partial t} + \sum_{j=1}^{3}\frac{\partial}{\partial x_j}\big(\pi_{ij}+\rho\,\mathsf{v}_i\mathsf{v}_j\big) = 0.$$

A simple calculation using the mass balance equation (I.68) to cancel two terms allows one to rewrite this relation as

$$\boxed{\rho\left[\frac{\partial\mathsf{v}_i}{\partial t} + \big(\vec{\mathsf{v}}\cdot\vec{\nabla}\big)\mathsf{v}_i\right] + \sum_{j=1}^{3}\frac{\partial\pi_{ij}}{\partial x_j} = 0.} \tag{I.69}$$

Since the stress tensor component π_{ij} represents the amount of momentum in direction i flowing per unit time through a unit surface perpendicular to direction j, it equals according to Newton's(o) second law the i-th component of the force per unit area on a surface normal to the direction j. Following Newton's third law, this equals the negative of the force per unit area exerted by the remainder of the fluid on the cell under study, so that this local balance equation actually expresses the fundamental principle of Newtonian dynamics.

Eventually, the energy density (I.62b) and flux (I.67) yield for the local balance equation for energy in the fixed reference frame

$$\frac{\partial}{\partial t}\left(e+\frac{1}{2}\rho\vec{\mathsf{v}}^2\right) + \vec{\nabla}\cdot\vec{\mathcal{J}}_E = \frac{\partial}{\partial t}\left(e+\frac{1}{2}\rho\vec{\mathsf{v}}^2\right) + \sum_{i=1}^{3}\frac{\partial}{\partial x_i}\left[\left(e+\frac{1}{2}\rho\vec{\mathsf{v}}^2\right)\mathsf{v}_i + \vec{\mathcal{J}}_U + \sum_{j=1}^{3}\pi_{ij}\mathsf{v}_j\right] = 0.$$

To simplify this equation, one views $\frac{1}{2}\rho\,\vec{\mathsf{v}}^2$ as ρ times $\frac{1}{2}\vec{\mathsf{v}}^2$ and then differentiates with the usual product rule. The terms proportional to $\frac{1}{2}\vec{\mathsf{v}}^2$ vanish thanks to the continuity equation (I.68). Those

---

(o) I. Newton, 1642–1727

proportional to ρ can be rewritten with the help of the identity $\vec{\nabla}\big(\frac{1}{2}\vec{\mathsf{v}}^2\big) = \big(\vec{\mathsf{v}}\cdot\vec{\nabla}\big)\vec{\mathsf{v}}+\vec{\mathsf{v}}\times\big(\vec{\nabla}\times\vec{\mathsf{v}}\big)$, and one recognizes the sum over i of $\rho\mathsf{v}_i$ multiplied with the term within square brackets in Eq. (I.69). A further application to the product rule leads then to

$$\boxed{\frac{\partial e}{\partial t} + \vec{\nabla}\cdot\big(e\,\vec{\mathsf{v}} + \vec{\mathcal{J}}_U\big) + \sum_{i,j=1}^{3}\pi_{ij}\frac{\partial \mathsf{v}_j}{\partial x_i} = 0.} \tag{I.70}$$

Interestingly, the kinetic energy density no longer appears in this equation, which can be interpreted as describing the change in the internal energy due to dissipation ($\vec{\mathcal{J}}_U$) and to the work of forces exerted on neighboring fluid cells (π_{ij}).

I.2.4 c Linear Markovian transport processes in a simple fluid

Let us now explicitly consider a model for the fluid, by assuming that the fluxes (I.67)—and more precisely the flux $\vec{\mathcal{J}}_U$ and the stress tensor $\boldsymbol{\pi}$—are linear functions of the affinities, i.e. of the first derivatives of the intensive variables (I.63) with respect to space coordinates. This assumption defines *Newtonian fluids*, which are those governed by the resulting dynamical laws derived from Eqs. (I.69) and (I.70).

In the linear regime, one only needs to introduce first-order kinetic coefficients L_{ab} relating the fluxes to the affinities as in Eq. (I.31). Enumerating the former, the two vectors $\vec{\mathcal{J}}_N$ and $\vec{\mathcal{J}}_E$ and the rank 2 tensor $\mathbf{J}_{\vec{P}}$ amount altogether to 15 components. Similarly, the gradients of the intensive variables (I.63) also represent 15 different scalar fields, so that a naive approach would necessitate 15×15 kinetic coefficients L_{ab}. Fortunately, the problem can be considerably simplified thanks to the general principles of Sec. I.2.2 and to system-specific properties, resulting in a small number of coefficients.

A first important remark is that by definition of the flow velocity $\vec{\mathsf{v}}$, the particle-number flux $\vec{\mathcal{J}}_N$ always remains equal to its equilibrium contribution $\vec{\mathcal{J}}_N^{\text{eq.}} = n\,\vec{\mathsf{v}}$, irrespective of the affinities. This means that all coefficients L_{Na} vanish. Thanks to the Onsager relations, all reciprocal coefficients L_{aN} also vanish. Thus a gradient in chemical potential does not lead to any dissipation in a simple fluid: chemical potential plays no direct role in entropy production.(18)

Using the second consequence listed in Sec. I.2.2 of Curie's symmetry principle applied to isotropic media, the vectorial flux (resp. affinity) for energy cannot couple to the affinity (resp. flux) for momentum, which is a tensor of rank 2. That is, the tensors of rank 3 coupling E and $\vec{P}$ identically vanish. All in all, this means that there is no indirect transport in a simple fluid.

According to the first of the consequences of isotropy given in Sec. I.2.2, the transport of energy—a scalar quantity—involves a tensor $\mathbf{L}_{EE}$ of rank 2 proportional to the identity, i.e. effectively a single kinetic coefficient L_{EE}. To write down this relation explicitly, it is convenient to move to a rest frame in which the fluid is locally at rest, so that the energy flux has a simple expression.

In the Galilean transformation from the fixed frame $\mathcal{R}_0$ to the frame $\mathcal{R}_{\vec{\mathsf{v}}_0}$ comoving with the fluid at point M_0, where the fluid velocity at time t_0 is $\vec{\mathsf{v}}_0$, the position and velocity respectively transform according to $\vec{r}\,' = \vec{r} - \vec{\mathsf{v}}_0(t-t_0)$, which implies the identity of derivatives $\partial/\partial x_i' = \partial/\partial x_i$, and $\vec{\mathsf{v}}\,'(t,\vec{r}\,') = \vec{\mathsf{v}}(t,\vec{r}) - \vec{\mathsf{v}}_0$, where primed resp. unprimed quantities refer to $\mathcal{R}_{\vec{\mathsf{v}}_0}$ resp. $\mathcal{R}_0$. Since temperature is the same in all frames, the affinity conjugate to the energy is $\vec{\nabla}(1/T)$, so that the relation between this affinity and the energy flux in the comoving rest frame reads

$$\vec{\mathcal{J}}_U = L_{EE}\vec{\nabla}\bigg(\frac{1}{T}\bigg), \tag{I.71a}$$

---

(18) A gradient in chemical potential, or equivalently in particle number density, at uniform temperature will lead via the equations of state to a gradient in pressure $\mathcal{P}$—cf. the example of an ideal gas—, which results in a macroscopic flow through Eq. (I.69). In turn, this motion will lead to dissipation due to the viscous effects described by the transport coefficients η and ζ.

where L_{EE} is as always nonnegative.

In turn, the i-th component in $\mathcal{R}_{\vec{\mathsf{v}}_0}$ of the affinity $\vec{\mathscr{F}}_{P_j}$ is given by

$$\frac{\partial}{\partial x'_i}\left(-\frac{\mathsf{v}'_j}{T}\right)\bigg|_{t_0,M_0} = \frac{\partial}{\partial x_i}\left(-\frac{\mathsf{v}_j-\mathsf{v}_{0,j}}{T}\right)\bigg|_{t_0,M_0} = -\frac{1}{T}\frac{\partial \mathsf{v}_j}{\partial x_i}\bigg|_{t_0,M_0},$$

where the term proportional to $\partial(1/T)/\partial x_i$ vanishes since it multiplies $\mathsf{v}_j-\mathsf{v}_{0,j}$ taken at time t_0 and at the point M_0. The remaining task is thus to express the components of stress tensor $\boldsymbol{\pi}$, which equals the momentum flux in the comoving local rest frame, as function of the derivatives $\partial\mathsf{v}_k/\partial x_l$. This necessitates a tensor $\mathbf{L}_{\vec{P}\vec{P}}$ of rank 4 with components $L^{ijkl}_{\vec{P}\vec{P}}$ such that

$$\pi_{ij} = \sum_{k,l=1}^{3} L^{ijkl}_{\vec{P}\vec{P}}\,\frac{\partial\mathsf{v}_k}{\partial x_l}.$$

Invoking again the local isotropy of the fluid, this tensor can only be a linear combination of the three rank-4 tensors invariant under rotations, namely those with (Cartesian) components $\delta_{ij}\delta_{kl}$, $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$, with three respective kinetic coefficients.

As noted above, $\boldsymbol{\pi}$ must be symmetric ($\pi_{ij}=\pi_{ji}$) to ensure angular momentum conservation, so that the coefficients of $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$ must be identical. Instead of considering π_{ij} as linear combination of $\delta_{ij}\delta_{kl}$ and $\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}$, one traditionally—and equivalently—writes down a linear combination of $\delta_{ij}\delta_{kl}$ and the traceless tensor $\frac{1}{2}\big(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}\big)-\frac{1}{3}\delta_{ij}\delta_{kl}$. Introducing the two necessary non-negative coefficients $L^{(1)}_{\vec{P}\vec{P}}$, $L^{(2)}_{\vec{P}\vec{P}}$, the relation between the stress tensor components and the affinities conjugate to momentum reads

$$\pi_{ij} = \mathcal{P}\,\delta_{ij} - \frac{L^{(1)}_{\vec{P}\vec{P}}}{T}\sum_{k,l=1}^{3}\delta_{ij}\delta_{kl}\frac{\partial\mathsf{v}_k}{\partial x_l} - \frac{L^{(2)}_{\vec{P}\vec{P}}}{T}\sum_{k,l=1}^{3}\left[\frac{1}{2}\big(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}\big)-\frac{1}{3}\delta_{ij}\delta_{kl}\right]\frac{\partial\mathsf{v}_k}{\partial x_l},$$

where we also included the equilibrium part $\mathcal{P}\,\delta_{ij}$. With the help of the traceless symmetric tensor $\boldsymbol{\sigma}$ with components

$$\sigma_{ij} \equiv \frac{1}{2}\left(\frac{\partial\mathsf{v}_i}{\partial x_j}+\frac{\partial\mathsf{v}_j}{\partial x_i}\right)-\frac{1}{3}\big(\vec{\nabla}\cdot\vec{\mathsf{v}}\big)\,\delta_{ij}, \tag{I.71b}$$

the stress tensor can be rewritten more concisely as

$$\boldsymbol{\pi} = \mathcal{P}\,\mathbb{1} - \frac{L^{(1)}_{\vec{P}\vec{P}}}{T}\big(\vec{\nabla}\cdot\vec{\mathsf{v}}\big)\,\mathbb{1} - \frac{L^{(2)}_{\vec{P}\vec{P}}}{T}\,\boldsymbol{\sigma} \tag{I.71c}$$

i.e. component-wise

$$\pi_{ij} = \mathcal{P}\,\delta_{ij} - \frac{L^{(1)}_{\vec{P}\vec{P}}}{T}\big(\vec{\nabla}\cdot\vec{\mathsf{v}}\big)\,\delta_{ij} - \frac{L^{(2)}_{\vec{P}\vec{P}}}{T}\,\sigma_{ij}. \tag{I.71d}$$

This relation and Eq. (I.71a) are the characteristic “constitutive equations” for a Newtonian fluid.

Let us now interpret the three kinetic coefficients L_{EE}, $L^{(1)}_{\vec{P}\vec{P}}$ and $L^{(2)}_{\vec{P}\vec{P}}$ in terms of more traditional transport coefficients.

First, Eq. (I.71a) is naturally reminiscent of Fourier's law

$$\vec{\mathcal{J}}_U = -\kappa\vec{\nabla}T \quad\text{with}\quad \kappa = \frac{L_{EE}}{T^2} \tag{I.72}$$

like in the case of an insulator [Eq. (I.37a)]. L_{EE} is thus related to the heat conductivity, which is nonnegative, as it should be.

As was already mentioned, π_{ij} is the i-component of the force per unit area acting on a surface normal to the j-direction. Phenomenologically, this force per unit area is related to the gradient along direction j of the i-th component of velocity through *Newton's law of viscosity*(19)

---

(19) This law is defined for a fluid flowing uniformly along the i-direction, so that the velocity only depends on x_j, which ensures that both $\partial\mathsf{v}_i/\partial x_i$—and thereby $\vec{\nabla}\cdot\vec{\mathsf{v}}$—and $\partial\mathsf{v}_j/\partial x_i$ vanish.

$$\pi_{ij} = -\eta\,\frac{\mathrm{d}\mathrm{v}_i}{\mathrm{d}x_j}, \tag{I.73}$$

with η the *shear viscosity* of the fluid. Identifying this empirical law with relation (I.71d) for $i \neq j$ yields

$$\eta = \frac{L^{(2)}_{\vec{P}\vec{P}}}{2T}. \tag{I.74}$$

Eventually, the parameter $L^{(1)}_{\vec{P}\vec{P}}$ is related to the transport parameter referred to as *volume viscosity* (or at times *second viscosity* or *bulk viscosity*(20)) ζ, which only plays a role in compressible flows ($\vec{\nabla}\cdot\vec{\mathrm{v}} \neq 0$), in particular in the damping of sound waves. To obtain the proper form for the equation of motion of a compressible flow, one must set

$$\zeta = \frac{L^{(1)}_{\vec{P}\vec{P}}}{T}. \tag{I.75}$$

With Eqs. (I.74) and (I.75), the stress tensor component (I.71c) becomes

$$\boldsymbol{\pi} = \mathcal{P}\,\mathbb{1} - \zeta\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big)\,\mathbb{1} - 2\eta\,\boldsymbol{\sigma}. \tag{I.76}$$

Rewriting Eqs. (I.69) and (I.70) in the linear regime in which the internal energy and momentum fluxes are respectively given by Eqs. (I.72) and (I.76), one obtains in the case of position-independent viscosity coefficients the *Navier*[p]*–Stokes*[q] *equation*

$$\boxed{\rho\bigg[\frac{\partial\vec{\mathrm{v}}}{\partial t} + \big(\vec{\mathrm{v}}\cdot\vec{\nabla}\big)\vec{\mathrm{v}}\bigg] = -\vec{\nabla}\mathcal{P} + \eta\bigg[\triangle\vec{\mathrm{v}} + \frac{1}{3}\vec{\nabla}\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big)\bigg] + \zeta\vec{\nabla}\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big)} \tag{I.77}$$

and the energy balance equation

$$\boxed{\begin{aligned}&\frac{\partial}{\partial t}\bigg(e+\frac{1}{2}\rho\vec{\mathrm{v}}^2\bigg) + \vec{\nabla}\cdot\bigg\{\bigg(e+\frac{1}{2}\rho\vec{\mathrm{v}}^2+\mathcal{P}\bigg)\vec{\mathrm{v}}\\ &\qquad -\eta\bigg[\big(\vec{\mathrm{v}}\cdot\vec{\nabla}\big)\vec{\mathrm{v}} + \vec{\nabla}\bigg(\frac{\vec{\mathrm{v}}^2}{2}\bigg) - \frac{2}{3}\vec{\mathrm{v}}\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big)\bigg] - \zeta\vec{\mathrm{v}}\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big) - \kappa\vec{\nabla}T\bigg\} = 0.\end{aligned}} \tag{I.78}$$

I.2.4 d Entropy production

To conclude this application of the thermodynamics of linear irreversible processes to simple fluids, we can write down the balance equation for entropy (I.20).

For that purpose, we first need the expression of the entropy flux density. Using the general formula (I.22) in the comoving local rest frame $\mathcal{R}_{\vec{\mathrm{v}}}$, in which $\vec{\mathcal{J}}_N = \vec{0}$ and $\mathcal{Y}_{\vec{P}} = \vec{0}$ since both are in the general case proportional to the fluid velocity, yields

$$(\vec{\mathcal{J}}_S)_{\mathcal{R}_{\vec{\mathrm{v}}}} = \frac{1}{T}\,\vec{\mathcal{J}}_U. \tag{I.79a}$$

Transforming back to the fixed frame $\mathcal{R}_0$, the entropy density s is invariant, so that the entropy flux only changes because surfaces at rest in $\mathcal{R}_{\vec{\mathrm{v}}}$ are now moving:

$$\vec{\mathcal{J}}_S = \frac{1}{T}\,\vec{\mathcal{J}}_U + s\,\vec{\mathrm{v}} \tag{I.79b}$$

---

(20) Some authors, as e.g. in Ref. [21], reserve the name “bulk viscosity” to the combination $\zeta + \frac{2}{3}\eta$. The lack of unity in the terminology shows how little this coefficient has actually been studied!

---

[p] C.L. Navier, 1785–1836 [q] G. G. Stokes, 1819–1903

Inserting in the general formula for the entropy production (I.30) the fluxes (I.67) from which one subtracts their equilibrium values (I.66) with the corresponding affinities, one finds(21)

$$\sigma_S = \vec{\nabla}\left(\frac{1}{T}\right)\cdot\vec{\mathcal{J}}_U + \sum_{i,j=1}^{3}\left(-\frac{1}{T}\frac{\partial \mathsf{v}_j}{\partial x_i}\right)\left(\pi_{ij} - \mathcal{P}\delta_{ij}\right). \tag{I.80}$$

In the linear approximation where the fluxes are respectively given by Eqs. (I.72) and (I.76), this becomes

$$\sigma_S = \kappa T^2\left[\vec{\nabla}\left(\frac{1}{T}\right)\right]^2 + \sum_{i,j=1}^{3}\left(-\frac{1}{T}\frac{\partial \mathsf{v}_j}{\partial x_i}\right)\left[-2\eta\sigma_{ij} - \zeta(\vec{\nabla}\cdot\vec{\mathsf{v}})\delta_{ij}\right].$$

The rightmost term between square brackets is symmetric under the exchange of i and j, so that one can replace $\partial \mathsf{v}_j/\partial x_i$ in front by half of the symmetrized version, i.e. according to definition (I.71b) by $\sigma_{ij} + \frac{1}{3}(\vec{\nabla}\cdot\vec{\mathsf{v}})\delta_{ij}$. Canceling out the minus signs, one thus has

$$\sigma_S = \kappa T^2\left[\vec{\nabla}\left(\frac{1}{T}\right)\right]^2 + \frac{1}{T}\sum_{i,j=1}^{3}\left[\sigma_{ij} + \frac{1}{3}(\vec{\nabla}\cdot\vec{\mathsf{v}})\delta_{ij}\right]\left[2\eta\sigma_{ij} + \zeta(\vec{\nabla}\cdot\vec{\mathsf{v}})\delta_{ij}\right].$$

The remaining product can then be expanded. Multiplying a rank 2 tensor by δ_{ij} and taking the sum over i and j amounts to taking the trace of the tensor. Since σ_{ij} is traceless, the products $\sigma_{ij}\delta_{ij}$ yield 0, while $\delta_{ij}\delta_{ij}$ gives 3. In the end, there remains

$$\sigma_S = \kappa T^2\left[\vec{\nabla}\left(\frac{1}{T}\right)\right]^2 + \frac{\zeta}{T}(\vec{\nabla}\cdot\vec{\mathsf{v}})^2 + \frac{2\eta}{T}\sum_{i,j=1}^{3}(\sigma_{ij})^2. \tag{I.81}$$

Since the three transport coefficients κ, η and ζ are nonnegative, the entropy production rate is always positive—as it should be.

Bibliography for Chapter I

- Blundell & Blundell, *Concepts in thermal physics* [22], chapter 34.
- Callen, *Thermodynamics and an introduction to Thermostatistics* [8], chapter 14.
- de Groot & Mazur, *Non-equilibrium thermodynamics* [1], chapters II–IV, VI, XI §1 & 2, XII §1.
- Le Bellac, Mortessagne & Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* [9], chapter 6.
- Pottier, *Nonequilibrium statistical physics* [6], chapter 2.

---

(21) This is most obvious in the comoving frame, holds however in a general frame.

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 such an approach in practice inevitably involves probabilities to be 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 cm^3 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-energy collisions of heavy nuclei are often considered as forming a statistical system—even worse, a system possibly in local thermodynamic equilibrium.

---

(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.

$*$ At the other end of the length spectrum, some astrophysical 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 $\lesssim 1\,\mu$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 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 t be about $\epsilon\,\mathrm{e}^{\lambda t}$ away from each other with some *Lyapunov*(r) *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 macroscopic quantities, defining a *macrostate*.

Such a macroscopic observable usually results from a sum over many particles of microscopic quantities, and 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 phenomenological 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.41b) 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 $\langle \vec{J}_{\rm el.} \rangle = \sigma_{\rm el.} \langle \vec{\mathscr{E}} \rangle$, which describes the actual procedure better than the traditional form (I.41b), and again emphasizes the need for a statistical approach.

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.

---

(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.

(r) A. Lyapunov, 1857–1918

After recalling the basics of the phase-space based formalism for exactly known classical systems (Sec. 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 (Sec. II.2.2). We next derive the evolution equations governing the dynamics of that density (Sec. II.2.3) and of observables (Sec. II.2.4). Eventually, we shortly mention the generalization to systems whose particle number is not fixed (Sec. 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 $3N$ space coordinates of the particles $q_1,\dots,q_{3N}$ 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 classical system, a measurable quantity—or (classical) *observable*—is defined as a phase-space function $O_N(\{q_i\},\{p_i\})$ of the $6N$ variables. We shall only consider observables without explicit time dependence, i.e. the mathematical function O_N remains constant over time.

Remark: 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_i and p_i 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 $(\{\mathfrak{q}_i(t)\},\{\mathfrak{p}_i(t)\})$ 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 $\{\dot{\mathfrak{q}}_i(t)\},\{\dot{\mathfrak{p}}_i(t)\}$.

The dynamics of the system—or equivalently, of the representing point in Γ-space—is fully determined by specifying the time-independent Hamilton(s) function $H_N(\{q_i\},\{p_i\})$. More precisely, the trajectory $(\{\mathfrak{q}_i(t)\},\{\mathfrak{p}_i(t)\})$ is governed by the *Hamilton equations*

$$\begin{aligned}\dot{\mathfrak{q}}_i(t) &\equiv \frac{\mathrm{d}\mathfrak{q}_i(t)}{\mathrm{d}t} = \frac{\partial H_N}{\partial p_i} = \{q_i,H_N\},\\ \dot{\mathfrak{p}}_i(t) &\equiv \frac{\mathrm{d}\mathfrak{p}_i(t)}{\mathrm{d}t} = -\frac{\partial H_N}{\partial q_i} = \{p_i,H_N\}, \qquad i=1,\dots,3N,\end{aligned} \tag{II.1}$$

where the derivatives of the Hamilton function, and accordingly the Poisson brackets, are to be computed at the point in Γ-space where the system sits at time t, i.e. at $(\{q_i=\mathfrak{q}_i(t)\},\{p_i=\mathfrak{p}_i(t)\})$ for $1\le i\le 3N$. The Poisson(t) bracket of two functions f, g defined on phase space is given by(25)

$$\{f,g\} \equiv \sum_{i=1}^{3N}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right), \tag{II.2}$$

wherethe arguments of the functions and their derivatives have been dropped for the sake of brevity.

---

(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].

(s) W. R. Hamilton, 1805–1865 (t) S. Poisson, 1781–1840

It is important to realize that the Hamilton equations are fully deterministic once the Hamilton function is fixed: given an initial condition $(\{\mathfrak{q}_i(0)\},\{\mathfrak{p}_i(0)\})$, the microstate at time t is uniquely determined by Eqs. (II.1). Accordingly, there is only a single trajectory through each individual point of the Γ-space, so that the notation $\mathbf{u}(\{q_i\},\{p_i\})$ 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 $\rho_N(t,\{q_i\},\{p_i\})$ on the Γ-space, the *N-particle phase space density*, which is as always nonnegative and normalized to unity

$$\rho_N\big(t,\{q_i\},\{p_i\}\big) \geq 0 \quad \forall\{q_i\},\{p_i\} \quad \text{and} \quad \int_\Gamma \rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathrm{d}^{6N}\mathcal{V} = 1, \tag{II.3}$$

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

$$\mathrm{d}^{6N}\mathcal{V} = C_N \prod_{i=1}^{3N} \mathrm{d}q_i\,\mathrm{d}p_i, \tag{II.4a}$$

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

$$\mathrm{d}^{6N}\mathcal{V} = \frac{1}{N!}\prod_{i=1}^{3N} \frac{\mathrm{d}q_i\,\mathrm{d}p_i}{2\pi\hbar}, \tag{II.4b}$$

with $\hbar$ the reduced Planck(u) constant. A further advantage of this choice is that $\mathrm{d}^{6N}\mathcal{V}$ 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 in Γ-space $(\{\mathfrak{q}_i(t)\},\{\mathfrak{p}_i(t)\})$ at time t can be viewed as a $6N$-dimensional random variable. Again, the microstate at t is *not* random if $(\{\mathfrak{q}_i(t=0)\},\{\mathfrak{p}_i(t=0)\})$ is known, but randomness enters due to our knowledge of the initial condition only on a statistical basis.

In turn, the value taken by an observable $O_N(\{q_i\},\{p_i\})$ for a given system—corresponding to a given trajectory in Γ-space—at time t, namely

$$O_N(t) \equiv O_N\big(\{q_i = \mathfrak{q}_i(t)\},\{p_i = \mathfrak{p}_i(t)\}\big), \tag{II.5}$$

is also a random variable, whose momenta are given by the usual formulae (see Appendix B). For instance, the average value at time t is

$$\langle O_N(t)\rangle_t = \int_\Gamma O_N\big(\{q_i\},\{p_i\}\big)\,\rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathrm{d}^{6N}\mathcal{V} \tag{II.6a}$$

and the variance reads

$$\sigma^2_{O_N}(t) = \big\langle O_N(t)^2\big\rangle_t - \langle O_N(t)\rangle_t^2\,, \tag{II.6b}$$

(u) M. Planck, 1858–1947

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

Remark: Even 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 $\mathcal{V}$ in the 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

$$\mathcal{N}(t) = N\int_{\mathcal{V}} \rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathrm{d}^{6N}\mathcal{V}, \tag{II.7}$$

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

$$\frac{\mathrm{d}\mathcal{N}(t)}{\mathrm{d}t} = N\int_{\mathcal{V}} \frac{\partial\rho_N\big(t,\{q_i\},\{p_i\}\big)}{\partial t}\,\mathrm{d}^{6N}\mathcal{V}. \tag{II.8}$$

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 $\mathbf{e}_\mathrm{n}(\{q_i\},\{p_i\})$ denote the unit vector normal to $\partial\mathcal{V}$ at a given point, oriented towards the outside of $\mathcal{V}$, and $\mathbf{u}(\{q_i\},\{p_i\})$ be the velocity-vector tangent to the trajectory passing through the point $(\{q_i\},\{p_i\})$. One then has

$$\frac{\mathrm{d}\mathcal{N}(t)}{\mathrm{d}t} = -N\int_{\partial\mathcal{V}} \rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathbf{u}(\{q_i\},\{p_i\})\cdot\mathbf{e}_\mathrm{n}(\{q_i\},\{p_i\})\,\mathrm{d}^{6N-1}\mathcal{S}.$$

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

$$\frac{\mathrm{d}\mathcal{N}(t)}{\mathrm{d}t} = -N\int_{\mathcal{V}} \boldsymbol{\nabla}\cdot\big[\rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathbf{u}(\{q_i\},\{p_i\})\big]\,\mathrm{d}^{6N}\mathcal{V}, \tag{II.9}$$

with $\boldsymbol{\nabla}$ 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 are not written)

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

The divergence in the left-hand side can be rewritten as

$$\boldsymbol{\nabla}\cdot\big(\rho_N\mathbf{u}\big) = \sum_{i=1}^{3N}\frac{\partial}{\partial q_i}\big(\rho_N\dot{\mathfrak{q}}_i\big) + \sum_{i=1}^{3N}\frac{\partial}{\partial p_i}\big(\rho_N\dot{\mathfrak{p}}_i\big) = \sum_{i=1}^{3N}\left(\frac{\partial\rho_N}{\partial q_i}\,\dot{\mathfrak{q}}_i + \frac{\partial\rho_N}{\partial p_i}\,\dot{\mathfrak{p}}_i\right) + \sum_{i=1}^{3N}\left(\frac{\partial\dot{\mathfrak{q}}_i}{\partial q_i} + \frac{\partial\dot{\mathfrak{p}}_i}{\partial p_i}\right)\rho_N.$$

With the help of the Hamilton equations (II.1) this gives the *Liouville*[v] *equation*[26]

$$\boxed{\frac{\partial\rho_N}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial\rho_N}{\partial q_i}\frac{\partial H_N}{\partial p_i} - \frac{\partial\rho_N}{\partial p_i}\frac{\partial H_N}{\partial q_i}\right) = \frac{\partial\rho_N}{\partial t} + \big\{\rho_N, H_N\big\} = 0.} \tag{II.10b}$$

---

[26] 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

$$\frac{\partial\dot{\mathfrak{q}}_i}{\partial q_i} = \frac{\partial}{\partial q_i}\frac{\partial H_N}{\partial p_i} = \frac{\partial}{\partial p_i}\frac{\partial H_N}{\partial q_i} = -\frac{\partial\dot{\mathfrak{p}}_i}{\partial p_i}.$$

---

[v] J. Liouville, 1809–1882

Introducing the *Liouville operator* (or *Liouvillian*) $\mathcal{L}$ defined by[(27)]

$$\mathrm{i}\mathcal{L} \equiv \sum_{i=1}^{3N} \left(\frac{\partial H_N}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H_N}{\partial q_i} \frac{\partial}{\partial p_i} \right) = \{\, \cdot\, , H_N \}, \tag{II.11}$$

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

$$\boxed{\frac{\partial \rho_N}{\partial t} + \mathrm{i}\mathcal{L}\rho_N = 0.} \tag{II.12}$$

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

$$\rho_N(t, \{q_i\}, \{p_i\}) = \mathrm{e}^{-\mathrm{i}\mathcal{L}t} \rho_N\big(t\!=\!0, \{q_i\}, \{p_i\}\big), \tag{II.13}$$

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

Remarks:

∗ An equivalent formulation of the Liouville equation, which follows from Eq. (II.10a) under consideration from the Hamilton equations, which yield $\boldsymbol{\nabla} \cdot \mathbf{u} = 0$ (see the identity in footnote 26), is

$$\frac{\partial \rho_N}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}\rho_N = \frac{\mathrm{D}\rho_N}{\mathrm{D}t} = 0, \tag{II.14}$$

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

∗ Equations (II.10b), (II.12) or (II.14) represent the Eulerian[(w)] viewpoint on the Liouville equation. Alternatively, one can adopt the Lagrangian[(x)] 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:

The volume in phase space occupied by a collection of microstates for a system obeying the Hamilton equations of motion remains constant in time. (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.

∗ 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, then it is stationary. This is for instance the case at equilibrium[(28)], but not only.

[(27)] 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.

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

[(w)] L. Euler, 1707–1783 [(x)] J.-L. Lagrange, 1736–1813

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

$$\frac{\mathrm{d}O_N(t)}{\mathrm{d}t} = \sum_{i=1}^{3N}\left[\frac{\partial O_N}{\partial q_i}\,\dot{\mathfrak{q}}_i(t) + \frac{\partial O_N}{\partial p_i}\,\dot{\mathfrak{p}}_i(t)\right],$$

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

$$\boxed{\frac{\mathrm{d}O_N}{\mathrm{d}t} = \sum_{i=1}^{3N}\left(\frac{\partial O_N}{\partial q_i}\frac{\partial H_N}{\partial p_i} - \frac{\partial O_N}{\partial p_i}\frac{\partial H_N}{\partial q_i}\right) = \{O_N, H_N\} = \mathrm{i}\mathcal{L}O_N,} \tag{II.16}$$

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, the differential equation (II.16) can formally be integrated as

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

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 the initial conditions at $t = 0$ then reads

$$\begin{aligned}\langle O_N(t)\rangle_0 &= \int_\Gamma \rho_N\big(t{=}0,\{q_i\},\{p_i\}\big)\,O_N(t)\,\mathrm{d}^{6N}\mathcal{V}\\ &= \int_\Gamma \rho_N\big(t{=}0,\{q_i\},\{p_i\}\big)\,\mathrm{e}^{\mathrm{i}\mathcal{L}t}\,O_N(t{=}0)\,\mathrm{d}^{6N}\mathcal{V}.\end{aligned} \tag{II.18}$$

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

$$\langle O_N(t)\rangle_t = \int_\Gamma O_N\big(\{q_i\},\{p_i\}\big)\,\rho_N\big(t,\{q_i\},\{p_i\}\big)\,\mathrm{d}^{6N}\mathcal{V}.$$

Using Eq. (II.13), this becomes

$$\langle O_N(t)\rangle_t = \int_\Gamma O_N\big(\{q_i\},\{p_i\}\big)\,\mathrm{e}^{-\mathrm{i}\mathcal{L}t}\,\rho_N\big(t{=}0,\{q_i\},\{p_i\}\big)\,\mathrm{d}^{6N}\mathcal{V}. \tag{II.19}$$

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.

This 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.

Remark: More generally, the identity

$$\int_\Gamma g^*(\{q_i\},\{p_i\})\mathcal{L}h(\{q_i\},\{p_i\})\,\mathrm{d}^{6N}\mathcal{V} = \int_\Gamma (\mathcal{L}g)^*(\{q_i\},\{p_i\})h(\{q_i\},\{p_i\})\,\mathrm{d}^{6N}\mathcal{V},$$

holds for every pair of phase-space functions $g(\{q_i\},\{p_i\})$ and $h(\{q_i\},\{p_i\})$ which 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 $\mathrm{e}^{\pm\mathrm{i}t\mathcal{L}}$, which govern the evolution of the density ρ_N or of observables $O_N(t)$, are *unitary* for this product, i.e. $\langle g, \mathrm{e}^{-\mathrm{i}t\mathcal{L}}h\rangle = \langle \mathrm{e}^{\mathrm{i}t\mathcal{L}}g, h\rangle$ or equivalently:

$$\int_\Gamma g(\{q_i\},\{p_i\})\,\mathrm{e}^{-\mathrm{i}t\mathcal{L}}h(\{q_i\},\{p_i\})\,\mathrm{d}^{6N}\mathcal{V} = \int_\Gamma \Big[\mathrm{e}^{\mathrm{i}t\mathcal{L}}g(\{q_i\},\{p_i\})\Big]h(\{q_i\},\{p_i\})\,\mathrm{d}^{6N}\mathcal{V}. \tag{II.20}$$

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 is the union—if one wants to be precise, the direct sum—of the individual N-particle phase spaces for every possible value of N, i.e. for $N \in \mathbb{N}$.(29) 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

$$\pi_N = \int \tilde{\rho}_N(t,\{q_i\},\{p_i\})\,\mathrm{d}^{6N}\mathcal{V}$$

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

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

II.3 Probabilistic description of quantum mechanical systems

In the “classical” formalism of quantum mechanics, the state of a system—if it is exactly known, in which case it is referred to as a *pure state*—is described by a normalized vector $|\Psi\rangle$ of a Hilbert(y) space $\mathscr{H}$. In turn, observables are modeled by Hermitian operators $\hat{O}(t)$ on $\mathscr{H}$.

In this section, we recall the basics of the formalism for the description of macroscopic quantum systems based on the density operator (Sec. II.3.1). We then discuss the time evolution of the latter (Sec. II.3.2) as well as that of the expectation values of observables (Sec. II.3.3). Eventually, we consider the case in which the Hamilton operator governing the evolution of the system can be split into two terms, namely a time-independent one and a time-dependent “perturbation” that is switched on at some initial instant (Sec. II.3.4).

II.3.1 Randomness in quantum mechanical systems

II.3.1 a Randomness in pure states

Experimentally, this pure state is entirely determined by the results of measurements of quantities associated with the operators of a complete set of commuting observables, where the latter are Hermitian linear operators on $\mathscr{H}$.

In quantum mechanics, the result of a measurement performed on a pure state $|\Psi\rangle$ might already be a random variable, in case $|\Psi\rangle$ is not an eigenstate of the observable $\hat{O}$ associated to the measured

---

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

(y) D. Hilbert, 1862–1943

quantity. In that case, repeated measurements will give the expectation value of the observable according to

$$\langle \hat{O} \rangle = \langle \Psi | \hat{O} | \Psi \rangle. \tag{II.21}$$

II.3.1 b Randomness in mixed states

In realistic cases, the microstate $|\Psi\rangle$ of a macroscopic system is not exactly determined. Instead, one only knows that the system can be with probability p_1 in microstate $|\Psi_1\rangle$, with probability p_2 in another microstate $|\Psi_2\rangle$, and so on, where the states $|\Psi_1\rangle, \dots, |\Psi_m\rangle, \dots$ are normalized, but not necessarily orthogonal, while the probabilities $p_1, \dots, p_m, \dots$ satisfy

$$p_m \geq 0 \quad \forall m \quad \text{and} \quad \sum_m p_m = 1.$$

One speaks then of a *statistical ensemble* or *statistical mixture* of states, or in short—and somewhat misleadingly—of a *mixed state*.

Remark: A mixed state should not be confused with a linear combination of states. In the latter case, the system is still in a pure state, corresponding to a single vector of the Hilbert space.

The expectation value of an observable for a system in a mixed state is the weighted sum of the expectation values in the pure states:

$$\langle \hat{O} \rangle = \sum_m p_m \langle \Psi_m | \hat{O} | \Psi_m \rangle. \tag{II.22a}$$

To express such expectation values in a convenient way, one introduces the *density operator* (also called *statistical operator* or *density matrix*)(30)

$$\hat{\rho} = \sum_m p_m |\Psi_m\rangle\langle\Psi_m| \,. \tag{II.22b}$$

One then has the identity

$$\langle \hat{O} \rangle = \mathrm{Tr}\big(\hat{\rho}\,\hat{O}\big), \tag{II.22c}$$

where Tr denotes the trace of an operator.

Using the matrix elements ρ_{ij} and O_{ij} of $\hat{\rho}$ and $\hat{O}$ in an arbitrary basis $\{|\phi_j\rangle\}$ of $\mathscr{H}$, as well as two decompositions of the identity, one finds at once

$$\langle \hat{O} \rangle = \sum_{i,j} \sum_m p_m \langle \Psi_m | \phi_i \rangle \langle \phi_i | \hat{O} | \phi_j \rangle \langle \phi_j | \Psi_m \rangle = \sum_{i,j} \rho_{ij} O_{ji} = \mathrm{Tr}\big(\hat{\rho}\,\hat{O}\big). \qquad \square$$

Remarks:

∗ The probabilities $p_1, \dots, p_m, \dots$ are clearly the eigenvalues of the density operator $\hat{\rho}$.

∗ The density-operator formalism easily accommodates the description of both mixed states and pure states. Thus, a microstate $|\Psi\rangle \in \mathscr{H}$ can be equivalently represented by the density operator $\hat{\rho} = |\Psi\rangle\langle\Psi|$ acting on the Hilbert space $\mathscr{H}$.

Properties of the density operator

1. $\hat{\rho}$ is Hermitian: $\hat{\rho}^\dagger = \hat{\rho}$.
   As a consequence, the expectation value of every observable is real.

   The proof follows from the hermiticity of $\hat{O}$ and the invariance of the trace under cyclic permutations: $\langle \hat{O} \rangle^* = \big[\mathrm{Tr}(\hat{\rho}\,\hat{O})\big]^* = \mathrm{Tr}\big(\hat{O}^\dagger \hat{\rho}^\dagger\big) = \mathrm{Tr}\big(\hat{O}\,\hat{\rho}\big) = \mathrm{Tr}\big(\hat{\rho}\,\hat{O}\big) = \langle \hat{O} \rangle$. $\square$

2. $\hat{\rho}$ is positive: $\forall\, |\phi\rangle \in \mathscr{H},\ \langle \phi | \hat{\rho} | \phi \rangle \geq 0$.
   Thus, the expectation value of every positive operator is a positive number.

---

(30) See e.g. Refs. [27] chapter E$_{\text{III}}$ or [3] § 5.

3. $\hat{\rho}$ is normalized to unity: $\operatorname{Tr}\hat{\rho} = 1$.
   This means that the expectation value of the identity equals 1.

The whole information on the system is encoded in its density operator $\hat{\rho}$. If one considers its matrix elements ρ_{ij} in an arbitrary basis $\{|\phi_j\rangle\}$, then each diagonal element, called *population*, ρ_{ii} is the probability to find the system in state $|\phi_i\rangle$.
The off-diagonal elements ρ_{ij} with $i \neq j$ are called *coherences*, and represent an information on the quantum-mechanical correlations between the possible states $|\phi_i\rangle$ and $|\phi_j\rangle$ of the system, which is absent in a classical description.

Remark: From the positivity of the density operator follows the positivity of each of its minors in a given basis, in particular the inequality $\rho_{ii}\rho_{jj} - \rho_{ij}\rho_{ji} \geq 0$. Since $\rho_{ji} = \rho_{ij}^*$, one sees that the coherence between two states can only be non-zero when the populations of these states do not vanish.

II.3.1 c Fluctuating number of particles

To account for possible fluctuations in the number of particles in a quantum-mechanical system, one introduces the *Fock*(z) *space*, that is the Hilbert space $\mathscr{H}$ defined as direct sum of the Hilbert spaces $\mathscr{H}_N$ corresponding to the N-particle problems, including the one-dimensional space $\mathscr{H}_0$ spanned by the *vacuum state* $|0\rangle$ describing the absence of particles.

The density operator $\hat{\rho}$ then simply acts on this Fock space and allows one to compute the expectation value of an observable—represented as an Hermitian operator on $\mathscr{H}$—through the usual formula (II.22c).

II.3.2 Time evolution of the density operator

Consider a macroscopic system, with the (possibly time-dependent) Hamilton operator $\hat{H}(t)$. Starting from the Schrödinger(aa) equation

$$i\hbar\frac{\partial|\Psi(t)\rangle}{\partial t} = \hat{H}(t)|\Psi(t)\rangle, \tag{II.23}$$

which holds for every pure state $|\Psi_m\rangle$ in which a statistical mixture can be, one finds that the time evolution of the density operator $\hat{\rho}(t)$ is governed by the *Liouville–von Neumann*(ab) *equation*

$$\boxed{\frac{\partial\hat{\rho}(t)}{\partial t} = \frac{1}{i\hbar}\big[\hat{H}(t),\hat{\rho}(t)\big],} \tag{II.24}$$

where the square brackets denote the commutator of two operators.

Considering the Hermitian conjugate equation to the Schrödinger equation, one finds

$$\frac{\partial\hat{\rho}(t)}{\partial t} = \sum_m p_m\left(\frac{\partial|\Psi_m\rangle}{\partial t}\langle\Psi_m| + |\Psi_m\rangle\frac{\partial\langle\Psi_m|}{\partial t}\right) = \sum_m p_m\left(\frac{1}{i\hbar}\hat{H}|\Psi_m\rangle\langle\Psi_m| + \frac{1}{-i\hbar}|\Psi_m\rangle\langle\Psi_m|\hat{H}\right),$$

which can be recast as Eq. (II.24). □

The solution of this differential equation for a given initial condition $\hat{\rho}(t_0)$ at some time $t = t_0$ can be expressed in terms of the *time-evolution operator* $\hat{\mathcal{U}}(t,t_0)$.[31] Recall that the latter evolves pure states of the system—described as vectors of $\mathscr{H}(t)$—between the initial time t_0 and time t

$$|\Psi(t)\rangle = \hat{\mathcal{U}}(t,t_0)|\Psi(t_0)\rangle. \tag{II.25a}$$

As such, the time-evolution operator is solution to the first-order differential equation

$$i\hbar\frac{\partial}{\partial t}\hat{\mathcal{U}}(t,t_0) = \hat{H}(t)\hat{\mathcal{U}}(t,t_0), \tag{II.25b}$$

---

[31] See e.g. Ref. [27] chapter FIII.

(z) V. A. Fock (or Fok), 1898–1974 (aa) E. Schrödinger, 1887–1961 (ab) J. von Neumann, 1903–1957 with the initial condition

$$\hat{\mathcal{U}}(t=t_0,t_0) = \hat{\mathbb{1}}. \tag{II.25c}$$

One then easily checks(32) that the solution to Eq. (II.24) is given by

$$\hat{\rho}(t) = \hat{\mathcal{U}}(t,t_0)\,\hat{\rho}(t_0)\,\hat{\mathcal{U}}(t_0,t), \tag{II.26}$$

where $\hat{\mathcal{U}}(t_0,t) = \hat{\mathcal{U}}(t,t_0)^{-1} = \hat{\mathcal{U}}(t,t_0)^\dagger$.

> This follows from differentiating this expression and using the Hermitian conjugate equation to Eq. (II.25b). □

Introducing the *Liouville operator* $\hat{\hat{\mathcal{L}}}(t)$ (or *superoperator*, since the vectors it acts upon are the operators on the Hilbert space $\mathscr{H}$) defined by

$$\mathrm{i}\hat{\hat{\mathcal{L}}}(t) \equiv \frac{1}{\mathrm{i}\hbar}\big[\,\cdot\,,\hat{H}(t)\big], \tag{II.27}$$

the Liouville–von Neumann equation takes the form

$$\boxed{\frac{\partial\hat{\rho}(t)}{\partial t} = -\mathrm{i}\hat{\hat{\mathcal{L}}}(t)\,\hat{\rho}(t),} \tag{II.28}$$

formally analogous to the classical Liouville equation (II.12).

II.3.2 a Isolated systems

If the system under consideration is isolated, its Hamilton operator $\hat{H}$ is actually time independent, so that the equation governing the time-evolution operator is readily integrated, yielding

$$\hat{\mathcal{U}}(t,t_0) = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}/\hbar}, \tag{II.29}$$

so that relation (II.26) becomes

$$\hat{\rho}(t) = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}/\hbar}\,\hat{\rho}(t_0)\,\mathrm{e}^{\mathrm{i}(t-t_0)\hat{H}/\hbar}, \tag{II.30}$$

with $\hat{\rho}(t_0)$ the "initial" density operator at $t = t_0$.

Denoting momentarily by $\{|\phi_j\rangle\}$ the orthonormal basis formed by the eigenstates of the Hamilton operator, with respective eigenvalues ϵ_j, and by ρ_{ij} the matrix elements of the density operator in this basis, Eq. (II.30) reads

$$\rho_{ii}(t) = \rho_{ii}(t_0), \qquad \rho_{ij}(t) = \rho_{ij}(t_0)\,\mathrm{e}^{-\mathrm{i}(\epsilon_i-\epsilon_j)(t-t_0)/\hbar} \;\text{ for } i\neq j. \tag{II.31}$$

That is, the populations in the energy eigenbasis do not evolve with time, while the corresponding coherences oscillate with the respective Bohr frequencies.

In turn, the Liouville superoperator is also time-independent, and the Liouville–von Neumann equation (II.28) can formally be integrated as

$$\hat{\rho}(t) = \mathrm{e}^{-\mathrm{i}(t-t_0)\hat{\hat{\mathcal{L}}}}\,\hat{\rho}(t_0), \tag{II.32}$$

which parallels Eq. (II.13) in the classical case and is totally equivalent to Eq. (II.30).

II.3.2 b Time-dependent Hamiltonian

When the Hamilton operator $\hat{H}$ depends on time, the corresponding time-evolution operator $\hat{\mathcal{U}}(t,t')$ is given by the *Dyson*(ac) *series*

$$\hat{\mathcal{U}}(t,t') = \hat{\mathbb{1}} - \frac{\mathrm{i}}{\hbar}\int_{t'}^{t}\hat{H}(t_1)\,\mathrm{d}t_1 + \left(\frac{-\mathrm{i}}{\hbar}\right)^2\int_{t'}^{t}\left[\int_{t'}^{t_1}\hat{H}(t_1)\hat{H}(t_2)\,\mathrm{d}t_2\right]\mathrm{d}t_1 + \cdots. \tag{II.33a}$$

---

(32) Note that this form can also be deduced from definition (II.22b) and equation (II.25a).

(ac) F. Dyson, born 1923

If t is a later time than t', then the time arguments on the right hand side obey $t' \leq \cdots \leq t_2 \leq t_1 \leq t$, i.e. the latest time argument is leftmost, the second latest is second from the left, and so on. Accordingly, one may write

$$\hat{\mathcal{U}}(t,t') = \mathcal{T} \exp\left[-\frac{\mathrm{i}}{\hbar}\int_{t'}^{t} \hat{H}(u)\,\mathrm{d}u\right] \quad \text{for } t \geq t', \tag{II.33b}$$

with $\mathcal{T}$ the Dyson time-ordering operator, which orders each product of (Hamilton) operators in the expansion of the exponential with growing time arguments from the right to the left.

On the other hand, if $t < t'$ in the Dyson series (II.33a), then the time arguments are actually ordered the other way round: $t \leq t_1 \leq t_2 \leq \cdots \leq t'$, i.e. the latest one is rightmost. Thus, one now writes

$$\hat{\mathcal{U}}(t,t') = \mathcal{T}^{\mathrm{a}} \exp\left[\frac{\mathrm{i}}{\hbar}\int_{t}^{t'} \hat{H}(u)\,\mathrm{d}u\right] \quad \text{for } t \leq t', \tag{II.33c}$$

with $\mathcal{T}^{\mathrm{a}}$ the anti-chronological time-ordering operator, which orders each product of operators in the expansion of the exponential with growing time arguments from the left to the right.

Armed with these results, we may now express the density operator (II.26). Assuming—since this is the case we shall in practice consider—that the instant t_0 at which the boundary condition is fixed is really the initial time, one has

$$\hat{\rho}(t) = \mathcal{T} \exp\left[-\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \hat{H}(u)\,\mathrm{d}u\right]\hat{\rho}(t_0)\,\mathcal{T}^{\mathrm{a}} \exp\left[\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \hat{H}(u)\,\mathrm{d}u\right] \quad \text{for } t \geq t_0. \tag{II.34}$$

Remark: If the values of the Hamilton operator at two arbitrary different times t', t'' always commute, $\big[\hat{H}(t'), \hat{H}(t'')\big] = 0$, then the time-ordering operators (chronological or anti-chronological) are not necessary.

II.3.3 Time evolution of observables and of their expectation values

Consider now an observable $\hat{O}$ for a time-dependent quantum-mechanical system with Hilbert space $\mathscr{H}$. As is well known, there are two approaches to compute its expectation value (II.21) in a pure state, namely either considering that the latter is described by a time-evolving vector $|\Psi(t)\rangle$ of $\mathscr{H}$ obeying the Schrödinger equation, or by letting the state-vector remain constant, while the time evolution is entirely attached to the observable.(33) Here we extend this dual point of view to the computation of expectations values of observables in statistical mixtures of states, described by a density operator $\hat{\rho}$.

II.3.3 a Schrödinger picture

Let $\hat{O}(t)$ denote an observable of the system, where for the sake of generality we allowed for an explicit time dependence of the operator. In the Schrödinger picture, the density operator $\hat{\rho}(t)$ evolves with time according to the Liouville–von Neumann equation, while $\hat{O}(t)$ is what it is—in particular, $\hat{O}$ remains constant in time if it has no explicit dependence on t.

Using the general formula (II.22c) under consideration of Eq. (II.26), the expectation value of the observable at time t then reads

$$\langle\hat{O}(t)\rangle = \mathrm{Tr}\big[\hat{\rho}(t)\,\hat{O}(t)\big] = \mathrm{Tr}\Big[\hat{\mathcal{U}}(t,t_0)\,\hat{\rho}(t_0)\,\hat{\mathcal{U}}(t_0,t)\,\hat{O}(t)\Big]. \tag{II.35}$$

In the following, (pure) states or operators without subscript will automatically refer to their representation in the Schrödinger picture.

II.3.3 b Heisenberg picture

In the Heisenberg(ad) picture, the state of the system is kept fixed at its value at a given reference time t_0: for a pure state, $|\Psi\rangle_{\mathrm{H}} \equiv |\Psi(t_0)\rangle$; for a statistical mixture of states, $\hat{\rho}_{\mathrm{H}} \equiv \hat{\rho}(t_0)$.

---

(33) See e.g. Refs. [27] chapter $\mathrm{G_{III}}$ or [28] § 13.

(ad) W. Heisenberg, 1901–1976

In turn, observables are represented by operators $\hat{O}_{\rm H}(t)$ related to those in the Schrödinger representation by

$$\hat{O}_{\rm H}(t) \equiv \hat{\mathcal{U}}(t_0,t)\,\hat{O}(t)\,\hat{\mathcal{U}}(t,t_0), \tag{II.36}$$

which ensures the identity ${}_{\rm H}\langle\Psi|\,\hat{O}_{\rm H}(t)\,|\Psi\rangle_{\rm H} = \langle\Psi(t)|\,\hat{O}(t)\,|\Psi(t)\rangle$ for the expectation value of the observable in a pure state in either picture. The operator $\hat{O}_{\rm H}(t)$ obeys the *Heisenberg equation*

$$\frac{{\rm d}\hat{O}_{\rm H}(t)}{{\rm d}t} = \frac{1}{{\rm i}\hbar}\big[\hat{O}_{\rm H}(t),\hat{H}_{\rm H}\big] + \left(\frac{\partial\hat{O}(t)}{\partial t}\right)_{\rm H}, \tag{II.37}$$

where the second term on the right-hand side vanishes when $\hat{O}$ has no explicit time dependence. Under this assumption and using the Liouville operator (II.27) computed with $\hat{H}_{\rm H}$, this equation becomes

$$\boxed{\frac{{\rm d}\hat{O}_{\rm H}(t)}{{\rm d}t} = {\rm i}\hat{\hat{\mathcal{L}}}_{\rm H}(t)\hat{O}_{\rm H}(t),} \tag{II.38}$$

to be compared with Eq. (II.16) in the classical case.

If the Hamiltonian $\hat{H}$ is time independent, then $\hat{H}_{\rm H} = \hat{H}$ and the Liouville operator ${\rm i}\hat{\hat{\mathcal{L}}}_{\rm H}$ is time-independent. Equation (II.38) is straightforwardly integrated as

$$\hat{O}_{\rm H}(t) = {\rm e}^{{\rm i}(t-t_0)\hat{\hat{\mathcal{L}}}_{\rm H}}\,\hat{O}_{\rm H}(t_0) = {\rm e}^{{\rm i}(t-t_0)\hat{H}/\hbar}\,\hat{O}_{\rm H}(t_0)\,{\rm e}^{-{\rm i}(t-t_0)\hat{H}/\hbar}, \tag{II.39}$$

which under consideration of Eq. (II.29) is exactly equivalent to relation (II.36) since $\hat{O}_{\rm H}(t_0) = \hat{O}(t_0)$.

Remarks:

∗ The evolution equations for the density operator and for observables, Eqs. (II.28) and (II.38), differ by a minus sign, which shows that the former, despite its possessing some of the "good" properties (hermiticity), is not an observable.(34)

∗ The Liouville operator is sometimes defined with a different convention from Eq. (II.27), namely as

$$\hat{\hat{\mathcal{L}}}' \equiv -\big[\,\cdot\,,\hat{H}\big], \tag{II.40}$$

without the factor $1/\hbar$. The advantage of this alternative definition is that the evolution equation of observables then becomes

$${\rm i}\hbar\frac{{\rm d}\hat{O}(t)}{{\rm d}t} = -\hat{\hat{\mathcal{L}}}'\hat{O}(t), \tag{II.41}$$

instead of Eq. (II.37). It is thus now quite similar—up to the minus sign—to the Schrödinger equation (II.23): the Liouville superoperator plays the role of the (negative of the) Hamilton operator, while the role of the kets of the Hilbert space $\mathscr{H}$ is taken by the operators on $\mathscr{H}$.
The drawback of this definition is that one loses the usual recipe for going from the quantum-mechanical to the classical case by replacing Poisson brackets $\{\,\cdot\,,\,\cdot\,\}$ with $[\,\cdot\,,\,\cdot\,]/{\rm i}\hbar$.

II.3.3 c Time evolution of expectation values in statistical ensembles

Multiplying both sides of Eq. (II.36) with the density operator $\hat{\rho}_{\rm H} = \hat{\rho}(t_0)$ and taking the trace yields the expectation value

$$\big\langle\hat{O}_{\rm H}(t)\big\rangle = {\rm Tr}\big[\hat{\rho}_{\rm H}\,\hat{O}_{\rm H}(t)\big] = {\rm Tr}\Big[\hat{\rho}(t_0)\,\hat{\mathcal{U}}(t_0,t)\,\hat{O}(t)\,\hat{\mathcal{U}}(t,t_0)\Big]. \tag{II.42}$$

Using the invariance of the trace under cyclic permutations, this is clearly the same as Eq. (II.35) in the Schrödinger picture, i.e. $\big\langle\hat{O}_{\rm H}(t)\big\rangle = \big\langle\hat{O}(t)\big\rangle$.

---

(34) A further hint to this difference is given by the fact that, in the interaction picture, observables evolve with the "unperturbed" Hamiltonian $\hat{H}_0$ while the density operator evolves with the perturbation $\hat{W}(t)$, see § II.3.4.

Using either picture—which we do by not specifying whether the time dependence is attached to $\hat{\rho}$ or $\hat{O}$ and by dropping the subscript H—one finds that the time derivative of the expectation value of an observable with no explicit time dependence obeys the equation

$$\frac{\mathrm{d}\langle\hat{O}\rangle}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\big[\mathrm{Tr}\big(\hat{\rho}\,\hat{O}\big)\big] = \frac{1}{\mathrm{i}\hbar}\,\mathrm{Tr}\big(\big[\hat{H},\hat{\rho}\big]\hat{O}\big) = \frac{1}{\mathrm{i}\hbar}\,\mathrm{Tr}\big(\hat{\rho}\big[\hat{O},\hat{H}\big]\big). \tag{II.43}$$

The third term is proven by attaching the time dependence to $\hat{\rho}$ (Schrödinger picture) and using the Liouville–von Neumann equation (II.24); The fourth term follows from differentiating $\hat{O}_{\mathrm{H}}(t)$ and inserting the Heisenberg equation (II.37) without the partial-derivative term. The equivalence between the third and fourth terms is easily checked and follows from the invariance of the trace under cyclic permutations.

II.3.3 d Time contour

Invoking the invariance of the trace under cyclic permutations, Eqs. (II.35) or (II.42) can be rewritten as

$$\big\langle\hat{O}(t)\big\rangle = \big\langle\hat{O}_{\mathrm{H}}(t)\big\rangle = \mathrm{Tr}\Big[\hat{\mathcal{U}}(t_0,t)\,\hat{O}(t)\,\hat{\mathcal{U}}(t,t_0)\,\hat{\rho}(t_0)\Big]. \tag{II.44}$$

Reading the operator product in the argument of the trace from right to left, one begins at an "initial" time t_0, evolves until time t—which might actually be prior to t_0, yet in practice we shall always take $t \geq t_0$—; then at time t the operator $\hat{O}$ acts. Eventually, the system evolves "back" from t to t_0.

The corresponding path "in time" is pictured in Fig. II.1 as a *time contour* or *Keldysh*(ae) *contour* (in the complex plane of the variable t) along the real axis, starting at t_0, going forward until t, then back to t_0. Note that for the readability of the figure both forward and backward parts of the contour have been slightly displaced away from the real axis, although they in fact lie on it.

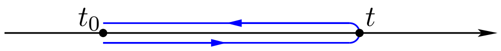

Figure II.1

II.3.4 Time evolution of perturbed systems

An often encountered scenario consists in the following setup:

- Until some "initial" time t_0, the quantum-mechanical system under consideration is governed by a time-independent Hamilton operator $\hat{H}_0$—whose eigenvalues and eigenstates are often assumed to be known—and the (macro)state of the system at t_0 is known: $\hat{\rho}(t_0)$.
- At t_0 a time-dependent "perturbation" is turned on, corresponding to an extra term $\hat{W}(t)$ in the Hamiltonian, resulting in the total Hamilton operator

$$\hat{H}(t) = \hat{H}_0 + \hat{W}(t) \quad \text{for } t \geq t_0. \tag{II.45}$$

The goal is then to compute the evolution of the system—in particular of $\hat{\rho}(t)$—or the expectation value of some observable(s) at $t > t_0$.

In that case, it is fruitful to work in the *interaction* or *Dirac picture*, introducing on one hand vectors of the Hilbert space $\mathscr{H}$

$$|\Psi(t)\rangle_{\mathrm{I}} \equiv \mathrm{e}^{\mathrm{i}(t-t_0)\hat{H}_0/\hbar}\,|\Psi(t)\rangle = \hat{\mathcal{U}}_0(t,t_0)^\dagger\,|\Psi(t)\rangle \tag{II.46a}$$

and on the other hand operators on $\mathscr{H}$

$$\hat{O}_{\mathrm{I}}(t) \equiv \mathrm{e}^{\mathrm{i}(t-t_0)\hat{H}_0/\hbar}\,\hat{O}(t)\,\mathrm{e}^{-\mathrm{i}(t-t_0)\hat{H}_0/\hbar} = \hat{\mathcal{U}}_0(t,t_0)^\dagger\,\hat{O}(t)\,\hat{\mathcal{U}}_0(t,t_0). \tag{II.46b}$$

(ae) L. V. Keldysh, born 1931

In these definitions, $\hat{\mathcal{U}}_0(t,t_0)$ denotes the time-evolution operator associated with $\hat{H}_0$ alone, here given by Eq. (II.29).

One then quickly finds that pure states $|\Psi(t)\rangle_{\rm I}$ evolve according to

$$\mathrm{i}\hbar\frac{\partial|\Psi(t)\rangle_{\rm I}}{\partial t} = \hat{W}_{\rm I}(t)|\Psi(t)\rangle_{\rm I}, \tag{II.47a}$$

i.e. under the influence of the "perturbation" $\hat{W}(t)$ alone, while accordingly the density operator in interaction representation $\hat{\rho}_{\rm I}(t)$ is governed by

$$\frac{\partial\hat{\rho}_{\rm I}(t)}{\partial t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{W}_{\rm I}(t),\hat{\rho}_{\rm I}(t)\big]. \tag{II.47b}$$

Eventually, definition (II.46b) leads to

$$\frac{\mathrm{d}\hat{O}_{\rm I}(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{O}_{\rm I}(t),\hat{H}_0\big] + \left(\frac{\partial\hat{O}(t)}{\partial t}\right)_{\rm I}. \tag{II.47c}$$

Remarks:

∗ The "initial time" t_0 is also often taken to lie in the infinitely remote past, $t_0\to-\infty$. This in particular allows one to consider if need be that the perturbation $\hat{W}(t)$ is turned on "adiabatically", i.e. slowly enough.(35)

∗ Irrespective of whether t_0 is finite or not, it is often assumed that the system is at time t_0 in a state of thermodynamic equilibrium with respect to the Hamiltonian $\hat{H}_0$. For instance, the system is in equilibrium with a thermostat at temperature T, leading to the canonical equilibrium distribution

$$\hat{\rho}(t_0) = \frac{1}{Z(\beta)}\mathrm{e}^{-\beta\hat{H}_0} \quad \text{with} \quad Z(\beta)\equiv \mathrm{Tr}\,\mathrm{e}^{-\beta\hat{H}_0}$$

and $\beta\equiv 1/k_BT$. The exponential term entering this density operator can be written in term of the time-evolution operator associated with $\hat{H}_0$

$$\mathrm{e}^{-\beta\hat{H}_0} = \hat{\mathcal{U}}_0(t_0,t_0-\mathrm{i}\hbar\beta),$$

corresponding formally to an evolution in imaginary time, from $t_0-\mathrm{i}\hbar\beta$ to t_0. Accordingly, the expectation value (II.44) becomes

$$\big\langle\hat{O}(t)\big\rangle = \big\langle\hat{O}_{\rm H}(t)\big\rangle = \frac{1}{Z(\beta)}\mathrm{Tr}\Big[\hat{\mathcal{U}}(t_0,t)\,\hat{O}(t)\,\hat{\mathcal{U}}(t,t_0)\,\hat{\mathcal{U}}_0(t_0,t_0-\mathrm{i}\hbar\beta)\Big], \tag{II.48}$$

where one has to pay attention that the system does not evolve with the same Hamiltonian "before" t_0 and afterwards, resulting in different time-evolution operators $\hat{\mathcal{U}}_0$, $\hat{\mathcal{U}}$. Corresponding to the time sequence, read from right to left, one can associate the Keldysh contour pictured in Fig. II.2, with a first part parallel to the imaginary axis.

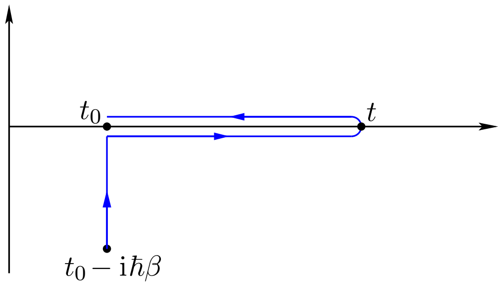

Figure II.2

---

[35] The corresponding *adiabaticity*, which takes here a different meaning from that of thermodynamics (absence of heat exchange), is discussed for instance in Refs. [29] chapter XVII § 7–14 or [30] chapter 10.

II.4 Statistiscal entropy

The probabilistic approach advocated in Sec. II.1 arises from a lack of knowledge on the microscopic state of the system. This raises the question of how much information is missing, if only discrete probabilities $\{p_k\}$ or a probability density ρ_N are known.

Intuitively, the missing information is "small" when the probability distribution takes significant values for only a few microstates, while it is larger in case the distribution extends over many states. To make sense of this intuition, a measure of (missing) information in probability theory, the statistical entropy, is introduced in Sec. II.4.1. This measure is then applied to the probability distributions that describe quantum-mechanical (Sec. II.4.2) or classical (Sec. II.4.3) systems.

II.4.1 Statistical entropy in information theory

For the sake of brevity, we shall in the following consider only discrete probability distributions, with the exception of definition (II.50). The generalization to the case of continuous distributions is straightforward.

Consider M events $\{\omega_1,\dots,\omega_M\}$ with respective probabilities $p_1,\dots,p_M$. In order to quantify the "uncertainty" (or "missing information", "ignorance") corresponding to the use of this probability distribution, C. Shannon(af) [31] has introduced the *statistical entropy*

$$\boxed{S_{\rm stat}(p_1,\dots,p_M) \equiv -k\sum_{m=1}^{M} p_m \ln p_m,} \tag{II.49}$$

with k an arbitrary positive constant. $S_{\rm stat}(p_1,\dots,p_M)$ is thus exactly equal to the average value of $-k\ln p_m$.

In information theory one usually takes $k=1/\ln 2$, so that

$$S_{\rm stat}(p_1,\dots,p_M) = \left\langle \log_2 \frac{1}{p_m} \right\rangle,$$

and the statistical entropy is dimensionless.(36)

> The quantity $-\log_2 p_m$ is the *information content* or *self-information* associated with the event ω_m (for Shannon, the events were possible meaningful messages in a given language). Thus, the occurrence of a message ω_m with low probability is more informative, more surprising than that of a more probable message.

In the case of a continuous probability distribution, described by a probability density $p(x)$, the statistical entropy is given by

$$\boxed{S_{\rm stat}\big(p(x)\big) \equiv -k\int p(x)\ln p(x)\,{\rm d}x.} \tag{II.50}$$

$S_{\rm stat}$ is here also equal to the expectation value of the logarithm of the distribution.

II.4.2 Statistical entropy of a quantum-mechanical system

The notion of statistical entropy can now be applied to the probability distributions introduced in the description of macroscopic quantum-mechanical systems.

---

(36) In this context, the unit of the statistical entropy is the bit.

(af) C. Shannon, 1916–2001

Let $\hat{\rho}$ denote the density operator associated with a quantum-mechanical mixture of states, with eigenvalues $p_1, \ldots, p_M$. The corresponding statistical entropy is defined as

$$\boxed{S(\hat{\rho}) \equiv -k_B \sum_m p_m \ln p_m = -k_B \operatorname{Tr}\big(\hat{\rho} \ln \hat{\rho}\big)}, \tag{II.51}$$

where the second identity follows at once in a basis in which $\hat{\rho}$ is diagonal. Here k_B is the Boltzmann[(ag)] constant, $k_B = 1.38 \times 10^{-23}$ J.K^{-1}.

Remark: The first identity matches the definition of the entropy by Gibbs (1878), the second corresponds to the von Neumann entropy (1927).

II.4.3 Statistical entropy of a classical system

The missing information associated with the use of a probability density on phase space, as is the case for macroscropic classical systems, can similarly be quantified with the statistical entropy. For a classical system with fluctuating particle number N, the statistical entropy reads

$$S(\rho) = -k_B \sum_N \int \tilde{\rho}_N(\{q_i\},\{p_i\}) \ln \tilde{\rho}_N(\{q_i\},\{p_i\})\, \mathrm{d}^{6N}\mathcal{V}. \tag{II.52}$$

This statistical entropy globally enjoys the same properties as its quantum-mechanical counterpart (II.51), up to an important exception. Since a probability density is not bounded from above to 1, $S(\rho)$ can be negative, and actually has no minimum. One can find phase-space densities such that $S(\rho)$ is as negative as one wants, i.e. $S(\rho)$ can go to $-\infty$. Such densities actually require a simultaneous knowledge of both positions and momenta of particles, which violate the Heisenberg uncertainty relations of quantum mechanics.

Bibliography for Chapter II

- Kubo *et al.*, *Statistical physics I: Equilibrium statistical physics* [32], chapter 1.3.
- Landau & Lifshitz, *Course of theoretical physics Vol. 1: Mechanics* [26], chapter VII § 46 and *Vol. 5: Statistical physics, part 1* [3], chapter I.
- Le Bellac, Mortessagne & Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* [9], chapter 2.1–2.2.
- Pottier, *Nonequilibrium statistical physics* [6], chapter 3.
- Zwanzig, *Nonequilibrium statistical mechanics* [7], chapters 2.1 & 6.1.

[(ag)] L. Boltzmann, 1844–1906

Chapter III

Reduced classical phase-space densities and their evolution

The description of a classical N-body systems relies on the N-particle phase space Γ: the evolution of the system is represented by a trajectory $(\{\mathfrak{q}_i(t)\},\{\mathfrak{p}_i(t)\})$ solution of the Hamilton equations (II.1)—corresponding to the sequence of microstates occupied by the system at successive instants—, while observables are functions $O_N\big(\{q_i\},\{p_i\}\big)$ on Γ. If, as is generally the case for macroscopic systems, the system state is only known on a statistical basis, one also needs the probability density $\rho_N\big(t,\{q_i\},\{p_i\}\big)$ on Γ, whose time evolution is governed by the Liouville equation (II.10b), see Sec. II.2.

In the large majority of cases, the macroscopic quantities of interest are a measure for the expectation value of microscopic observables that only involve few particles, often only one or two. For instance, the temperature is related to the average kinetic energy of the particles in the system, which involves a single-particle observable. To compute the expectation values of such observables, it is worthwhile to introduce *reduced phase-space densities*, which integrate out the irrelevant degrees of freedom (Sec. III.1). From the Liouville equation for the Γ-space probability density ρ_N, one can then derive the equations of motion governing the dynamics of the reduced densities (Sec. III.2).

In this chapter and the following one, the positions and momenta of the N particles will be denoted by $\vec{r}_i$ and $\vec{p}_i$ respectively, where $i=1,\dots,N$ labels the particle. Accordingly, the argument of a function on the Γ-space of the N particles will be written $\{\vec{r}_i\},\{\vec{p}_i\}$ or $\vec{r}_1,\vec{p}_1,\dots,\vec{r}_N,\vec{p}_N$ instead of $\{q_j\},\{p_j\}$.

III.1 Reduced phase-space densities

Consider a system of N particles, which for simplicity are first assumed to be identical. Starting from the N-particle density $\rho_N(t,\vec{r}_1,\vec{p}_1,\dots,\vec{r}_N,\vec{p}_N)$ and integrating out the space and momentum coordinates of $N-1$ particles, one defines the *single-particle phase-space density*

$$f_1(t,\vec{r},\vec{p}) \equiv \alpha_{N,1}\int\rho_N(t,\vec{r},\vec{p},\vec{r}_2,\vec{p}_2,\dots,\vec{r}_N,\vec{p}_N)\,\mathrm{d}^{6(N-1)}\mathcal{V} \tag{III.1a}$$

where $\mathrm{d}^{6(N-1)}\mathcal{V}$ is the infinitesimal volume element in the subspace spanned by the variables $\vec{r}_2,\vec{p}_2,\dots,\vec{r}_N,\vec{p}_N$. The remaining single-particle phase space is often referred to as μ-space, so as to distinguish it from the Γ-space.

Similarly, one introduces the *two-particle phase-space density*

$$f_2(t,\vec{r},\vec{p},\vec{r}\,',\vec{p}\,') \equiv \alpha_{N,2}\int\rho_N(t,\vec{r},\vec{p},\vec{r}\,',\vec{p}\,',\vec{r}_3,\vec{p}_3,\dots,\vec{r}_N,\vec{p}_N)\,\mathrm{d}^{6(N-2)}\mathcal{V}. \tag{III.1b}$$

More generally, integrating out the positions and momenta of $N-k$ particles in the Γ-space density ρ_N leads to a k-particle phase-space density f_k.

Remark: For a system of identical—i.e. quantum-mechanically indistinguishable—particles, the density ρ_N should be symmetric under the exchange of any two particles, i.e. of the variables $\vec{r}_i$, $\vec{p}_i$

and $\vec{r}_j$, $\vec{p}_j$ for any i, j. Accordingly, which $N-k$ positions and momenta are being integrated out in the definition of the reduced k-particle density f_k is actually unimportant, and any choice leads to the same reduced density.

The definitions (III.1a) and (III.1b) involve constants $\alpha_{N,1}$, $\alpha_{N,2}$. These are often chosen so that the single- and two-particle phase-space densities are respectively normalized to

$$\int f_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p} = N, \tag{III.2a}$$

i.e. the total number of particles, and

$$\int f_2(t,\vec{r},\vec{p},\vec{r}\,',\vec{p}\,')\,\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}\;\mathrm{d}^3\vec{r}\,'\,\mathrm{d}^3\vec{p}\,' = N(N-1), \tag{III.2b}$$

that is, the total number of ordered pairs. More generally, one requires that the integral of f_k be the total number of ordered k-tuples. This requirement dictates for identical particles the choice $\alpha_{N,k} = 1/(2\pi\hbar)^{3k}$, independent of N. With these normalizations, $f_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{r}\,\mathrm{d}^3\vec{p}$ is the average number of particles in the infinitesimal volume element $\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}$ around the μ-space point $(\vec{r},\vec{p})$ at time t, while $f_2(t,\vec{r},\vec{p},\vec{r}\,',\vec{p}\,')\,\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}\;\mathrm{d}^3\vec{r}\,'\,\mathrm{d}^3\vec{p}\,'$ is the average number of ordered pairs with one particle in the volume element $\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}$ around $(\vec{r},\vec{p})$ and the other in $\mathrm{d}^3\vec{r}\,'\,\mathrm{d}^3\vec{p}\,'$ around $(\vec{r}\,',\vec{p}\,')$. The various f_k are thus number densities in the respective phase spaces, rather than probability densities.

Integrating the single-particle phase-space density (III.1a) normalized according to Eq. (III.2a) over momentum, one recovers the usual *particle number density* (in position space)

$$n(t,\vec{r}) \equiv \int f_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{p}. \tag{III.3a}$$

Remarks:

∗ Up to the normalization $[N!/(N-k)!$ instead of $1]$, the reduced phase-space densities f_k are marginal distributions (see Appendix B.4.1) of the Γ-space density ρ_N.

∗ Given the normalization adopted here, the densities f_k are dimensionful quantities. It is sometimes preferable to work with dimensionless densities, namely $\mathsf{f}_k \equiv (2\pi\hbar)^{3k} f_k$, which correspond to the simpler normalization constant $\alpha_{N,k} = 1$ for all k. Integrating f_1 over a spatial volume on which the system is homogeneous then yields the so-called *phase-space occupancy*—where "phase space" actually means momentum space.

The various results of this chapter, as e.g. the BBGKY hierarchy (III.14), still hold when replacing the reduced densities f_k by the dimensionless f_k, provided every momentum-space volume element $\mathrm{d}^3\vec{p}$ appearing in an integral is replaced by $\mathrm{d}^3\vec{p}/(2\pi\hbar)^3$. For instance,

$$n(t,\vec{r}) \equiv \int \mathsf{f}_1(t,\vec{r},\vec{p})\,\frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^3}. \tag{III.3b}$$

∗ The reduced densities f_k depend on the underlying statistical ensemble. Here, since the number of particles N is fixed, we are implicitly considering the canonical ensemble, yet this is just for the sake of simplicity.

In the "grand-canonical" case of a fluctuating particle number, integrating the phase-space density ρ defined in § II.2.5 over the irrelevant degrees of freedom yields the reduced k-particle density

$$f_k = \sum_{N=k}^{\infty} \pi_N f_{k|N\text{–}k}$$

where $f_{k|N\text{–}k}$ is the ("canonical") reduced k-particle distribution under the condition that there are exactly N particles in the system while π_N is the probability that this is the case. Accordingly, f_1 is then normalized to $\langle N\rangle$, f_2 to $\langle N(N-1)\rangle$, and so on.

The reader can check that for a classical ideal gas, $f_2(t,\vec{r},\vec{p},\vec{r}\,',\vec{p}\,') = f_1(t,\vec{r},\vec{p})f_1(t,\vec{r}\,',\vec{p}\,')$ in the grand-canonical case—which can be interpreted as signaling the absence of correlations between particles—, while the identity does not hold when the total particle number is fixed.

With the help of the μ-space density f_1, the expectation value of a single-particle observable $O_1(\vec{r},\vec{p})$ reads

$$\left\langle O_1(t)\right\rangle_t = \frac{\displaystyle\int O_1(\vec{r},\vec{p})\, f_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}}{\displaystyle\int f_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}}. \tag{III.4}$$

The expectation value of a two-particle observable can similarly be expressed as an average weighted by the two-particle density f_2.

III.2 Time evolution of the reduced phase-space densities

We now deduce from the Liouville equation (III.5) the equations of motion for the reduced phase-space densities, first for a system of particles in the absence of vector potential (Sec. III.2.2), then for charged particles in a external vector potential (Sec. III.2.3).

III.2.1 Description of the system

Consider identical pointlike particles of mass m, without internal degrees of freedom, whose positions and canonical momenta will be denoted by $\vec{r}_i$ and $\vec{p}_i$. For a system of N such particles, the Liouville equation (II.10b) can be recast as

$$\boxed{\frac{\partial\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big)}{\partial t} + \sum_{i=1}^{N}\dot{\vec{r}}_i\cdot\vec{\nabla}_{\vec{r}_i}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big) + \sum_{i=1}^{N}\dot{\vec{p}}_i\cdot\vec{\nabla}_{\vec{p}_i}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big) = 0,} \tag{III.5a}$$

where $\vec{\nabla}_{\vec{r}_i}$, $\vec{\nabla}_{\vec{p}_i}$ stand for the three-dimensional gradients whose coordinates are the derivatives with respect to the canonical variables $\vec{r}_i$, $\vec{p}_i$. In turn, $\dot{\vec{r}}_i$—i.e. the velocity $\vec{v}_i$ of particle i—and $\dot{\vec{p}}_i$ denote the time derivatives of the canonical variables. Invoking the Hamilton equations (II.1) and dropping the variables, the Liouville equation becomes

$$\frac{\partial\rho_N}{\partial t} + \sum_{i=1}^{N}\vec{\nabla}_{\vec{p}_i}H_N\cdot\vec{\nabla}_{\vec{r}_i}\rho_N - \sum_{i=1}^{N}\vec{\nabla}_{\vec{r}_i}H_N\cdot\vec{\nabla}_{\vec{p}_i}\rho_N = 0 \tag{III.5b}$$

with H_N the Hamilton function of the system, which we shall specify later in this section.

In both this chapter and the following we shall implicitly work in the "thermodynamic limit" of an infinitely many of particles occupying an infinitely large volume, thereby allowing for an infinitely large total energy of the system. As a result, both the Cartesian components of spatial and momentum variables can take values spanning the whole real axis from $-\infty$ to $+\infty$. Now, to ensure its normalizability [Eq. (II.3)], the probability density ρ_N must vanish when one of its phase-space variables goes to infinity: the partial differential equation (III.5a) is thus to be complemented with the boundary conditions

$$\lim_{|\vec{r}_i|\to\infty}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big) = 0 \quad\text{and}\quad \lim_{|\vec{p}_i|\to\infty}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big) = 0 \qquad \forall i\in\{1,\dots,N\}. \tag{III.6}$$

Accordingly, integrals of the type

$$\int\vec{\nabla}_{\vec{r}_i}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big)\,\mathrm{d}^3\vec{r}_i \quad\text{resp.}\quad \int\vec{\nabla}_{\vec{p}_i}\rho_N\big(t,\{\vec{r}_j\},\{\vec{p}_j\}\big)\,\mathrm{d}^3\vec{p}_i \tag{III.7}$$

over the whole allowed range for $\vec{r}_i$ resp. $\vec{p}_i$ identically vanish.

For the sake of completeness, let us briefly mention some of the necessary changes taking place if the spatial volume $\mathcal{V}$ occupied by the system is not infinite—as is e.g. the case of a gas enclosed in a fixed box. In short, the treatment becomes slightly more tedious.

First, the total number N of particles is then finite, and so is the total energy of the system. As a result, the (kinetic) energy of an individual particle cannot be infinitely large, which means that the range of allowed momentum values is finite: the behavior of ρ_N at the edges of the momentum range is no longer obvious, although it is natural to assume that ρ_N vanishes.

Since the spatial volume is finite, the variables $\vec{r}_i$ are also restricted to a finite range. It is then necessary to specify the boundary conditions at these spatial edges ("walls"). A useful choice is to assume that the boundaries behave like perfect mirrors for the particles, which are thus reflected elastically according to the Snell[ah]–Descartes[ai] law of reflection: this has the advantage of ensuring the conservation of kinetic energy; however such a particle-wall collision does not conserve (linear) momentum, which is unsatisfactory. The "good" approach is then to resort to a microscopic description of the walls themselves, which become part of the system under study.

Let us now specify the Hamilton function H_N of the system. For the sake of simplicity, we shall assume that the particles can only interact with each other pairwise, i.e. we discard possible genuine three-, four-, …, many-body interaction terms. These two-body interactions can be described in terms of an interaction potential W,(37) which will be assumed to depend on the inter-particle distance only.

We also allow for the possible presence of external potentials acting on the particles. Here one has to distinguish between two possibilities. Thus, a scalar potential $V(t,\vec{r})$—e.g. electrostatic, (Newtonian) gravitational, or a potential well enclosing particles in a specified volume—only constitutes a minor modification, since it does not affect the canonical momentum $\vec{p}_i$ conjugate to position. In contrast, any vector potential $\vec{A}(t,\vec{r})$ directly enters the canonical momentum, so that we shall have to discuss the gauge invariance of the evolution equations for the reduced phase-space densities.

In the absence of vector potential, the Hamilton function of the system is of the type

$$H_N = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} + \sum_{i=1}^{N} V(t,\vec{r}_i) + \sum_{1\le i<j\le N} W\big(|\vec{r}_i - \vec{r}_j|\big). \tag{III.8}$$

For a system of electrically charged particles with charge q in an external electromagnetic field described by scalar and vector potentials $(\phi,\vec{A})$, the Hamiltonian reads

$$H_N = \sum_{i=1}^{N} \frac{\big[\vec{p}_i - q\vec{A}(t,\vec{r}_i)\big]^2}{2m} + \sum_{i=1}^{N} q\phi(t,\vec{r}_i) + \sum_{1\le i<j\le N} W\big(|\vec{r}_i - \vec{r}_j|\big), \tag{III.9}$$

where the energy corresponding to the scalar potential has been denoted by $q\phi$ instead of V.

In the remainder of this Chapter, we shall use the shorthand notations $W_{ij} \equiv W\big(|\vec{r}_i - \vec{r}_j|\big)$ as well as

$$\vec{F}_i \equiv -\vec{\nabla}_{\vec{r}_i} V(\vec{r}_i), \tag{III.10a}$$

$$\vec{K}_{ij} \equiv -\vec{\nabla}_{\vec{r}_i} W\big(|\vec{r}_i - \vec{r}_j|\big), \tag{III.10b}$$

for the forces upon particle i due to the external potential V and to particle $j \neq i$, respectively. In accordance with Newton's third law,

$$\vec{K}_{ij} = -\vec{K}_{ji}, \tag{III.10c}$$

which follows from $\vec{\nabla}_{\vec{r}_j} W_{ij} = -\vec{\nabla}_{\vec{r}_i} W_{ij}$ and the relabeling of particles.

Remarks:

∗ The above forms of the Hamilton function implicitly assume that the forces at play in the system are additive, as is always the case in Newtonian physics.

---

(37) More precisely, $W\big(|\vec{r}_i - \vec{r}_j|\big)$ is the interaction potential energy.

[ah] W. Snellius, 1580–1626 [ai] R. Descartes, 1596–1650

$*$ In the absence of interactions, the Hamilton function (III.8) clearly reduces to that of a classical ideal gas.

III.2.2 System of neutral particles

In the absence of external vector potential, the Hamilton equations (II.1) with the Hamilton function (III.8) read

$$\dot{\vec{r}}_i = \vec{\nabla}_{\vec{p}_i} H_N = \frac{\vec{p}_i}{m}, \qquad \dot{\vec{p}}_i = -\vec{\nabla}_{\vec{r}_i} H_N = \vec{F}_i + \sum_{j\neq i} \vec{K}_{ij}. \tag{III.11}$$

The first equation simply states that linear momentum and canonical momentum coincide, from where it follows that the second equation is Newton's second law.

Accordingly, the Liouville equation (III.5) becomes

$$\frac{\partial \rho_N}{\partial t} + \sum_{i=1}^{N} \vec{v}_i \cdot \vec{\nabla}_{\vec{r}_i} \rho_N + \sum_{i=1}^{N} \vec{F}_i \cdot \vec{\nabla}_{\vec{p}_i} \rho_N + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \vec{K}_{ij} \cdot \vec{\nabla}_{\vec{p}_i} \rho_N = 0. \tag{III.12}$$

III.2.2 a BBGKY hierarchy

The reduced k-particle phase density $f_k(t, \vec{r}_1, \vec{p}_1, \ldots, \vec{r}_k, \vec{p}_k)$ follows from the Γ-space probability density ρ_N after integrating out the degrees of freedom of the remaining $N-k$ particles. Similarly, integrating the Liouville equation (III.12) over the positions and momenta $\vec{r}_j$, $\vec{p}_j$ of $N-k$ particles gives the evolution equation obeyed by f_k.

This integration is made simpler by the boundary conditions (III.6): since the velocity $\vec{v}_i$ resp. the force $\vec{F}_i$ is independent of $\vec{r}_i$ resp. $\vec{p}_i$, it can be factored outside of the integral

$$\int \vec{v}_i \cdot \vec{\nabla}_{\vec{r}_i} \rho_N \, \mathrm{d}^3\vec{r}_i \quad \text{resp.} \quad \int \vec{F}_i \cdot \vec{\nabla}_{\vec{p}_i} \rho_N \, \mathrm{d}^3\vec{p}_i, \tag{III.13}$$

leaving an integral of the type (III.7), which identically vanishes.
Additionally, we shall assume that we are allowed to exchange the integration and partial differentiation operations when needed.

Integrating the Liouville equation (III.12) over the $N-1$ particles labeled 2, 3, …, N with the proper normalization factor $\alpha_{N,1} = 1/(2\pi\hbar)^3$ yields for the evolution of the single-particle density (III.1a) the equation

$$\frac{\partial f_1(t, \vec{r}_1, \vec{p}_1)}{\partial t} + \vec{v}_1 \cdot \vec{\nabla}_{\vec{r}_1} f_1(t, \vec{r}_1, \vec{p}_1) + \vec{F}_1 \cdot \vec{\nabla}_{\vec{p}_1} f_1(t, \vec{r}_1, \vec{p}_1)$$
$$+ \sum_{j=2}^{N} \frac{1}{(2\pi\hbar)^3} \int \vec{K}_{1j} \cdot \vec{\nabla}_{\vec{p}_1} \rho_N\big(t, \{\vec{r}_i\}, \{\vec{p}_i\}\big) \, \mathrm{d}^{6(N-1)}\mathcal{V}(2, \ldots, N) = 0,$$

where the notation $\mathrm{d}^{6(N-1)}\mathcal{V}(2, \ldots, N)$ emphasizes the labels of the particles which are integrated out. In the integral in the second line, it is convenient to write

$$\mathrm{d}^{6(N-1)}\mathcal{V}(2, \ldots, N) = \frac{1}{N-1} \frac{\mathrm{d}^3\vec{r}_j \, \mathrm{d}^3\vec{p}_j}{(2\pi\hbar)^3} \, \mathrm{d}^{6(N-2)}\mathcal{V}(2, \ldots, j-1, j+1, \ldots, N)$$

to isolate particle j from the $N-2$ particles $i \neq j$. The latter are then straightforwardly dealt with: integrating over their phase-space coordinates yields, together with the factor $1/(2\pi\hbar)^6 = \alpha_{N,2}$, precisely the two-particle phase-space density $f_2(t, \vec{r}_1, \vec{p}_1, \vec{r}_j, \vec{p}_j)$, i.e.

$$\frac{1}{(2\pi\hbar)^3} \int \vec{K}_{1j} \cdot \vec{\nabla}_{\vec{p}_1} \rho_N\big(t, \{\vec{r}_i\}, \{\vec{p}_i\}\big) \, \mathrm{d}^{6(N-1)}\mathcal{V}(2, \ldots, N) = \frac{1}{N-1} \int \vec{K}_{1j} \cdot \vec{\nabla}_{\vec{p}_1} f_2(t, \vec{r}_1, \vec{p}_1, \vec{r}_j, \vec{p}_j) \, \mathrm{d}^3\vec{r}_j \, \mathrm{d}^3\vec{p}_j.$$

Thanks to the indistinguishability of the particles, the remaining integral over $\mathrm{d}^3\vec{r}_j \, \mathrm{d}^3\vec{p}_j$ is actually independent of the value of the index j: the sum over j gives $N-1$ times the same contribution, which may be rewritten with the dummy integration variables $\vec{r}_2, \vec{p}_2$, while the factor $N-1$ cancels that present in the denominator of $\mathrm{d}^{6(N-1)}\mathcal{V}$. All in all, one thus obtains

$$\boxed{\left(\frac{\partial}{\partial t} + \vec{v}_1 \cdot \vec{\nabla}_{\vec{r}_1} + \vec{F}_1 \cdot \vec{\nabla}_{\vec{p}_1}\right) f_1(t,\vec{r}_1,\vec{p}_1) = -\int \vec{K}_{12} \cdot \vec{\nabla}_{\vec{p}_1} f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2)\, \mathrm{d}^3\vec{r}_2\, \mathrm{d}^3\vec{p}_2.} \qquad \text{(III.14a)}$$

That is, the equation governing the evolution of the single-particle density f_1 involves the two-particle density f_2.

By integrating out $N-2$ particles in the Liouville equation (III.12), one finds in a similar manner the evolution equation for the dynamics of the reduced two-particle density(38)

$$\boxed{\begin{aligned}\left[\frac{\partial}{\partial t} + \vec{v}_1 \cdot \vec{\nabla}_{\vec{r}_1} + \vec{v}_2 \cdot \vec{\nabla}_{\vec{r}_2} + \vec{F}_1 \cdot \vec{\nabla}_{\vec{p}_1} + \vec{F}_2 \cdot \vec{\nabla}_{\vec{p}_2} + \vec{K}_{12} \cdot \left(\vec{\nabla}_{\vec{p}_1} - \vec{\nabla}_{\vec{p}_2}\right)\right] f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2) = \\ -\int \left(\vec{K}_{13} \cdot \vec{\nabla}_{\vec{p}_1} + \vec{K}_{23} \cdot \vec{\nabla}_{\vec{p}_2}\right) f_3(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2,\vec{r}_3,\vec{p}_3)\, \mathrm{d}^3\vec{r}_3\, \mathrm{d}^3\vec{p}_3.\end{aligned}} \qquad \text{(III.14b)}$$

In turn, the evolution equation for f_2 involves the three-particle phase-space density f_3.

> The only trick in deriving this form of the evolution equation consists in using relation (III.10c) to rewrite $\vec{K}_{21} \cdot \vec{\nabla}_{\vec{p}_2}$ as $-\vec{K}_{12} \cdot \vec{\nabla}_{\vec{p}_2}$.

This generalizes to the equation of motion for the reduced k-particle density f_k—obtained by integrating out $N-k$ particles in the Liouville equation (III.12)—, which is not closed, but depends on the $(k+1)$-particle density f_{k+1} for $1 \leq k < N$:

$$\boxed{\begin{aligned}\left[\frac{\partial}{\partial t} + \sum_{j=1}^{k}\left(\vec{v}_k \cdot \vec{\nabla}_{\vec{r}_j} + \vec{F}_j \cdot \vec{\nabla}_{\vec{p}_j}\right) + \sum_{1\leq i<j\leq k} \vec{K}_{ij} \cdot \left(\vec{\nabla}_{\vec{p}_i} - \vec{\nabla}_{\vec{p}_j}\right)\right] f_k(t,\vec{r}_1,\vec{p}_1,\ldots,\vec{r}_k,\vec{p}_k) = \\ -\int \sum_{j=1}^{k} \vec{K}_{j,k+1} \cdot \vec{\nabla}_{\vec{p}_j} f_{k+1}(t,\vec{r}_1,\vec{p}_1,\ldots,\vec{r}_k,\vec{p}_k,\vec{r}_{k+1},\vec{p}_{k+1})\, \mathrm{d}^3\vec{r}_{k+1}\, \mathrm{d}^3\vec{p}_{k+1}.\end{aligned}} \qquad \text{(III.14c)}$$

The meaning of this equation is rather clear: The left-hand side describes the "free", or *streaming* evolution of f_k, involving only the k particles under consideration—including their reciprocal interactions. On the other hand, the right-hand side is a *collision term* describing the interaction between any of these k particles and a partner among the group of the other, not measured particles.

Eventually, the N-particle phase-space density f_N is simply ρ_N multiplied by a normalization factor, so that the evolution equation for f_N is closed, since it is the Liouville equation (III.12) itself

$$\left[\frac{\partial}{\partial t} + \sum_{j=1}^{N}\left(\vec{v}_j \cdot \vec{\nabla}_{\vec{r}_j} + \vec{F}_j \cdot \vec{\nabla}_{\vec{p}_j}\right) + \sum_{1\leq i<j\leq N} \vec{K}_{ij} \cdot \left(\vec{\nabla}_{\vec{p}_i} - \vec{\nabla}_{\vec{p}_j}\right)\right] f_N(t,\vec{r}_1,\vec{p}_1,\ldots,\vec{r}_N,\vec{p}_N) = 0. \qquad \text{(III.14d)}$$

Together, the coupled equations (III.14a)–(III.14d) constitute the so-called *BBGKY hierarchy*, where the initials stand for Bogolioubov(aj)–Born(ak)–Green(al)–Kirkwood(am)–Yvon(an) (in alphabetical and reverse chronological order).

The system of equations (III.14) is exact, in the sense that it is strictly equivalent to the Liouville equation for the probability density in Γ-space. Yet since the hierarchy involves all reduced densities, and accordingly N equations, one loses the simplification aimed at when considering single-or two-particle densities only. To recover some simplicity—and thereby be able to actually compute the evolution of the system—, one has to choose a closure procedure, based on physical arguments,

---

(38) Note that there is an erroneous factor of $\frac{1}{2}$ in Eq. (3.60) of Ref. [33].

(aj) N. N. Bogolioubov, 1909–1992 (ak) M. Born, 1882–1970 (al) H. S. Green, 1920–1999 (am) J. G. Kirkwood, 1907–1959 (an) J. Yvon, 1903–1979 which amounts to truncating the hierarchy at some level. This is most often done after Eq. (III.14a), sometimes after Eq. (III.14b).

The most drastic closure prescription, which is discussed in § III.2.2 b below, consists in fully neglecting interactions. A similar, yet physically richer, scheme is to keep Eq. (III.14a) intact, yet to assume that the right-hand side of Eq. (III.14b) vanishes. Alternative procedures will be considered in § III.2.3 b and in the next chapter.

Remarks:

∗ If the number of particles in the system is not fixed, i.e. in a grand-canonical description, there is no upper bound to the hierarchy.

∗ The generalization to a system consisting of several particle types, labeled a, $b\dots$, interacting with each other over respective potentials W^{ab}, is quite straightforward. One first needs to introduce the reduced densities f_k^b, $f_k^b\dots$, for having k particles of a given type in an infinitesimal k-particle phase-space volume, as well as "mixed" densities $f^{ab\cdots}_{k_a k_b\dots}$ for having k_a particles of type a, k_b particles of type $b\dots$ within corresponding phase-space volume elements. These various densities—including the mixed ones—then obey coupled equations similar to those of the hierarchy (III.14), with in addition sums over the various particle types. For instance, the evolution equation for f_1^a involves on its right-hand side not only f_2^a, but also all $f_{1,1}^{ab}$ involving the particles of type $b \neq a$ which couple to those of type a.

III.2.2 b System of non-interacting neutral particles: single-particle Liouville equation

Consider first the case where the particles in the system are not interacting, i.e. when W vanishes identically. The Hamilton function H_N can then be expressed as the sum over the N particles of a single-particle Hamiltonian $h \equiv \vec{p}^{\,2}/2m + V(t,\vec{r})$.

Under these conditions, the various equations of the BBGKY hierarchy (III.14) decouple from each other. For instance, the evolution equation (III.14a) for the dynamics of the single-particle density f_1 becomes the *single-particle Liouville equation* (or *collisionless Boltzmann equation*)

$$\frac{\partial f_1(t,\vec{r},\vec{p})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}} f_1(t,\vec{r},\vec{p}) + \vec{F}(t,\vec{r})\cdot\vec{\nabla}_{\vec{p}} f_1(t,\vec{r},\vec{p}) = 0, \tag{III.15}$$

with $\vec{F} = -\vec{\nabla}_{\vec{r}} V$ the external force acting on particles. The contribution $\big(\vec{v}\cdot\vec{\nabla}_{\vec{r}} + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\big) f_1$—which is actually the Poisson bracket in μ-space of $f_1(t,\vec{r},\vec{p})$ and the single-particle Hamiltonian $h(\vec{r},\vec{p})$—is often referred to as *drift term*.

The latter implicitly defines two time scales of the system, namely

$$\tau_{\rm s} \sim \big(\vec{v}\cdot\vec{\nabla}_{\vec{r}}\big)^{-1}, \qquad \tau_{\rm e} \sim \big(\vec{F}\cdot\vec{\nabla}_{\vec{p}}\big)^{-1}. \tag{III.16}$$

$\tau_{\rm s}$ is the time for a particle to cross the typical distance over which the single-particle density f_1 is varying. As such, it is also the characteristic time scale for smoothing out spatial inhomogeneities of the single-particle distribution.

In turn, $\tau_{\rm e}$ is the typical time associated with the gradient imposed by the external potential V. It is thus also the characteristic time scale over which the system inhomogeneities will relax under the influence of $\vec{F}$ to the equilibrium solution matching the external potential. $\tau_{\rm e}$ is generally larger than $\tau_{\rm s}$.

Remark: One sees at once that the other equations of the hierarchy involve the same two time scales, and no further one. That is, any f_k evolves with the same time scale as f_1, which will no longer be the case in the presence of interactions.

III.2.2 c Influence of inter-particle interactions

Let us now take into account the effect of interactions between particles, considering the first equation (III.14a) of the BBGKY hierarchy with a non-vanishing right-hand side. On the other hand, we assume that the right member of the second equation (III.14b) vanishes, i.e. we neglect the influence of the three-particle density on the evolution of f_2.

Inspecting the resulting evolution equations, one finds that there is a new time scale besides $\tau_{\rm s}$ and $\tau_{\rm e}$, set by the intensity of the pairwise interaction, namely

$$\tau_{\rm c} \sim \big(\vec{K} \cdot \vec{\nabla}_{\vec{p}}\big)^{-1}. \tag{III.17}$$

From the evolution equation for the two-particle density f_2, the *collision duration* $\tau_{\rm c}$ appears as the characteristic time scale over which collisions between two particles smooth out the differences in their respective momentum distributions. $\tau_{\rm c}$ is actually the smallest time scale, much smaller than $\tau_{\rm s}$ and $\tau_{\rm e}$.

As a result, the evolution of f_2 is actually quicker than that of f_1 in the non-interacting case, for this quick time scale is absent from the single-particle Liouville equation (III.15). In turn, it means that when particles are allowed to interact with each other, the pace for the evolution of f_1 is not set by its slow drift term, but by the fast collision term in the right-hand side.

Remark: One could then wonder whether the time scale for the evolution of f_2 is not governed by the collision term involving f_3, which we are neglecting here. It can be checked that this is not the case in the dilute regime in which only pairwise interactions are important.

In the derivation of the BBGKY hierarchy (III.14), we have from the start assumed that particles only interact with each other in pairs. It is quite straightforward to see what would happen if we also allowed for genuine interactions between three, four or more particles at once.

The effect of the latter is simply to add further contributions to the collision and drift terms of the equations of motion. For instance, when allowing for three-particle interactions, the equation governing the evolution of f_k would not only involve f_{k+1}, but also f_{k+2}, corresponding to the case where one particle among the k under consideration is interacting with two “outsiders”. Despite this complication, one can still derive an exact hierarchy of equations.

III.2.2 d BBGKY hierarchy in position-velocity space

Instead of characterizing a pointlike particle by a point in its μ-space, i.e. by its position and canonical momentum, one sometimes rather makes use of a point in *position-velocity space*, i.e. one specifies the positions and velocities of the particles.

In that representation, one can similarly introduce k-particle densities over a k-particle position-velocity space, hereafter denoted as $f_{\vec{v},k}$, such that $f_{\vec{v},k}(t,\vec{r}_1,\vec{v}_1,\dots,\vec{r}_k,\vec{v}_k)\,{\rm d}^3\vec{r}_1\,{\rm d}^3\vec{v}_1\cdots{\rm d}^3\vec{r}_k\,{\rm d}^3\vec{v}_k$ is the number of particles in the infinitesimal volume element ${\rm d}^3\vec{r}_1\,{\rm d}^3\vec{v}_1\cdots{\rm d}^3\vec{r}_k\,{\rm d}^3\vec{v}_k$ around the point $(\vec{r}_1,\vec{v}_1,\dots,\vec{r}_k,\vec{v}_k)$. The density $f_{\vec{v},k}$ is related to the reduced phase-space density f_k by the relation

$$f_{\vec{v},k}(t,\vec{r}_1,\vec{v}_1,\dots,\vec{r}_k,\vec{v}_k)\,{\rm d}^3\vec{r}_1\,{\rm d}^3\vec{v}_1\cdots\,{\rm d}^3\vec{r}_k\,{\rm d}^3\vec{v}_k = f_k(t,\vec{r}_1,\vec{p}_1,\dots,\vec{r}_k,\vec{p}_k)\,{\rm d}^3\vec{r}_1\,{\rm d}^3\vec{p}_1\cdots\,{\rm d}^3\vec{r}_k\,{\rm d}^3\vec{p}_k. \tag{III.18}$$

In the case considered in this section of a system of particles in absence of an external vector potential, the velocity $\vec{v}$ of a particle is simply proportional to its canonical momentum $\vec{p}$, so that $f_{\vec{v},k}$ is trivially proportional to f_k and $\vec{\nabla}_{\vec{v}}$ to $\vec{\nabla}_{\vec{p}}$. One checks at once that the position-velocity space densities obey a similar BBGKY hierarchy as the phase-space densities, the only modification being the substitutions of $\vec{\nabla}_{\vec{p}_j}$ by by $\vec{\nabla}_{\vec{v}_j}$ and $\dot{\vec{p}}_j$ by $\dot{\vec{v}}_j$. For instance, the first equation of the hierarchy reads

$$\left(\frac{\partial}{\partial t} + \vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1} + \frac{\vec{F}_1}{m}\cdot\vec{\nabla}_{\vec{v}_1}\right) f_{\vec{v},1}(t,\vec{r}_1,\vec{v}_1) = -\int \frac{\vec{K}_{12}}{m}\cdot\vec{\nabla}_{\vec{v}_1} f_{\vec{v},2}(t,\vec{r}_1,\vec{v}_1,\vec{r}_2,\vec{v}_2)\,{\rm d}^3\vec{r}_2\,{\rm d}^3\vec{v}_2. \tag{III.19}$$

The real advantage of the formulation in position-velocity space is for the case, addressed in next section, of a system of charged particles in the presence of an external vector potential.

III.2.3 System of charged particles

We now briefly introduce the subtleties which appear when the particles in the system are coupled to an external vector potential, as is for instance the case of electrically charged particles in an external (electro)magnetic field.

In this subsection, we shall exceptionally denote the canonical momentum conjugate to position $\vec{r}_i$ as $\vec{\pi}_i$, and the reduced phase-space densities as $f_{\vec{\pi},k}$.

III.2.3 a Phase-space vs. position-velocity space description

The evolution equations of the BBGKY hierarchy (III.14) have been derived starting from the Liouville equation on phase space, i.e. within the framework of a Hamiltonian formalism. A similar approach can also be adopted in the presence of an external vector field, yet there appear two complications, which can be traced back to the expression of the canonical momentum, derived from the Hamilton equations following from the Hamilton function (III.9)

$$\dot{\vec{r}}_i = \frac{\vec{\pi}_i - q\vec{A}(t,\vec{r}_i)}{m}, \tag{III.20a}$$

$$\dot{\vec{\pi}}_i = -q\vec{\nabla}_{\vec{r}}\,\phi(t,\vec{r}_i) + q\big(\dot{\vec{r}}_i\cdot\vec{\nabla}_{\vec{r}}\big)\vec{A}(t,\vec{r}_i) + q\dot{\vec{r}}_i\times\Big[\vec{\nabla}_{\vec{r}}\times\vec{A}(t,\vec{r}_i)\Big] + \sum_{j\neq i}\vec{K}_{ij}, \tag{III.20b}$$

where in the second equation we have made use of the first one.

- The time derivatives $\vec{v}_i = \dot{\vec{r}}_i$ and $\dot{\vec{\pi}}_i$ are no longer independent of $\vec{r}_i$ or $\vec{\pi}_i$, respectively. As a consequence, the integrals (III.13)—in which $\vec{\nabla}_{\vec{p}_i}$ is to be replaced by $\vec{\nabla}_{\vec{\pi}_i}$—are no longer trivial. Yet by using the Hamilton equations, one can show at the cost of a few integration by parts that the contributions of these integrals vanish, so that the evolution of the reduced phase-space density $f_{\vec{\pi},k}$ is eventually governed by an equation with the same structure as Eq. (III.14c), although with different factors in front of the gradients.

  For example, the collisionless evolution equation for the single-particle density $f_{\vec{\pi},1}$ is

  $$\frac{\partial f_{\vec{\pi},1}(t,\vec{r},\vec{\pi})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}f_{\vec{\pi},1}(t,\vec{r},\vec{\pi}) + \dot{\vec{\pi}}\cdot\vec{\nabla}_{\vec{\pi}}f_{\vec{\pi},1}(t,\vec{r},\vec{\pi}) = 0, \tag{III.21}$$

  which differs from Eq. (III.15) because of the Hamilton equations (III.20).

- The second issue with the evolution equations for the reduced phase-space densities derived within the Hamiltonian formalism is that they are not manifestly gauge invariant, which is quite natural since the canonical momenta $\vec{\pi}_i$ themselves are gauge-dependent, see Eq. (III.20).

To remedy the latter problem, it is convenient to consider the densities in position-velocity space, instead of phase space. The resulting equations—which take the same form as in the absence of vector potential—only involve gauge invariant quantities. For instance, the collisionless evolution equation for the single-particle density $f_{\vec{v},1}$ is

$$\frac{\partial f_{\vec{v},1}(t,\vec{r},\vec{v})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}f_{\vec{v},1}(t,\vec{r},\vec{v}) + \frac{\vec{F}_{\rm L}}{m}\cdot\vec{\nabla}_{\vec{v}}f_{\vec{v},1}(t,\vec{r},\vec{v}) = 0, \tag{III.22}$$

with $\vec{F}_{\rm L} = q(\vec{\mathscr{E}}_{\rm ext} + \vec{v}\times\vec{\mathscr{B}}_{\rm ext})$ the Lorentz(ao) force due to external electromagnetic fields $(\vec{\mathscr{E}}_{\rm ext},\vec{\mathscr{B}}_{\rm ext})$. This equation is exactly the same as Eq. (III.19) in the absence of collision term. One can check that Eqs. (III.21) under consideration of Eq. (III.20) and Eq. (III.22) are actually equivalent.

Alternatively, one most often works in the space spanned by the positions $\vec{r}_i$ and the *linear* momenta $\vec{p}_i = m\vec{v}_i$. The densities f_k on that space obey exactly the same equations (III.14) as in the case of neutral particles. This is obvious in case there is no vector potential—conjugate and linear momenta $\vec{\pi}_i$ and $\vec{p}_i$ then coincide—and justifies our retaining the same notation as in Sec. III.2.2 for the densities.

(ao) H. A. Lorentz, 1853–1926

III.2.3 b Vlasov equation

An important example of macroscopic system of electrically charged particles is that of a plasma—where one actually has to consider two different types of particles, with positive and negative charges respectively, to ensure the mechanical stability.

In that context, Vlasov(ap) has introduced a closure prescription for the corresponding BBGKY hierarchy, which consists in assuming that the two-particle density factorizes into the product of single particle densities(39)

$$f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2) \simeq f_1(t,\vec{r}_1,\vec{p}_1)\, f_1(t,\vec{r}_2,\vec{p}_2). \tag{III.23}$$

Under this assumption, Eq. (III.19) becomes closed and can be rewritten as

$$\left[\frac{\partial}{\partial t} + \vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1} + \left(\vec{F}_1 + \int \vec{K}_{12} f_1(t,\vec{r}_2,\vec{p}_2)\,\mathrm{d}^3\vec{r}_2\,\mathrm{d}^3\vec{p}_2\right)\cdot\vec{\nabla}_{\vec{p}_1}\right] f_1(t,\vec{r}_1,\vec{p}_1) = 0, \tag{III.24}$$

with $\vec{F}_1 = q\big(\vec{\mathscr{E}}_{\text{ext}} + \vec{v}\times\vec{\mathscr{B}}_{\text{ext}}\big)$ the Lorentz force due to the external electromagnetic field. This constitutes the *Vlasov equation*.

The integral term inside the brackets can be viewed as an average force exerted by the partner particle 2 on particle 1. In the specific case of an electromagnetic plasma—assuming for simplicity that the particles in the system have velocities much smaller than the speed of light, so that their mutual interaction is mostly of Coulombic(aq) nature—this average force is that due to the mean electrostatic field created by the other particles:

$$\left[\frac{\partial}{\partial t} + \vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1} + q\big(\vec{\mathscr{E}}' + \vec{\mathscr{E}}_{\text{ext}} + \vec{v}\times\vec{\mathscr{B}}_{\text{ext}}\big)\cdot\vec{\nabla}_{\vec{p}_1}\right] f_1(t,\vec{r}_1,\vec{p}_1) = 0$$

with

$$\vec{\mathscr{E}}' \equiv \frac{1}{q}\int \vec{K}_{12} f_1(t,\vec{r}_2,\vec{p}_2)\,\mathrm{d}^3\vec{r}_2\,\mathrm{d}^3\vec{p}_2.$$

The Vlasov assumption is thus an instance of *mean-field approximation*.

Remark: The Vlasov equation (III.24) is nonlinear, and thus non-trivial. In practice it must be solved in a self-consistent way, since the mean field in which f_1 evolves depends on f_1 itself.

Bibliography for Chapter III

- Huang, *Statistical Mechanics* [33], chapter 3.5.
- Landau & Lifshitz, *Course of theoretical physics. Vol. 10: Physical kinetics* [5], chapter I § 16.
- Pottier, *Nonequilibrium statistical physics* [6], chapter 4.

---

(39) We do not write this hypothesis as an identity with an = sign to accommodate the possible mismatch between the normalizations of f_1 and f_2.

(ap) A. A. Vlasov, 1908–1975 (aq) C. A. Coulomb, 1736–1806

Chapter IV

Boltzmann equation

In the previous chapter, it has been shown that the Liouville equation describing the evolution in Γ-space of the probability density for a collection of many classical particles can be equivalently recast as a system of coupled evolution equations, the BBGKY hierarchy, for the successive reduced phase-space densities f_k. To become tractable, the hierarchy has to be truncated, by closing the system of equations on the basis of some physical ansatz.

A non-trivial example of such a prescription is the so-called assumption of *molecular chaos*, introduced by L. Boltzmann. This hypothesis provides a satisfactory recipe for systems of *weakly interacting* particles, the paradigmatic example thereof being a dilute gas. In such systems, the time scales characterizing the typical duration of a collision or the interval between two successive scatterings of a given particle are well separated, as discussed in Sec. IV.1. Together with an analysis at the microscopic level of the collision processes which induce changes in the single-particle density, Boltzmann's ansatz yields a generic kinetic equation for the evolution of the phase space distribution of a particle (Sec. IV.2).

Irrespective of the precise form of the interaction responsible for the scattering processes—at least as long as the latter remain elastic and are invariant under time and space parity—, the kinetic Boltzmann equation leads to balance relations for various macroscopic quantities, which are first introduced in Sec. IV.3. In addition, we shall see that the Boltzmann equation describes a macroscopic time evolution which is not invariant under time reversal, even though the underlying microscopic interactions are.

Once an equation has been obtained comes the natural question about its solutions (Sec. IV.4). One easily shows that the Boltzmann equation is fulfilled by a stationary solution, which describes thermodynamic equilibrium at the microscopic level. It can even be shown that this solution often constitutes the long-time behavior of an evolution. Apart from that one, extremely few other analytical solutions are known, and one uses various methods to find distributions that approximate true solutions for out-of-equilibrium situations.

The latter distributions are especially interesting, inasmuch as they allow the computation within the framework of the Boltzmann equation to the of macroscopic quantities, namely for instance of transport coefficients (Sec. IV.5) and of the dissipative fluxes in a fluid (Sec. IV.6), thereby also yielding the form of the equations of motion for the fluid.

Throughout this chapter, $\vec{p}$ denotes the *linear* momentum, not the canonical one, so that the various equations hold irrespective of the presence of a external vector potential.

IV.1 Description of the system

The kinetic Boltzmann equation is a simplification of the BBGKY hierarchy (III.14), or equivalently of the N-particle Liouville equation (II.10b), which implies the introduction of assumptions. In this first section, we introduce the generic hypotheses underlying *weakly interacting systems* of classical particles, starting with a discussion of the various length and time scales in the system. In a second step, we discuss the interactions in the system, adopting for the remainder of the chapter the most simple, non-trivial choice—which is that originally made by Boltzmann.

IV.1.1 Length and time scales

Throughout this chapter, we shall consider systems of classical particles, i.e. we assume that we may meaningfully use the notion of phase space and of distributions over it. As is known from equilibrium statistical mechanics, this holds at thermodynamic equilibrium provided the mean interparticle distance $d \sim n^{-1/3}$, with n the particle number density in position space, is large compared to the thermal de Broglie[(ar)] wavelength $\lambda_{\rm th} = 2\pi\hbar/\sqrt{2\pi m k_B T}$. In a non-equilibrium case, as we are interested in here, d should be much larger than the de Broglie wavelength $2\pi\hbar/\langle|\vec{p}|\rangle$, with $\langle|\vec{p}|\rangle$ the typical momentum of particles.

Since we want to allow for scatterings between particles, this condition is not always fulfilled. Colliding partners might come close enough to each other that their wave functions start overlapping significantly. In that regime, classical notions like the particle trajectories or the impact parameter of a scattering process become meaningless, as they have no (exact) equivalent in quantum mechanics.

As we discuss shortly, the kinetic equations we shall consider hereafter hold for dilute systems, in which the successive collisions of a given particle are independent of each other. Accordingly, one only need to consider the ingoing and outgoing states of a scattering event, as well as the probability rate to transition from the former to the latter, irrespective of the detailed description of the process. In particular, even though the scattering rate may (and indeed should!) be computed quantum mechanically, the evolution of a particle between collisions remains classical.

As we shall see in Chap. ??, it is possible to formulate fully quantum-mechanical kinetic equations, from which the classical Boltzmann equation can be recovered in a specific limit. In that respect, the assumption of classical particles is only a simplification, not a necessity.

A more crucial assumption underlying kinetic approaches is that of the "diluteness" of the system under consideration. More precisely, the typical distance $n^{-1/3}$ between two neighboring particles is assumed to be much larger than the characteristic range r_0 of the force between them. Again, this obviously cannot hold when the particles collide. Yet the requirement $r_0 \ll n^{-1/3}$ ensures that scatterings remain rare events, i.e. that most of the time any given particle is moving freely. Accordingly, the potential energy associated with the interactions between particles is much smaller than their kinetic energy: this condition defines a *weakly interacting* system. The paradigm for such a system is a dilute gas, yet we shall see further examples in Sec. IV.5.

The reader should be aware of the difference between "weakly interacting"—which is a statement on the diluteness of the system, as measured by a dimensionless number (*diluteness parameter*) like $n r_0^3$ or $n\sigma_{\rm tot}^{3/2}$ which should be much smaller than unity, where $\sigma_{\rm tot}$ denotes the total cross section—and "weakly coupled", which relates to the characteristic strength of the coupling of the interactions at play.

Thanks to the assumption of diluteness, the typical distance travelled by a particle between two successive collisions, the mean free path $\ell_{\rm mfp}$, must be much larger than the interaction range r_0. In consequence, any scattering event will happen long after the previous collisions (with different partners) of each of the two participating particles; that is, the successive scatterings of any particle are independent processes, they are said to be *incoherent*.

Since we do not wish to describe the details of collisions, which involve length scales of the order of r_0 and smaller, we may simply consider a *coarse-grained* version of position space, in which points separated by a distance of the order of r_0 or smaller are taken as one and the same. In the coarse-grained position space, scatterings thus only take place when two particles are at the same point, i.e. they are *local* processes.

In parallel, we introduce a coarse-grained version of the time axis [34], by restricting the temporal resolution of the description to some scale larger than the collision duration τ_0, defined as the typical

---

[(ar)] L. de Broglie, 1892–1987

time which a particle needs to travel a distance r_0. As scatterings actually take place on a time scale of the order of τ_0, they are instantaneous in the coarse-grained time. Conversely, all the time intervals δt we shall consider—even infinitesimal ones, which will as usual be denoted as dt—will implicitly fulfill the condition $\delta t \gg \tau_0$.

Remark: If the particles interact through long-range interactions, as e.g. the Coulomb force, the diluteness condition $n r_0^3 \ll 1$ is clearly violated. For such systems, the proper approach is to leave out such interactions from the term in the evolution equation that describes collision processes, and to reintroduce them "on the left-hand side of the equation" in the form of a mean field in which the particles are moving, as was done in § III.2.3 b when writing down the Vlasov equation.

Let $\vec{p}$ denote the linear momentum of a particle. We introduce the dimensionless single-particle distribution (or density) $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$, such that[(40)]

$$\bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\frac{\mathrm{d}^3\vec{r}\;\mathrm{d}^3\vec{p}}{(2\pi\hbar)^3} \tag{IV.1}$$

is the average number of particles which at time t are in the infinitesimal (coarse-grained) volume element $\mathrm{d}^3\vec{r}$ around position $\vec{r}$ and possess a momentum equal to $\vec{p}$ up to $\mathrm{d}^3\vec{p}$. The kinetic equation we shall derive and study hereafter will be the equation that governs the dynamics of $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$.

Integrating $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ over all possible momenta yields the local particle-number density on the coarse-grained position space at time t:

$$n(t,\vec{r}) = \int \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^3} \tag{IV.2}$$

[cf. Eq. (III.3b)]. In turn, the integral of $n(t,\vec{r})$ over position yields the total number N of particles in the system, which will be assumed to remain constant.

Remarks:

∗ The distribution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ is obviously a coarse-grained version of the dimensionless single-particle density on μ-space $\mathsf{f}_1(t,\vec{r},\vec{p})$ introduced in the previous chapter. One might already anticipate that f_1 provides a better description than $\bar{\mathsf{f}}$, since it corresponds to an increased resolution. The implicit loss of information when going from f_1 to $\bar{\mathsf{f}}$ should manifest itself when measuring the knowledge associated with each distribution for a given physical situation, i.e. when considering the corresponding statistical entropies.

> The notation $\bar{\mathsf{f}}$ is naturally suggestive of an average, so one may wonder whether f_1 can be meaningfully decomposed as $\mathsf{f}_1 = \bar{\mathsf{f}} + \delta\mathsf{f}$, with $\delta\mathsf{f}$ a "fluctuating part" whose coarse-grained value vanishes. That is, there might be a well-defined prescription for splitting f_1—which obeys the exact BBGKY hierarchy—into a "kinetic part" $\bar{\mathsf{f}}$—which satisfies the assumptions leading to the Boltzmann equation, in particular the position-space locality and the Stoßzahlansatz (IV.14)—and a "non-kinetic part" $\delta\mathsf{f}$, which should be irrelevant for weakly interacting systems. This decomposition can indeed be performed, with the help of projection operators (Chap. ??).

∗ Till now, no upper bounds were specified for the scales of the space-time cells which constitute the points $(t,\vec{r})$ in the coarse-grained description. In Sec. IV.4.1, we shall define local equilibrium distributions, which depend on various local fields. The latter should be slowly varying functions of t and $\vec{r}$, which implicitly sets upper bounds on the size of local cells. Thus the typical time between two collisions of a particle should be much larger than the time size of a local cell, while accordingly the mean free path ℓ_{mfp}—i.e. the characteristic length traveled by a particle between two collisions—should be much larger than the spatial size of a local cell.

---

(40) In Sec. IV.2.1–IV.2.3 we shall also use the notation $\bar{\mathsf{f}}_1$, instead of $\bar{\mathsf{f}}$, thus emphasizing the "single-particle" aspect of the distribution.

IV.1.2 Collisions between particles

Let us now discuss the interactions between particles. As already stated above, in the coarse-grained description the collisions between particles are local and instantaneous, i.e. the participants have to be at the same space-time point $(t,\vec{r})$. If we allow the presence of an external potential—which for the sake of consistency has to vary slowly, so that the coarse-graining procedure makes sense—, then it is assumed to have no influence on the microscopic scattering processes.

Formulating the assumption of a weakly interacting system at the particle-scattering level, it implies that the probability that two particles collide is already small, so that the probability for collisions between three or more particles becomes totally negligible. Accordingly, we shall from now on only consider *binary collisions* in the system.

For the sake of simplicity, we shall assume that the collisions are *elastic*. That is, we consider that the energy exchanged in a collision can neither excite internal degrees of freedom of the particles—which are thus considered structureless—, nor can it be transformed from kinetic into mass energy or vice-versa. As a consequence, the conservation of energy in a collision becomes the conservation of kinetic energy.

In addition, linear momentum is also conserved in each scattering process(41), as is angular momentum. However, the latter does not contribute any additional constraint on the kinematics—and since the collisions are from the start assumed to be elastic, it does not provide selection rules for the possible final states of a process.

Remarks:

∗ The statement of the separate conservation of kinetic energy is less innocent than it seems and is closely related to the weakly-interacting assumption, according to which potential energy is negligible compared to the kinetic one.

∗ In the (Boltzmann–)Lorentz model—that is, a special case of the Boltzmann-gas model with two (or more) types of particles, among which one species is assumed to consist of infinitely massive particles—momentum conservation takes a particular twist, since the momentum of one participant in the collision is infinite... which has the effect of lifting the corresponding kinematic constraint.

In summary, we only consider so-called “two-to-two elastic processes”, with two particles labeled 1 and 2 in the initial state and the same two particles, now labeled 3 and 4 for reasons that will be explained below, in the final state. Denoting by $\vec{p}_1$, $\vec{p}_2$, $\vec{p}_3$ and $\vec{p}_4$ the particle linear momenta before and after the collision—i.e., to be precise, far from the collision zone—, the scattering process will be symbolically written $\vec{p}_1,\vec{p}_2 \to \vec{p}_3,\vec{p}_4$, and the conservation laws trivially read

$$\frac{\vec{p}_1^{\,2}}{2m_1}+\frac{\vec{p}_2^{\,2}}{2m_2}=\frac{\vec{p}_3^{\,2}}{2m_3}+\frac{\vec{p}_4^{\,2}}{2m_4}, \qquad \vec{p}_1+\vec{p}_2=\vec{p}_3+\vec{p}_4, \tag{IV.3}$$

where m_1, m_2, m_3 and m_4 are the respective masses of the particles.

Given a specific model for the microscopic interactions, one can compute the transition rate from the initial state with two particles with momenta $\vec{p}_1$, $\vec{p}_2$, to the final state in which the particles acquire the momenta $\vec{p}_3$ and $\vec{p}_4$. For scatterings, this rate is characterized by a differential cross section, as we recall at the end of the section. Yet to be more general and use notations which easily generalize to the case where particles may decay, we introduce a related quantity

$$\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\equiv$$
$$w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{2\pi}{\hbar}\,\delta\!\left(\frac{\vec{p}_1^{\,2}}{2m_1}+\frac{\vec{p}_2^{\,2}}{2m_2}-\frac{\vec{p}_3^{\,2}}{2m_3}-\frac{\vec{p}_4^{\,2}}{2m_4}\right)(2\pi\hbar)^3\delta^{(3)}\big(\vec{p}_1+\vec{p}_2-\vec{p}_3-\vec{p}_4\big) \tag{IV.4a}$$

---

(41) This follows from the conservation of canonical momentum and the fact that collisions are local: if the particles are charged, the contributions from the vector potential to their canonical momenta are identical in the initial and final states.

such that

$$\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3} \tag{IV.4b}$$

is the number of collisions per unit time for a unit density (in position space) of colliding particle pairs, with final momenta in the range $\mathrm{d}^3\vec{p}_3\ \mathrm{d}^3\vec{p}_4$, while the Dirac(as) distributions encode the conditions (IV.3)

Remarks:

$*$ The various factors of 2π and $\hbar$ in relation (IV.4a) are there to ensure that $\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$ has a simple physical meaning, while $w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$ is a quantity which naturally emerges from a quantum mechanical calculation. As a matter of fact, $w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$ is then the squared modulus of the relevant element of the T-matrix (transition matrix), computed for initial single-particle states normalized to one particle per unit volume. Accordingly, the identities (IV.5) below are actually relations between T-matrix elements.

$*$ When the colliding partners are identical, and thus—to respect the underlying quantum mechanical description—indistinguishable, then two final states that only differ through the exchange of the particle labels are actually a single state. To account for this, we shall later have to divide $\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$ by 2 when we integrate over both $\vec{p}_3$ and $\vec{p}_4$, to avoid double counting.

As is customary, we make a further assumption on the interactions involved in the scatterings, namely that they are invariant under space parity and under time reversal. The transition rates for processes are thus unchanged when all position vectors $\vec{r}$ are replaced by $-\vec{r}$, as well as under the transformation $t\to -t$.

Therefore, we first have, thanks to the invariance under space parity

$$w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)=w(-\vec{p}_1,-\vec{p}_2\to-\vec{p}_3,-\vec{p}_4), \tag{IV.5a}$$

where we used that $\vec{p}=m\,\mathrm{d}\vec{r}/\mathrm{d}t$ is transformed into $-\vec{p}$. In turn, the invariance under time reversal yields

$$w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)=w(-\vec{p}_3,-\vec{p}_4\to-\vec{p}_1,-\vec{p}_2), \tag{IV.5b}$$

where we took into account both the transformations of the individual momenta and of the time direction of the scattering.

Combining both invariances together, one finds the identity

$$w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)=w(\vec{p}_3,\vec{p}_4\to\vec{p}_1,\vec{p}_2), \tag{IV.5c}$$

which relates scattering processes which are both space- and time-reversed with respect to each other. The process on the right-hand side of Eq. (IV.5c) is often called the *inverse collision* to that on the left-hand side. The identity of the transition rates for a process and the "inverse" one is referred to as *microscopic reversibility* or *microreversibility*.

Reminder: Classical scattering theory

Consider the scattering $\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4$. The description of the process is simpler in the center-of-momentum frame(42) of the colliding particles, where the linear momenta become $\vec{p}_1{}'=-\vec{p}_2{}'$ and $\vec{p}_3{}'=-\vec{p}_4{}'$. In that frame—in which kinematic quantities will be denoted with a prime—one defines the *differential scattering cross section* $\mathrm{d}^2\sigma/\mathrm{d}^2\Omega'$, which characterizes the number of particles from an incident flux density which are deflected per unit time into a solid angle element $\mathrm{d}^2\Omega'$ around

---

(42) Remember that the "center of mass" has momentum $\vec{p}_1+\vec{p}_2=\vec{p}_3+\vec{p}_4$ and mass $m_1+m_2=m_3+m_4$.

(as) P. A. M. Dirac, 1902–1984

some direction with respect to the initial density. More precisely, the differential cross section is defined as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi') = \frac{|\vec{\jmath}\,'_{\rm out}(\theta',\varphi')|}{|\vec{\jmath}\,'_{\rm in}|}\,r^2, \tag{IV.6}$$

where $\vec{\jmath}\,'_{\rm in}$ is the flux density falling on a single scattering center while $\vec{\jmath}\,'_{\rm out}(\theta,\varphi)$ is the outgoing flux density in the direction (θ',φ'), so that out of the $|\vec{\jmath}\,'_{\rm in}|$ particles crossing a unit area pro unit time, a number $|\vec{\jmath}\,'_{\rm out}(\theta,\varphi)|\,\mathrm{d}^2\mathcal{S}$ leave the collision zone through a surface element $\mathrm{d}^2\mathcal{S}$ in the direction (θ',φ') situated at a distance r from the scattering center (see Fig. IV.1).

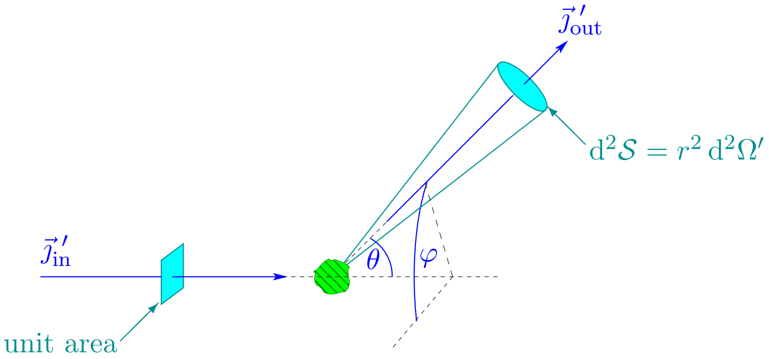

Figure IV.1 – Representation of the quantities entering the definition (IV.6) of the differential cross section.

If the colliding particles are non-identical, and thus distinguishable, say $m_3 = m_1 \neq m_2 = m_4$, then the scattering angle θ' is that between the (Galilei-invariant) incoming and outgoing relative velocities $\vec{v}_2 - \vec{v}_1$ and $\vec{v}_4 - \vec{v}_3$.(43) The solid-angle element $\mathrm{d}^2\Omega'$ is then equivalent to a volume element $\mathrm{d}^3\vec{p}_3\,\mathrm{d}^3\vec{p}_4$ in the joint momentum space of the two particles. More precisely, the number of collisions per unit time with final momenta in that range for a unit phase-space density of incoming particle pairs, $\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\mathrm{d}^3\vec{p}_3\,\mathrm{d}^3\vec{p}_4$, is related to the differential cross section by the identity (in the sense of distributions)

$$\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3} = \big|\vec{v}_2-\vec{v}_1\big|\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi')\,\mathrm{d}^2\Omega', \tag{IV.7}$$

which may be see as a definition for $w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$ in the case of classical scattering.

Remark: Integrating the differential cross section over all possible directions of the scattered particle yields the total cross section $\sigma_{\rm tot}$, which classically is the area of the effective... cross section of the target as seen by the projectile. For short-range interactions, this total cross section is of the order of the squared range of the interaction, $\sigma_{\rm tot}\sim r_0^2$. In turn, the total cross section allows to estimate the mean free path $\ell_{\rm mfp}$ as $\ell_{\rm mfp}\sim 1/n\sigma_{\rm tot}$, with n the particle number density. The assumption $n^{-1/3}\gg r_0$ is then equivalent to $\ell_{\rm mfp}\gg n^{-1/3}$, i.e. the mean free path is much larger than the typical inter-particle distance.(44) Accordingly, the diluteness of the system is sometimes measured with the dimensionless parameter $n\ell_{\rm mfp}^3$, instead of $nr_0^3\sim n\sigma_{\rm tot}^{3/2}$.

IV.2 Boltzmann equation

We now derive the equation governing the dynamics of the single-particle density $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$—denoted by $\bar{\mathsf{f}}_1$ in Sec. IV.2.1–IV.2.3—for the system with the properties presented in the previous section, where we only consider elastic two-to-two scattering processes.

---

(43) For indistinguishable colliding particles, the final states with scattering angle θ' defined as above or with the supplementary angle $\pi-\theta'$, which is the angle between $\vec{v}_2-\vec{v}_1$ and $\vec{v}_3-\vec{v}_4$, are one and the same.

(44) In due fairness, the invoked relation between $\ell_{\rm mfp}$ and $\sigma_{\rm tot}$ actually assumes that the system under study is dilute, so there is a kind of circular reasoning at play.

The derivation presented in this section is of a rather heuristic spirit, emphasizing the physical meaning of the terms on the right hand side of the Boltzmann equation. An alternative derivation, starting from the BBGKY hierarchy, is given in the appendix (IV.A) to this Chapter.

IV.2.1 General formulation

Since $\bar{\mathsf{f}}_1(t,\vec{r},\vec{p})$ is an instance of single-particle density—admittedly, on a coarse-grained version of space-time, yet this makes no difference here—, its evolution equation could be derived in the same manner as in the previous chapter III. Accordingly, in the absence of collisions $\bar{\mathsf{f}}_1$ obeys the single-particle Liouville equation [cf. (III.15)]

$$\frac{\partial\bar{\mathsf{f}}_1(t,\vec{r},\vec{p})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}_1(t,\vec{r},\vec{p}) + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}_1(t,\vec{r},\vec{p}) = 0.$$

This result can also be derived by counting particles within a volume $\mathrm{d}^3\vec{r}\ \mathrm{d}^3\vec{p}$ around point $(\vec{r},\vec{p})$ at time t, then by investigating where these particles are at a later time $t+\mathrm{d}t$, invoking Liouville's theorem (II.15) to equate the new volume they occupy to the old one.(45)

Traditionally, the influence of collisions on the evolution is expressed by introducing a symbolic *collision term* $(\partial\bar{\mathsf{f}}_1/\partial t)_{\text{coll.}}$ in the right member

$$\frac{\partial\bar{\mathsf{f}}_1}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}_1 + \vec{F}\cdot\vec{\nabla}_{\vec{v}}\bar{\mathsf{f}}_1 = \left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{coll.}}. \tag{IV.8}$$

The role of the collision term is to describe the change induced by scatterings in the number of particles $\bar{\mathsf{f}}_1(t,\vec{r},\vec{p})\,\mathrm{d}^3\vec{r}\ \mathrm{d}^3\vec{p}$ inside an infinitesimal volume element around point $(\vec{r},\vec{p})$.

The purpose of next subsection will be to give substance to this as yet empty notation. In particular, we shall split the collision term into a "gain term"—describing the particles that enter the volume $\mathrm{d}^3\vec{r}\ \mathrm{d}^3\vec{p}$ after a collision—and a "loss term"—corresponding to the particles which are scattered away from $\mathrm{d}^3\vec{r}\ \mathrm{d}^3\vec{p}$:

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{coll.}} \equiv \left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{gain}} - \left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{loss}}. \tag{IV.9}$$

Remark: In the same spirit as the right-hand side of Eq. (IV.8), one can designate the second and third terms of the left member as the rates of change of $\bar{\mathsf{f}}_1$ respectively caused by the motion of the particles and by the external force:

$$\frac{\partial\bar{\mathsf{f}}_1}{\partial t} = -\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{motion}} - \left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{force}} + \left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\text{coll.}}. \tag{IV.10}$$

IV.2.2 Computation of the collision term

We now derive the form of the two contributions to the collision term (IV.9), starting with the loss term.

IV.2.2 a Loss term

Consider a volume element $\mathrm{d}^3\vec{r}\ \mathrm{d}^3\vec{p}_1$ around a point $(\vec{r},\vec{p}_1)$ at time t. A particle inside this range can scatter on a partner also situated at $\vec{r}$—collisions are local—having a momentum $\vec{p}_2$ up to $\mathrm{d}^3\vec{p}_2$. After the collision, the outgoing particles have momenta $\vec{p}_3$, $\vec{p}_4$, with a probability related to the transition rate $w(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)$. Integrating over all possible final momenta yields the total scattering probability for initial particles with momenta $\vec{p}_1$, $\vec{p}_2$. Since $\mathrm{d}^3\vec{p}_1$ is infinitesimally small, any scattering process will give both colliding particles a different final momentum, so that any collision automatically leads to a decrease of the number of particles inside $\mathrm{d}^3\vec{p}_1$.

---

(45) This proof can for instance be found in Huang [33, Chap. 3.1] or Reif [35, Chap. 13.2].

To obtain the number of particles which are scattered away from $\mathrm{d}^3\vec{p}_1$, one has to multiply the transition rate per unit volume for the collision of one pair of particles with momenta $\vec{p}_1$, $\vec{p}_2$ by the total number of particles 1 and 2 per unit volume in the respective momentum ranges at time t. Very generally, this number is given by(46)

$$\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2)\,\mathrm{d}^3\vec{r}\;\frac{\mathrm{d}^3\vec{p}_1}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3},$$

with $\bar{\mathsf{f}}_2$ the (coarse-grained) joint two-particle density. Eventually, one integrates over all possible momenta $\vec{p}_2$ of the partner, which yields, after dividing by $\mathrm{d}^3\vec{r}\,\mathrm{d}^3\vec{p}_1/(2\pi\hbar)^3$

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\mathrm{loss}}(t,\vec{r},\vec{p}_1) = \int\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2)\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3}. \tag{IV.11a}$$

In terms of the differential cross section, this loss term reads

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\mathrm{loss}}(t,\vec{r},\vec{p}_1) = \int\bar{\mathsf{f}}^{(2)}(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2)\,\big|\vec{v}_2-\vec{v}_1\big|\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi')\,\mathrm{d}^2\Omega'\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}. \tag{IV.11b}$$

Remark: The integrand of the latter expression (IV.11b) actually involves quantities measured in different rest frames: $\vec{p}_1$, $\vec{p}_2$ are with respect to the rest frame in which the system is studied, while primed quantities are in the respective centre-of-momentum frames of the binary collisions—which, for a fixed $\vec{p}_1$, depend on $\vec{p}_2$!

IV.2.2 b Gain term

The gain term describes particles which at time t acquire the momentum $\vec{p}_1$ up to $\mathrm{d}^3\vec{p}_1$ in the final state of a collision. We thus need to consider scattering processes $\vec{p}_3,\vec{p}_4\to\vec{p}_1,\vec{p}_2$, where the values of the initial momenta and of $\vec{p}_2$ are irrelevant and thus will be integrated over.

For fixed $\vec{p}_3$, $\vec{p}_4$ and for a given $\vec{p}_2$ known up to $\mathrm{d}^3\vec{p}_2$, the number of particles with final momenta in the proper range for a unit number density of incoming particles is given by [cf. Eq. (IV.4b)]

$$\widetilde{w}(\vec{p}_3,\vec{p}_4\to\vec{p}_1,\vec{p}_2)\,\frac{\mathrm{d}^3\vec{p}_1}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}.$$

Multiplying by the two-particle distribution $\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4)$, which gives the density of particles with the respective momenta in the initial state, and integrating over these momenta as well as over the momentum $\vec{p}_2$ of the partner particle, one finds the number of "gained" particles per unit volume

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\mathrm{gain}}(t,\vec{r},\vec{p}_1)\frac{\mathrm{d}^3\vec{p}_1}{(2\pi\hbar)^3} = \frac{\mathrm{d}^3\vec{p}_1}{(2\pi\hbar)^3}\int\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4)\,\widetilde{w}(\vec{p}_3,\vec{p}_4\to\vec{p}_1,\vec{p}_2)\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3}.$$

Dividing both sides by $\mathrm{d}^3\vec{p}_1/(2\pi\hbar)^3$ and invoking the microreversibility property (IV.5c), this may be recast as

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\mathrm{gain}}(t,\vec{r},\vec{p}_1) = \int\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4)\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3}. \tag{IV.12a}$$

Equivalently, one may write

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\mathrm{gain}}(t,\vec{r},\vec{p}_1) = \int\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4)\,\big|\vec{v}_2-\vec{v}_1\big|\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi')\,\mathrm{d}^2\Omega'\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}, \tag{IV.12b}$$

where relation (IV.7) was used.

---

(46) The reader upset by the presence of the factor $\mathrm{d}^3\vec{r}$ despite the fact that we are interested in the number of pairs per unit volume may want to consider the number of pairs with both particles in the volume element $\mathrm{d}^3\vec{r}$, writing it first in the form

$$\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r}_2,\vec{p}_2)\,\mathrm{d}^3\vec{r}\,\mathrm{d}^3\vec{r}_2\,\frac{\mathrm{d}^3\vec{p}_1}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}$$

and then letting $\vec{r}_2=\vec{r}$—and accordingly $\mathrm{d}^3\vec{r}_2=\mathrm{d}^3\vec{r}$. The announced number of pairs per unit volume is then obtained by dividing by (a single factor of) $\mathrm{d}^3\vec{r}$.

IV.2.3 Closure prescription: molecular chaos

Gathering the loss and gain terms (IV.11a) and (IV.12a) together, the collision term, or *collision integral*, in the right-hand side of the Boltzmann equation reads

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\text{coll.}}\!(t,\vec{r},\vec{p}_1) = \int\Big[\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4) - \bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3}, \tag{IV.13a}$$

or equivalently [cf. Eqs (IV.11b) and (IV.12b)]

$$\left(\frac{\partial\bar{\mathsf{f}}_1}{\partial t}\right)_{\!\text{coll.}}\!(t,\vec{r},\vec{p}_1) = \int\Big[\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_3,\vec{r},\vec{p}_4) - \bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2)\Big]\,\big|\vec{v}_2-\vec{v}_1\big|\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi')\,\mathrm{d}^2\Omega'\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}. \tag{IV.13b}$$

As anticipated from the discussion of the BBGKY hierarchy in the previous chapter, the collision integral for the evolution of the single-particle density involves the two-particle density $\bar{\mathsf{f}}_2$. In turn, one can derive the collision term for the dynamics of the latter, which depends on $\bar{\mathsf{f}}_3$, and so forth.

Boltzmann's proposal was to transform the collision integral (IV.13a) into a term involving $\bar{\mathsf{f}}_1$ only, by invoking the assumption of *molecular chaos*, or *Stoßzahlansatz*, according to which the velocities of the particles *before the collision* are uncorrelated

$$\bar{\mathsf{f}}_2(t,\vec{r},\vec{p}_1,\vec{r},\vec{p}_2) = \bar{\mathsf{f}}_1(t,\vec{r},\vec{p}_1)\,\bar{\mathsf{f}}_1(t,\vec{r},\vec{p}_2)\quad\text{before a collision at instant } t. \tag{IV.14}$$

Just after a collision, two particles which have scattered on each other are correlated—inverting their velocities, one makes them collide, which is a rare event. Yet before they meet and collide again, they will undergo many scatterings with other, random particles, which wash out any trace of this correlation, and justifies the above assumption.

Remark: Molecular chaos is thus a weaker assumption that the factorization (III.23) in the Vlasov equation, which holds at any instant and for all positions of the two particles.

Under this assumption and inserting the resulting collision integral in the right-hand side of Eq. (IV.8), one obtains the *Boltzmann kinetic equation*[47]

$$\boxed{\begin{aligned}&\frac{\partial\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)}{\partial t} + \vec{v}_1\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1) + \vec{F}\cdot\vec{\nabla}_{\vec{p}_1}\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1) =\\ &\quad\left(1-\frac{\delta_{1,2}}{2}\right)\int\Big[\bar{\mathsf{f}}(t,\vec{r},\vec{p}_3)\bar{\mathsf{f}}(t,\vec{r},\vec{p}_4) - \bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)\bar{\mathsf{f}}(t,\vec{r},\vec{p}_2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3},\end{aligned}} \tag{IV.15a}$$

where the prefactor $1-\delta_{1,2}/2$ was introduced to ensure that the formula also holds without double counting when particles 1 and 2 are identical (in which case $\delta_{1,2}=1$, otherwise is $\delta_{1,2}=0$).
Equivalently, the Boltzmann equation may recast as

$$\begin{aligned}&\frac{\partial\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)}{\partial t} + \vec{v}_1\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1) + \vec{F}\cdot\vec{\nabla}_{\vec{p}_1}\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1) =\\ &\quad\left(1-\frac{\delta_{1,2}}{2}\right)\int\Big[\bar{\mathsf{f}}(t,\vec{r},\vec{p}_3)\bar{\mathsf{f}}(t,\vec{r},\vec{p}_4) - \bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)\bar{\mathsf{f}}(t,\vec{r},\vec{p}_2)\Big]\,\big|\vec{v}_2-\vec{v}_1\big|\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega'}(\theta',\varphi')\,\mathrm{d}^2\Omega'\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}.\end{aligned} \tag{IV.15b}$$

Remarks:

∗ One often introduces the abbreviations $\bar{\mathsf{f}}(1)\equiv\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)$, $\bar{\mathsf{f}}(2)\equiv\bar{\mathsf{f}}(t,\vec{r},\vec{p}_2)$, and so on, so that the collision integral is shortly written as

---

(47) From now on, we drop the notation $\bar{\mathsf{f}}_1$ and only use $\bar{\mathsf{f}}$.

$$\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int \Big[\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4) - \bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2 \to \vec{p}_3,\vec{p}_4)\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}\frac{\mathrm{d}^3\vec{p}_3}{(2\pi\hbar)^3}\frac{\mathrm{d}^3\vec{p}_4}{(2\pi\hbar)^3}. \tag{IV.15c}$$

To shorten expressions further, we shall also use $\int_{\vec{p}_i} \equiv \int \frac{\mathrm{d}^3\vec{p}_i}{(2\pi\hbar)^3}$, leading for instance to

$$\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big[\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4) - \bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2 \to \vec{p}_3,\vec{p}_4). \tag{IV.15d}$$

∗ The generalization of the Boltzmann equation to the case of a mixture of substances is straightforward: in the collision term for the evolution of the position-velocity-space density of a component, one has to sum the contribution from the (elastic two-to-two) scattering processes of the particles of that substance with each other—taking into account the $\frac{1}{2}$ factor to avoid double-counting—, and the contributions from collisions with particles of other components.

IV.2.4 Phenomenological generalization to fermions and bosons

The collision term of the Boltzmann equation can easily be modified so as to accommodate the Pauli exclusion principle between particles with half-integer spins.(48) Considering the two-to-two collision $\vec{p}_i,\vec{p}_j \to \vec{p}_k,\vec{p}_l$, where all particles are fermions,(at) the "repulsive" behavior of the latter can be phenomenologically accounted for by preventing the scattering process to happen when one of the final states $\vec{p}_k$ or $\vec{p}_l$ is already occupied. That is, one postulates that the rate for the process is not only proportional to the product $\bar{\mathsf{f}}(i)\bar{\mathsf{f}}(j)$ of the phase-space densities of the initial particles—where we use the same shorthand notation as in Eq. (IV.15c)—, but also to the product $[1-\bar{\mathsf{f}}(k)][1-\bar{\mathsf{f}}(l)]$ involving the densities of the final state particles. The collision integral of the Boltzmann equation thus reads

$$\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big(\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)\big[1-\bar{\mathsf{f}}(1)\big]\big[1-\bar{\mathsf{f}}(2)\big] - \bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\big[1-\bar{\mathsf{f}}(3)\big]\big[1-\bar{\mathsf{f}}(4)\big]\Big)\widetilde{w}(\vec{p}_1,\vec{p}_2 \to \vec{p}_3,\vec{p}_4). \tag{IV.16}$$

A similar generalization, which simulates the "attractive" character of bosons,(au) consists in enhancing the rate of the process $\vec{p}_i,\vec{p}_j \to \vec{p}_k,\vec{p}_l$, when there are already particles in the final state. This is done by multiplying the rate by the factor $[1+\bar{\mathsf{f}}(k)][1+\bar{\mathsf{f}}(l)]$, which yields

$$\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big(\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)\big[1+\bar{\mathsf{f}}(1)\big]\big[1+\bar{\mathsf{f}}(2)\big] - \bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\big[1+\bar{\mathsf{f}}(3)\big]\big[1+\bar{\mathsf{f}}(4)\big]\Big)\widetilde{w}(\vec{p}_1,\vec{p}_2 \to \vec{p}_3,\vec{p}_4). \tag{IV.17}$$

We shall see below that these seemingly ad hoc generalizations (IV.16)–(IV.17) lead for instance to the proper equilibrium distributions.

> An issue is naturally the actual meaning of $\bar{\mathsf{f}}$ in the generalized kinetic equations obtained with the above collision terms, since phase space is usually not considered as an interesting concept in quantum mechanics, where the Heisenberg uncertainties prevent a particle from being localized at a well-defined point in μ-space.

IV.2.5 Additional comments and discussions

Now that we have established the actual form of the Boltzmann equation, especially of its collision term, we wish to come back to the assumptions made in Sec. IV.1.1, to discuss their role in a new light.

---

(48) This idea seems to date back to Landau, in his work on the theory of Fermi liquids [36].

(at) E. Fermi, 1901–1954 (au) S. N. Bose, 1894–1974

An important point is the coarse graining of both time and position space. Thanks to it, the momenta of the colliding particles skip instantaneously from their initial values $\vec{p}_1$, $\vec{p}_2$ to the final ones, without going through intermediate values as would happen otherwise—except in the unrealistic case when the particles are modelled as hard spheres. If this transition were *not* instantaneous, particle 1—a similar reasoning holds for the other colliding particle (2), as well as for particles 3 and 4 in the gain term—would at the time t of the collision no longer have the momentum $\vec{p}_1$ it had "long" before the scattering. Accordingly, the distribution of particle 1 in the loss part of the collision integral should not be $\bar{\mathsf{f}}(t,\vec{r},\vec{p}_1)$, but rather one of the following possibilities:

- $\bar{\mathsf{f}}$ evaluated at time t, yet for the position $\vec{r}_1'$ and momentum $\vec{p}_1'$ of particle 1 at that very instant: in a classical description of the scattering process, $\vec{r}_1'$ and $\vec{p}_1'$ depend for instance on the impact parameter of the collision; whereas they are not even well-defined in a quantum mechanical description.
- $\bar{\mathsf{f}}$ evaluated at momentum $\vec{p}_1$, yet at a time $t-\tau$, before the collision, at which particle 1 still had this momentum, and accordingly at some position $\vec{r}_1 \neq \vec{r}$.

In the former case, one loses the locality in position space, while in the latter one has to abandon locality both in time and space. The advantage of adopting a coarse-grained description is thus to provide an evolution equation which is local both in t and $\vec{r}$, as is the case of Eq. (IV.15).

> Thanks to the time locality of the Boltzmann equation, the evolution of $\bar{\mathsf{f}}$ is "Markovian" in the wide sense of Sec. I.2.1, i.e. its rate of change is memoryless and only depends on $\bar{\mathsf{f}}$ at the same instant.

Another assumption is that the time scale on which the coarse graining is performed is much smaller than the average duration between two successive collisions of a particle, and similarly that the spatial size of the coarse-grained cells is much smaller than the mean free path. This allows one to meaningfully treat $\bar{\mathsf{f}}$ as a continuous—and even differentiable—function of t and $\vec{r}$, and thus amounts to assuming that the system properties do not change abruptly in time or spatially.

Eventually, one can note that the molecular chaos assumption (IV.14) provides a closed equation for $\bar{\mathsf{f}}$, yet at the cost of introducing nonlinearity, whereas the successive equations of the BBGKY hierarchy (III.14) are all linear.

IV.3 Balance equations derived from the Boltzmann equation

We now investigate various balance equations that hold in a system obeying the Boltzmann equation, beginning with conservation laws, then going on with the celebrated H-theorem. Motivated by this theorem, we then define various equilibrium distributions.

IV.3.1 Conservation laws

IV.3.1 a Properties of the collision integral

Let $\mathcal{I}_{\text{coll.}}(1,2,3,4)$ denote the integrand of the collision integral on the right-hand side of the Boltzmann equation (IV.15):

$$\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} \mathcal{I}_{\text{coll.}}(1,2,3,4) \qquad \text{(IV.18a)}$$

with

$$\mathcal{I}_{\text{coll.}}(1,2,3,4) \equiv \Big[\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4) - \bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4) \qquad \text{(IV.18b)}$$

in the case of "classical" particles—the expression in the case of fermions resp. bosons can be read at once off Eq. (IV.16) resp. (IV.17).

This integrand obeys various properties that appear when the role of the particles are exchanged, irrespective of whether they are indistinguishable or not.

- $(\mathfrak{S})$ The integrand in the collision term of the Boltzmann equation is symmetric under the exchanges $\vec{p}_1 \leftrightarrow \vec{p}_2$ and $\vec{p}_3 \leftrightarrow \vec{p}_4$.

  This symmetry is trivial.

- $(\mathfrak{A})$ The integrand in the collision term of the Boltzmann equation is antisymmetric under the simultaneous exchanges of $\vec{p}_1$, $\vec{p}_2$ with $\vec{p}_3$, $\vec{p}_4$.

  This property is straightforward when one considers on the one hand the (mathematical) change of labels $1 \leftrightarrow 3$, $2 \leftrightarrow 4$, which gives a minus sign, and on the other hand the physical microreversibility property encoded in Eq. (IV.5c).

Let $\chi(t,\vec{r},\vec{p})$ denote a *collisional invariant*, i.e. a microscopic quantity which is conserved in every binary collision. Examples are particle number, linear momentum, or kinetic energy since the collisions are elastic: the quantity can thus be scalar or vectorial. One then has the general identity

$$\boxed{\int \chi(t,\vec{r},\vec{p}) \left(\frac{\partial \bar{\mathsf{f}}}{\partial t}\right)_{\text{coll.}}\!(t,\vec{r},\vec{p})\, \frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^3} = 0.} \tag{IV.19}$$

Let us use the notations $\bar{\mathsf{f}}(1)$, $\bar{\mathsf{f}}(2)$, … as in Eq. (IV.15c), and accordingly $\chi(1) \equiv \chi(t,\vec{r},\vec{p}_1)$, $\chi(2) \equiv \chi(t,\vec{r},\vec{p}_2)$, and so on. Using Eq. (IV.18a) then gives

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} \chi(1)\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Exchanging first the dummy labels 1 and 2 of the integration variables, and invoking then the symmetry property $(\mathfrak{S})$ of the integrand of the collision term, one finds

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_2} \chi(2) \int_{\vec{p}_1}\int_{\vec{p}_3}\int_{\vec{p}_4} \mathcal{I}_{\text{coll.}}(2,1,3,4) = \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} \chi(2)\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Combining the previous two equations, we may thus write

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \frac{1}{2}\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} [\chi(1)+\chi(2)]\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Invoking now the antisymmetry $(\mathfrak{A})$ after exchanging the dummy indices $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$, one obtains

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_3} \chi(3) \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_4} \mathcal{I}_{\text{coll.}}(3,4,1,2) = -\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} \chi(3)\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Similarly the exchange $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$ yields:

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \int_{\vec{p}_4} \chi(4) \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_4} \mathcal{I}_{\text{coll.}}(4,3,2,1) = -\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} \chi(4)\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Both previous lines lead to

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = -\frac{1}{2}\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} [\chi(3)+\chi(4)]\mathcal{I}_{\text{coll.}}(1,2,3,4).$$

Gathering all intermediate results, there eventually comes

$$\int_{\vec{p}_1} \chi(1) \left(\frac{\partial \bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}} = \frac{1}{4}\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4} [\chi(1)+\chi(2)-\chi(3)-\chi(4)]\mathcal{I}_{\text{coll.}}(1,2,3,4) = 0,$$

where the last identity comes from the local conservation property $\chi(1)+\chi(2) = \chi(3)+\chi(4)$ expressing the invariance of χ under binary collisions. □

We can now replace χ by various conserved quantities.

IV.3.1 b Particle number conservation

As already mentioned, the (local) particle number density is the integral over momenta of the phase space distribution

$$\boxed{n(t,\vec{r}) \equiv \int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p}).} \tag{IV.20a}$$

On the other hand, the particle-number flux density is naturally given by

$$\boxed{\vec{\mathcal{J}}_N(t,\vec{r}) \equiv \int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\vec{v},} \tag{IV.20b}$$

where each "phase-space cell" contributes with its velocity $\vec{v}$, weighted with the corresponding distribution.

Integrating now the Boltzmann equation (IV.8) over $\vec{p}$, and exchanging the order of the derivatives with respect to time or space and of the integral over velocities, the first term in the left-hand side gives $\partial n/\partial t$, the second term equals $\vec{\nabla}_{\vec{r}} \cdot \vec{\mathcal{J}}_N$, while the third gives a vanishing contribution since $\bar{\mathsf{f}}$ vanishes at the boundaries of velocity space, to ensure the convergence of the integral in the normalization condition (III.2a). In turn, considering identity (IV.19) with the collisional invariant $\chi = 1$, the integral over $\vec{p}$ of the collision term vanishes. All in all, one obtains the local conservation law

$$\boxed{\frac{\partial n(t,\vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{\mathcal{J}}_N(t,\vec{r}) = 0,} \tag{IV.20c}$$

where we have dropped the now unnecessary subscript $\vec{r}$ on the gradient. This relation, known as *continuity equation*, is obviously of the general type (I.18). The immense progress is that n and $\vec{\mathcal{J}}_N$ can now be computed starting from a microscopic theory—if $\bar{\mathsf{f}}$ is known!—, instead of being postulated at the macroscopic level.

IV.3.1 c Energy conservation

As a second application, we can consider the collisional invariant $\chi(t,\vec{r},\vec{p}) = \vec{p}^{\,2}/2m$, i.e. the kinetic energy.(49) We introduce the local kinetic-energy density

$$\boxed{e_{\text{kin.}}(t,\vec{r}) = \int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\frac{\vec{p}^{\,2}}{2m}} \tag{IV.21a}$$

and the local kinetic-energy flux density

$$\boxed{\vec{\mathcal{J}}_{E_{\text{kin.}}}(t,\vec{r}) = \int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\frac{\vec{p}^{\,2}}{2m}\,\vec{v}.} \tag{IV.21b}$$

We first assume for simplicity that there is no external force acting on the particles.

Multiplying each term of the Boltzmann equation by $\vec{p}^{\,2}/2m$ and integrating over velocity, one finds the local conservation law

$$\boxed{\frac{\partial e_{\text{kin.}}(t,\vec{r})}{\partial t} + \vec{\nabla} \cdot \vec{\mathcal{J}}_{E_{\text{kin.}}}(t,\vec{r}) = 0,} \tag{IV.21c}$$

where identity (IV.19) has again been used.

---

(49) … in the case of neutral particles.

In the presence of an external force $\vec{F}$ independent from the particle momentum, a straightforward integration by parts shows that there comes an extra term—the integral over velocity of $\vec{F}\cdot\vec{v}$ multiplied by $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$—, with a + sign if it is written as right-hand side of Eq. (IV.21c). This trivially corresponds to the work exerted by the external force per unit time.

Remarks:

∗ Alternatively, one can take for χ the sum of the kinetic energy $\vec{p}^{\,2}/2m$ and the potential energy due to the external force. The resulting balance equation is then the local conservation of total energy, of the type (IV.21c) with different energy density and flux density, even in the presence of the external force.

∗ As already noted in Sec. IV.1.2, the conservation of kinetic energy alone is conserved is related to the assumption of a weakly-interacting system, in which the relative amount of potential energy is small.

IV.3.1 d Momentum balance

Eventually, let χ be the i-th component p_i of linear momentum. Let

$$\boxed{\vec{\mathsf{v}}(t,\vec{r}) \equiv \frac{1}{n(t,\vec{r})}\vec{\mathcal{J}}_N(t,\vec{r}) = \frac{1}{m\,n(t,\vec{r})}\int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\vec{p}} \tag{IV.22a}$$

be the local *flow velocity* and

$$\boxed{\big(\mathsf{J}_{\vec{P}}\big)^{ij}(t,\vec{r}) = \frac{1}{n(t,\vec{r})}\int_{\vec{p}} \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,p^i\,v^j} \tag{IV.22b}$$

the j-th component of the local flux density of the i-th component of linear momentum.

By integrating the Boltzmann equation multiplied by p^i over momentum, one easily finds

$$\boxed{\frac{\partial[m\,n(t,\vec{r})\mathsf{v}^i(t,\vec{r})]}{\partial t} + \sum_{j=1}^{3}\frac{\partial}{\partial x^j}\Big[n(t,\vec{r})\big(\mathsf{J}_{\vec{P}}\big)^{ij}(t,\vec{r})\Big] = n(t,\vec{r})\,F^i,} \tag{IV.22c}$$

which describes the rate of change of linear momentum under the influence of a force.
The property (IV.19) and the balance equations (IV.20)–(IV.22) will be recast in a different form in Sec. IV.6.1 below.

Remark: Inspecting Eqs. (IV.20b) and (IV.22a), one recognizes that the „local flow velocity“ is actually the average local velocity of particles, i.e. it is related to particle flow. In a relativistic theory, where particle number is not conserved—the corresponding scalar conserved quantity is rather a quantum number, like e.g. electric charge or baryon number—one may rather choose (after Landau) to define the flow velocity as the velocity derived from the flux of energy.

IV.3.2 H-theorem

A further consequence of the properties of the collision integral is the so-called H-theorem, which dates back to Boltzmann himself.

Given a solution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ of the Boltzmann equation, a macroscopic quantity $H(t)$ is defined by(50)

$$H(t) \equiv \int \bar{\mathsf{f}}(t,\vec{r},\vec{p}) \ln \bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\mathrm{d}^6\mathcal{V}, \tag{IV.23}$$

---

(50) $H(t)$ is not to be confused with the Hamilton function of the system...

with $\mathrm{d}^6\mathcal{V} \equiv \mathrm{d}^3\vec{r}\,\mathrm{d}^3\vec{p}/(2\pi\hbar)^3$ the measure on the (coarse-grained) single-particle phase space, as defined in Eq. (II.4b).

One also defines a related quantity, which we shall for the time being without further justification call *Boltzmann entropy*, as

$$S_B(t) \equiv k_B\!\int \bar{\mathsf{f}}(t,\vec{r},\vec{p})\big[1-\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big]\,\mathrm{d}^6\mathcal{V}, \tag{IV.24}$$

with k_B the Boltzmann constant.

We can check at once the simple relation

$$S_B(t) = -k_B H(t) + \text{constant}, \tag{IV.25}$$

where the unspecified constant is actually simply related to the total number of particles.

Remark: When adopting the generalizations (IV.16)–(IV.17) of the collision integral to fermions or bosons, one should accordingly modify the expressions of $H(t)$ and $S_B(t)$. For instance, one should consider

$$H(t) \equiv \int\Big\{\bar{\mathsf{f}}(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p}) \pm \big[1\mp\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big]\ln\big[1\mp\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big]\Big\}\,\mathrm{d}^6\mathcal{V}, \tag{IV.26}$$

where the upper (resp. lower) sign holds for fermions resp. bosons.

We can now turn to the *H-theorem*, which states that if just before a given time t_0 the system under study obeys the Boltzmann equation—and in particular fulfills the molecular chaos assumption—, then $H(t)$ decreases at time t_0

$$\boxed{\frac{\mathrm{d}H(t)}{\mathrm{d}t} \leq 0.} \tag{IV.27}$$

There ensues at once that the Boltzmann entropy is increasing

$$\boxed{\frac{\mathrm{d}S_B(t)}{\mathrm{d}t} \geq 0.} \tag{IV.28}$$

The proof of the H-theorem relies again on the properties of the collision integral. First, a straightforward differentiation yields, after exchanging the order of time derivative and integration over position and velocity,

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = \int\frac{\partial}{\partial t}\Big[\bar{\mathsf{f}}(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\Big]\mathrm{d}^6\mathcal{V} = \int\frac{\partial\bar{\mathsf{f}}(t,\vec{r},\vec{p})}{\partial t}\big[1+\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big]\,\mathrm{d}^6\mathcal{V}.$$

Since $\bar{\mathsf{f}}$ is a solution to the Boltzmann equation, the partial derivative $\partial\bar{\mathsf{f}}/\partial t$ can be rewritten with the help of Eq. (IV.8). The integral over $\vec{r}$ of the term $\vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}$ yields the difference of the values of $\bar{\mathsf{f}}$ at the boundaries of position space, where $\bar{\mathsf{f}}$ vanishes, so that the corresponding contribution is zero. The same reasoning and result hold for the integral over $\vec{p}$ of the term proportional to $\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}$—this is trivial if the force is velocity independent, and still holds when $\vec{F}$ depends on $\vec{v}$. All in all, one thus quickly finds

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = \int\bigg(\frac{\partial\bar{\mathsf{f}}}{\partial t}\bigg)_{\!\text{coll.}}\!(t,\vec{r},\vec{p})\big[1+\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big]\,\mathrm{d}^6\mathcal{V}. \tag{IV.29}$$

This shows that $H(t)$ does not evolve in the absence of collision: the decrease in $H(t)$ is entirely due to the scattering processes.

To deal with the remaining integral in Eq. (IV.29), one can first use property (IV.19) with $\chi = 1$, to get rid of the constant 1 in the angular brackets of the integrand. We are then left with

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = \int\bigg(\frac{\partial\bar{\mathsf{f}}}{\partial t}\bigg)_{\!\text{coll.}}\!(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\mathrm{d}^6\mathcal{V} = \int\bigg[\int_{\vec{p}_1}\bigg(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\bigg)_{\!\text{coll.}}\ln\bar{\mathsf{f}}(1)\bigg]\mathrm{d}^3\vec{r},$$

where the second identity follows from renaming the integration variable $\vec{p}$ as $\vec{p}_1$ and using the shorthand notations already introduced in the previous section.

As in the proof of relation (IV.19) one can find by writing explicitly the collision integral in terms of its integrand and using the symmetry properties of the latter and the change of labels $1 \leftrightarrow 2$ the identities

$$\begin{aligned}\int_{\vec{p}_1}\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}}\ln\bar{\mathsf{f}}(1) &= \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\ln\bar{\mathsf{f}}(1) = \int_{\vec{p}_2}\int_{\vec{p}_1}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(2,1,3,4)\ln\bar{\mathsf{f}}(2)\\ &= \int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\ln\bar{\mathsf{f}}(2)\\ &= \frac{1}{2}\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\big[\ln\bar{\mathsf{f}}(1)+\ln\bar{\mathsf{f}}(2)\big].\end{aligned}$$

Similarly one finds

$$\begin{aligned}\int_{\vec{p}_1}\left(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\right)_{\text{coll.}}\ln\bar{\mathsf{f}}(1) &= -\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\ln\bar{\mathsf{f}}(3) = -\int_{\vec{p}_2}\int_{\vec{p}_1}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\ln\bar{\mathsf{f}}(4)\\ &= -\frac{1}{2}\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\big[\ln\bar{\mathsf{f}}(3)+\ln\bar{\mathsf{f}}(4)\big].\end{aligned}$$

This eventually gives

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = \frac{1}{4}\int\left\{\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\mathcal{I}_{\text{coll.}}(1,2,3,4)\big[\ln\bar{\mathsf{f}}(1)+\ln\bar{\mathsf{f}}(2)-\ln\bar{\mathsf{f}}(3)-\ln\bar{\mathsf{f}}(4)\big]\right\}\mathrm{d}^3\vec{r}.$$

Replacing the integrand of the collision integral by its expression (IV.18b) and performing some straightforward algebra, one obtains

$$\begin{aligned}\frac{\mathrm{d}H(t)}{\mathrm{d}t} &= \frac{1}{4}\int\left\{\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big[\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)-\bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\ln\frac{\bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)}{\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)}\right\}\mathrm{d}^3\vec{r}\\ &= \frac{1}{4}\int\left\{\int_{\vec{p}_1}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)\left[1-\frac{\bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)}{\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)}\right]\ln\frac{\bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)}{\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)}\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4)\right\}\mathrm{d}^3\vec{r}.\end{aligned} \tag{IV.30}$$

Now, the integrand is always negative—since $(1-x)\ln x \leq 0$ for all x, while all other factors are positive—, which proves the H-theorem. □

Remarks:

∗ Boltzmann's contemporaries strongly objected to his ideas, and in particular to the H theorem, due to their incomplete understanding of its content. One of the objections was that the assumed invariance of interactions under time reversal, combined with the invariance of the equations of motion (for instance the Hamilton equations or the Liouville equation) under time reversal, should lead to the equivalence of both time directions, while the H-theorem selects a time direction.

The answer to this apparent paradox is that $H(t)$ is not decreasing at any time, but only when the system satisfies the assumption of molecular chaos. Actually, the existence of a preferred time direction was somehow postulated from the beginning by Boltzmann, when he made the difference between the state of the system before a collision (molecular chaos, the particles are uncorrelated) and after (the particles are then correlated). There is thus no inconsistency if $H(t)$ distinguishes between both time directions.

∗ Defining the quantities

$$h(t,\vec{r}) \equiv \int_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p}), \tag{IV.31a}$$

$$\vec{\mathcal{J}}_H(t,\vec{r}) \equiv \int_{\vec{p}}\vec{v}\,\bar{\mathsf{f}}(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p}), \tag{IV.31b}$$

and

$$\sigma_H(t,\vec{r}) \equiv \int_{\vec{p}} \left(\frac{\partial\bar{\mathsf{f}}}{\partial t}\right)_{\text{coll.}}\!(t,\vec{r},\vec{p})\big[1+\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p})\big] = \int_{\vec{p}} \left(\frac{\partial\bar{\mathsf{f}}}{\partial t}\right)_{\text{coll.}}\!(t,\vec{r},\vec{p})\ln\bar{\mathsf{f}}(t,\vec{r},\vec{p}), \tag{IV.31c}$$

one easily finds that they obey the local balance equation

$$\frac{\partial h(t,\vec{r})}{\partial t} + \vec{\nabla}\cdot\vec{\mathcal{J}}_H(t,\vec{r}) = \sigma_H(t,\vec{r}), \tag{IV.31d}$$

which suggests the interpretation of h, $\mathcal{J}_H$ and σ_H as density, flux density and source density(51) of the H-quantity, respectively.

IV.4 Solutions of the Boltzmann equation

We now turn to a discussion of the solution of the Boltzmann equation.

IV.4.1 Equilibrium distributions

According to the H-theorem, $H(t)$ is a decreasing function of time—at least when the assumption of molecular chaos holds. Besides, the defining formula (IV.23) shows that $H(t)$ is bounded from below when the spatial volume $\mathcal{V}$ occupied by the system, and accordingly its total energy, is finite.

This follows from the finiteness of the available phase space—the finite energy of the system translates into an upper bound on momenta—and the fact that the product $x\ln x$ is always larger than $-1/\mathrm{e}$.

As a consequence, $H(t)$ converges in the limit $t\to\infty$, i.e. its derivative vanishes:

$$\frac{\mathrm{d}H(t)}{\mathrm{d}t} = 0 \tag{IV.32}$$

or equivalently $\mathrm{d}S_B(t)/\mathrm{d}t = 0$. In this section, we define and discuss two types of distributions that fulfill this condition:

- the *global equilibrium distribution* $\bar{\mathsf{f}}_{\text{eq.}}(\vec{r},\vec{p})$ is defined as a stationary solution of the Boltzmann equation

$$\frac{\partial\bar{\mathsf{f}}_{\text{eq.}}}{\partial t} = 0 \tag{IV.33}$$

  obeying Eq. (IV.32). In that case, the system has reached global thermodynamic equilibrium, with uniform values of the temperature T and of the average velocity $\vec{\mathsf{v}}$ of particles, as well as of their number density $\mathcal{n}$ (or equivalently the chemical potential μ) when there is no external force. In particular, $\bar{\mathsf{f}}_{\text{eq.}}$ cancels the collision integral (IV.15c).

- *local equilibrium distributions*, hereafter denoted as $\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})$, cancel the collision term of the Boltzmann equation—and thereby obey condition (IV.32)—, yet are in general not solutions to the whole equation itself.

In practice, an out-of-equilibrium system first tends, under the influence of collisions alone, towards a state of local equilibrium. In a second step, the interplay of collisions and the drift terms drives the system towards global equilibrium.

Remarks:

$*$ If the spatial volume occupied by the system is unbounded, then H resp. S_B is not necessarily bounded from below resp. above, and thus a state of global equilibrium may never be reached. This is for instance the case for a collection of particles expanding initially confined into a small region of space left free to expand into the vacuum.

---

(51) In this case it might be more appropriate to call it sink density, since $H(t)$ is decreasing.

$*$ Even if the system is confined in a finite volume, it may still be "stuck" in a persistent non-equilibrium setup. An example is that of a gas performing undamped "monopole" oscillations in a spherically symmetric harmonic trap, as studied theoretically e.g. in Ref. [37] and realized experimentally (to a good approximation) with cold ^{87}Rb atoms [38].

IV.4.1 a Global equilibrium

For simplicity, we begin with the case when the system is homogeneous and there is no external force. Consistency then implies that the distribution function $\bar{\mathsf{f}}$ is independent of $\vec{r}$. This in particular holds for the equilibrium distribution, $\bar{\mathsf{f}}_{\rm eq.}(\vec{p})$, and the corresponding particle density n.

Inspecting the derivative (IV.30), one sees that the equilibrium condition $\mathrm{d}H(t)/\mathrm{d}t = 0$ can only be fulfilled if the always-negative integrand is actually vanishing itself, which implies the identity

$$\bar{\mathsf{f}}_{\rm eq.}(\vec{p}_1)\bar{\mathsf{f}}_{\rm eq.}(\vec{p}_2) = \bar{\mathsf{f}}_{\rm eq.}(\vec{p}_3)\bar{\mathsf{f}}_{\rm eq.}(\vec{p}_4),$$

for every quadruplet $(\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4)$ satisfying the kinetic-energy- and momentum-conservation relationships (IV.3) and allowed by the requirements on the system, in case the latter has a finite energy. From now on we assume that $\bar{\mathsf{f}}_{\rm eq.}$ is everywhere non-zero.(52) Taking the logarithm of the above relation gives

$$\ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p}_1) + \ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p}_2) = \ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p}_3) + \ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p}_4),$$

which expresses that $\ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p})$ is a quantity conserved in any elastic two-to-two collision, and is thus a linear combination of additive collision invariants $\chi(\vec{p})$. Assuming that particle number—amounting to $\chi(\vec{p}) = 1$—, momentum $[\chi(\vec{p}) = \vec{p}]$ and kinetic energy $[\chi(\vec{p}) = \vec{p}^{\,2}/2m]$ form a basis of the latter,(53) one can write $\ln \bar{\mathsf{f}}_{\rm eq.}(\vec{p}) = A' + \vec{B}\,' \cdot \vec{p} + C'\vec{p}^{\,2}$, with constant A', $\vec{B}\,'$, C'. Equivalently, one has (with new, related constants A, C and $\vec{p}_0$)

$$\bar{\mathsf{f}}_{\rm eq.}(\vec{p}) = C\,\mathrm{e}^{-A(\vec{p}-\vec{p}_0)^2}. \tag{IV.34}$$

Following relation (IV.2), the momentum-space integral of $\bar{\mathsf{f}}_{\rm eq.}(\vec{p})$ should yield the particle number density n, which necessitates $C = (4\pi\hbar^2 A)^{3/2} n$. The integral of the product $\bar{\mathsf{f}}_{\rm eq.}(\vec{p})\,\vec{p}$ with $\bar{\mathsf{f}}_{\rm eq.}$ of the form (IV.34) over the whole phase space then gives $\vec{p}_0$, which should equal the total momentum of the particles; the latter vanishes if the particles are studied in their global rest frame, yielding $\vec{p}_0 = \vec{0}$. Eventually, the momentum-space integral of $\bar{\mathsf{f}}_{\rm eq.}(\vec{p})\,\vec{p}^{\,2}/2m$ should give n times the average kinetic energy per particle, while Eq. (IV.34), together with the relation between C and A found above, leads to $3n/4mA$. Denoting the average kinetic energy per particle by $\frac{3}{2}k_BT$, one thus finds $A = 1/2mk_BT$. All in all, one obtains

$$\boxed{\bar{\mathsf{f}}_{\rm eq.}(\vec{p}) = n\left(\frac{2\pi\hbar^2}{mk_BT}\right)^{3/2} \mathrm{e}^{-\vec{p}^{\,2}/2mk_BT},} \tag{IV.35}$$

which is the Maxwell(av)–Boltzmann distribution.(54)

One easily checks that the latter actually also cancels the left-hand side of the Boltzmann equation in the absence of external force, i.e. it represents the global equilibrium distribution.

---

(52) The identically vanishing distribution $\bar{\mathsf{f}} \equiv 0$ is clearly stationary and does fulfill condition (IV.32), yet it of course leads to a vanishing particle number, i.e. it represents a very boring system!

(53) A proof was given by Grad [39], which is partly reproduced by Sommerfeld [40, § 42]. An alternative argument is that if there were yet another kinematic collisional invariant in two-to-two scatterings, it would yield a fifth condition on the outgoing momenta—the requirement of having two particles in the final state leaves the constraints from energy and momentum conservation—, and thus restrict their angles, which is not the case.

(54) possibly up to the normalization factor, since in these lecture notes $\bar{\mathsf{f}}$ is dimensionless and normalized to the total number of particles, while the "usual" Maxwell–Boltzmann distribution is rather a probability distribution.

(av) J. C. Maxwell, 1831–1879

Remark: The Maxwell–Boltzmann distribution (IV.35) was probably known to the reader as the single-particle phase-space distribution of a classical ideal monoatomic gas at thermodynamic equilibrium studied in the canonical ensemble. Here it emerges as the (almost) universal long-time limit of the out-of-equilibrium evolution of a Boltzmann gas, in which the average energy per particle was "arbitrarily" denoted by $\frac{3}{2}k_BT$, where T has no other interpretation. This constitutes a totally different vision, yet the fact that both approaches yield the same distribution hints at the consistency of their results.

In case the system is subject to an external force deriving from a time-independent scalar potential, $\vec{F} = -\vec{\nabla}V(\vec{r})$, the stationary equilibrium solution of the Boltzmann equation is no longer spatially homogeneous, but becomes

$$\bar{\mathsf{f}}_{\text{eq.}}(\vec{r},\vec{p}) = n(\vec{r})\left(\frac{2\pi\hbar^2}{mk_BT}\right)^{3/2} \mathrm{e}^{-\vec{p}^{\,2}/2mk_BT} \quad \text{with} \quad n(\vec{r}) \equiv \frac{N\,\mathrm{e}^{-V(\vec{r})/k_BT}}{\displaystyle\int \mathrm{e}^{-V(\vec{r})/k_BT}\,\mathrm{d}^3\vec{r}}. \tag{IV.36}$$

Boltzmann entropy of the global equilibrium distribution

Inserting the equilibrium distribution (IV.36) into the Boltzmann entropy (IV.24), one finds with the help of Eq. (A.1c)

$$S_B = Nk_B \ln\left[\frac{\mathcal{V}}{N}\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}\right] + \frac{5}{2}Nk_B, \tag{IV.37}$$

where $\mathcal{V}$ is the volume occupied by the particles. This result coincides with the *Sackur*[aw]*–Tetrode*[ax] *formula* [41, 42] for the thermodynamic entropy of a classical ideal gas under the same conditions. This shows that at equilibrium the Boltzmann entropy, which is also defined out of equilibrium, coincides with the thermodynamic entropy.

IV.4.1 b Local equilibrium

A non-equilibrated macroscopic system of weakly-interacting classical particles will not reach the global equilibrium distribution (IV.35) at once, the approach to equilibrium can be decomposed in "successive" (not in a strict chronological sense, but rather logically) steps involving various time scales [cf. Eqs. (III.16), (III.17)]:

i. Over the shortest time scale τ_c, neighbouring particles scatter on each other. These collisions—which strictly speaking are *not* described by the Boltzmann equation—lead to the emergence of molecular chaos.

ii. Once molecular chaos holds, scatterings tend to drive the system to a state of local equilibrium, described by local thermodynamic variables as defined in chapter I: particle number density $n(t,\vec{r})$, temperature $T(t,\vec{r})$ and average particle velocity $\vec{\mathsf{v}}(t,\vec{r})$. That is, the single-particle distribution $\bar{\mathsf{f}}$ relaxes to a *local equilibrium distribution*, of the form

$$\boxed{\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p}) = n(t,\vec{r})\left[\frac{2\pi\hbar^2}{mk_BT(t,\vec{r})}\right]^{3/2} \exp\left\{-\frac{\left[\vec{p}-m\vec{\mathsf{v}}(t,\vec{r})\right]^2}{2mk_BT(t,\vec{r})}\right\}}, \tag{IV.38}$$

which cancels the collision integral on the right-hand side of the Boltzmann equation. This relaxation takes place over a time scale τ_r, which physically should be (much) larger than τ_c since it involves "many" local scatterings.

iii. Eventually, the state of local equilibrium relaxes to that of global equilibrium over the time scales τ_s, τ_e defined in Eq. (III.16). This step in the evolution of the single-particle distribution $\bar{\mathsf{f}}$, which is now rather driven by the slow drift terms on the left-hand side of the Boltzmann

---

[aw] H. M. Sackur, 1895–1931 [ax] O. Tetrode, 1880–1914

equation, is more conveniently described in terms of the evolution of local thermodynamic quantities obeying macroscopic equations, in particular those of hydrodynamics, as will be described in Sec. IV.6.

Remark: In general, the distribution (IV.38) does not cancel the left-hand side of the Boltzmann equation and is thus *not* a solution of the equation. This is however not a problem, since the single-particle distribution $\bar{f}$ never really equals a given $\bar{f}^{(0)}$, it only tends momentarily towards it: the three steps which above have been described as successive actually take place simultaneously.

IV.4.1 c Equilibrium distributions for bosons and fermions

The kinetic equations with a collision integral which simulates either the Pauli[(ay)] principle of particles with half-integer spin [Eq. (IV.16)] or the “gregariousness” of particles with integer spin [Eq. (IV.17)] also obey an H-theorem, and thus tend at large time to respective stationary equilibrium distributions $\bar{f}^{\rm F}_{\rm eq.}$ and $\bar{f}^{\rm B}_{\rm eq.}$.

Repeating a reasoning analogous to that in paragraph IV.4.1 a, with $\bar{f}_{\rm eq.}$ replaced by $\bar{f}^{\rm F}_{\rm eq.}/(1-\bar{f}^{\rm F}_{\rm eq.})$ or $\bar{f}^{\rm B}_{\rm eq.}/(1+\bar{f}^{\rm B}_{\rm eq.})$, one finds after some straightforward algebra

$$\boxed{\bar{f}^{\rm F}_{\rm eq.}(\vec{p}) = \frac{1}{C^{\rm F}\exp(\vec{p}^{\,2}/2mk_BT)+1}} \tag{IV.39a}$$

and

$$\boxed{\bar{f}^{\rm B}_{\rm eq.}(\vec{p}) = \frac{1}{C^{\rm B}\exp(\vec{p}^{\,2}/2mk_BT)-1},} \tag{IV.39b}$$

with $C^{\rm F}$, $C^{\rm B}$ two constants ensuring the proper normalization of the distributions. These are respectively the Fermi–Dirac and Bose–Einstein distributions as found within equilibrium statistical mechanics for ideal quantum gases.

IV.4.2 Approximate solutions

The Boltzmann equation (IV.15) identifies the rate of change of the single-particle distribution $\bar{f}(t,\vec{r},\vec{p})$ to a collision integral, which accounts for microscopic two-body elastic collisions. It is a non-linear partial integro-differential equation, and as thus especially difficult to solve. Accordingly, few exact solutions—apart from the equilibrium Maxwell–Boltzmann distribution (IV.35)–(IV.36)—are known analytically. Instead, one has to resort to approximate solutions, based on various schemes, two of which are shortly presented here.

IV.4.2 a Orders of magnitude

As already mentioned in § III.2.2 b–III.2.2 c, the various terms of the Boltzmann equation a priori involve different time or length scales. Introducing the typical size $\bar{f}_c$ of $\bar{f}$ (at a given momentum), one may introduce characteristic times $\tau_{\rm evol.}$, $\tau_{\rm s}$, $\tau_{\rm e}$, and $\tau_{\rm coll.}$ such that the dimensionless terms

$$\frac{\tau_{\rm evol.}}{\bar{f}_c}\frac{\partial\bar{f}}{\partial t} \quad,\quad \frac{\tau_{\rm s}}{\bar{f}_c}\vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{f} \quad,\quad \frac{\tau_{\rm e}}{\bar{f}_c}\vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{f} \quad,\quad \frac{\tau_{\rm coll.}}{\bar{f}_c}\left(\frac{\partial\bar{f}}{\partial t}\right)_{\rm coll.}$$

are all of order unity. $\tau_{\rm e}$—imposed by the external force—is generally of the order of $\tau_{\rm s}$ or larger, and will no longer be considered in the following discussion.

- By construction, $\tau_{\rm evol.}$ is of the order of the inverse rate of change of $\bar{f}/\bar{f}_c$.

[(ay)] W. Pauli, 1900–1958

- $\tau_{\rm s}$ is of the order of the characteristic length L over which the macroscopic system properties vary divided by the typical particle velocity v_c in the system:
$$\tau_{\rm s} \sim \frac{L}{v_c}. \tag{IV.40}$$
- Starting from expression (IV.13b) of the collision integral in terms of the differential cross section, one finds that the collision term is of order
$$\left(\frac{\partial\bar{\mathsf{f}}}{\partial t}\right)_{\rm coll.} \sim n\bar{\mathsf{f}}_c\, v_c\, \sigma_{\rm tot}, \tag{IV.41}$$
where the number density n comes from integrating $\bar{\mathsf{f}}$ over momentum, while the total cross section $\sigma_{\rm tot.}$ results from the integration of the differential cross section over all possible directions. Accordingly, one finds
$$\tau_{\rm coll.} \sim \frac{1}{n\, v_c\, \sigma_{\rm tot}}. \tag{IV.42a}$$
As could have been argued on physical grounds, this time scale associated with the collision term is of the same order as the mean free time between two successive scatterings of a given particle. In terms of the mean free path $\ell_{\rm mfp} \sim 1/n\sigma_{\rm tot}$, one may write
$$\tau_{\rm coll.} \sim \frac{\ell_{\rm mfp}}{v_c} \tag{IV.42b}$$

With the three above time scales one may construct two independent dimensionless numbers, namely on the one hand the *Knudsen*[(az)] *number* Kn
$$\mathrm{Kn} \equiv \frac{\ell_{\rm mfp}}{L} \tag{IV.43}$$
or equivalently $\mathrm{Kn} \sim \tau_{\rm coll.}/\tau_{\rm s}$, which compares the collision integral and the drift term; and on the other hand the ratio
$$\frac{\tau_{\rm coll.}}{\tau_{\rm evol.}} \sim \frac{\ell_{\rm mfp}}{v_c\,\tau_{\rm evol.}} = \frac{L}{v_c\,\tau_{\rm evol.}}\mathrm{Kn},$$
in which the prefactor multiplying the Knudsen number is sometimes referred to as *Strouhal*[(ba)] *number*.

The two dimensionless parameters can take a priori (almost) any value. A small Knudsen number thus corresponds to a rather dense system, in which collision occur relatively often; while a large value of Kn corresponds to the free-streaming limit. In turn, small values of $\tau_{\rm coll.}/\tau_{\rm evol.}$ correspond to slow evolutions of $\bar{\mathsf{f}}$—thus $\tau_{\rm evol.} \to \infty$, yielding $\tau_{\rm coll.}/\tau_{\rm evol.} \to 0$, in stationary systems. On the other hand, on physical grounds $\tau_{\rm coll.}/\tau_{\rm evol.}$ cannot be much larger than Kn, since this would imply the existence of an emergent "macroscopic" velocity scale $L/\tau_{\rm evol.}$ much larger than the characteristic velocity v_c of the particles, which is hard to conceive.

Taking for simplicity $\tau_{\rm coll.}/\tau_{\rm evol.}$ of order Kn, so has to have a single dimensionless quantity in the system, approximate solutions can be found in a systematic way when the Knudsen number is either very small, $\mathrm{Kn} \ll 1$, or very large ($\mathrm{Kn} \gg 1$), since in either case there is a small parameter in the problem—either Kn or Kn^{-1}—, suggesting the use of perturbative expansions in that parameter.

Remark: The case of very small Knudsen numbers has to be taken with a grain of salt: to remain within the region of validity of the Boltzmann equation as a physically motivated model, the mean free path cannot become infinitely small, but has to remain (much) larger than the interaction range (Sec. IV.1.1). Kn can thus only be very small by letting the system size become large—which is perfectly acceptable.

[(az)] M. Knudsen, 1871–1949 [(ba)] V. Strouhal, 1850–1922

IV.4.2 b Hilbert expansion

Let us discuss a first instance of perturbative expansion for the solution of the Boltzmann equation, based on the assumption that the Knudsen number (IV.43) is small—and that the ratio $\tau_{\rm coll.}/\tau_{\rm evol.}$ is of the same order of magnitude.

Throughout this paragraph, we shall omit the $(t,\vec{r},\vec{p})$ variables. In addition, we introduce two new notations: the collision integral on the right hand side of the Boltzmann equation will be denoted by

$$\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}},\bar{\mathsf{f}}\big] \equiv \bigg(\frac{\partial\bar{\mathsf{f}}(1)}{\partial t}\bigg)_{\rm coll.} = \int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big[\bar{\mathsf{f}}(3)\bar{\mathsf{f}}(4)-\bar{\mathsf{f}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4). \tag{IV.44a}$$

Besides, we introduce the bilinear operator

$$\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}},\bar{\mathsf{g}}\big] \equiv \frac{1}{2}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\Big[\bar{\mathsf{f}}(3)\bar{\mathsf{g}}(4)+\bar{\mathsf{g}}(3)\bar{\mathsf{f}}(4)-\bar{\mathsf{f}}(1)\bar{\mathsf{g}}(2)-\bar{\mathsf{g}}(1)\bar{\mathsf{f}}(2)\Big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4). \tag{IV.44b}$$

The functional space on which $\mathcal{C}_{\rm coll.}$ is operating will be specified below.

To account for the physical interpretation according to which the left hand side of the Boltzmann equation is governed by a larger time scale than the right hand side—which corresponds to having a small Knudsen number—, we rewrite the equation with a dimensionless parameter ε, which will be treated as much smaller than 1, in front of the former:

$$\varepsilon\bigg(\frac{\partial\bar{\mathsf{f}}}{\partial t}+\vec{v}\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}+\vec{F}\cdot\vec{\nabla}_{\vec{p}}\,\bar{\mathsf{f}}\bigg) = \mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}},\bar{\mathsf{f}}\big]. \tag{IV.45a}$$

In turn, the phase space distribution $\bar{\mathsf{f}}$ solution to this equation is written as an expansion in powers of ε:

$$\bar{\mathsf{f}} = \sum_{n=0}^{\infty}\varepsilon^n\,\bar{\mathsf{f}}^{(n)}. \tag{IV.45b}$$

The reader should view this expansion as a formal one. The underlying idea is clearly that by knowing enough of the first terms, one can approximate the true solution (for a given problem, e.g. with specified initial and boundary conditions) with a high precision. Yet the mathematical problem of the convergence of the expansion—for which values of ε? in which phase-space region? what type of convergence?—depends on the inter-particle interactions, and is beyond the scope of these notes.

With the ansatz (IV.45b), known as the *Hilbert expansion*, the collision term (IV.44a) reads

$$\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}},\bar{\mathsf{f}}\big] = \sum_{n=0}^{\infty}\varepsilon^n\,\mathcal{C}^{(n)} \quad\text{with}\quad \mathcal{C}^{(n)} \equiv \sum_{k=0}^{n}\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}}^{(k)},\bar{\mathsf{f}}^{(n-k)}\big], \tag{IV.45c}$$

where the expression of $\mathcal{C}^{(n)}$ follows from some straightforward algebra. If the "leading term" $\bar{\mathsf{f}}^{(0)}$ is an equilibrium distribution, global or local, it cancels the collision integral, yielding

$$\mathcal{C}^{(0)} = \mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}}^{(0)},\bar{\mathsf{f}}^{(0)}\big] = 0. \tag{IV.45d}$$

In turn, the higher-order $\mathcal{C}^{(n)}$ are conveniently rewritten as

$$\mathcal{C}^{(n)} = 2\,\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}}^{(0)},\bar{\mathsf{f}}^{(n)}\big] + \sum_{k=1}^{n-1}\mathcal{C}_{\rm coll.}\big[\bar{\mathsf{f}}^{(k)},\bar{\mathsf{f}}^{(n-k)}\big] \quad\text{for } n\geq 1. \tag{IV.45e}$$

Inserting now the Hilbert expansion (IV.45b) and the collision term (IV.45c) in the Boltzmann equation leads at once to

$$\sum_{n=0}^{\infty}\varepsilon^{n+1}\bigg(\frac{\partial\bar{\mathsf{f}}^{(n)}}{\partial t}+\vec{v}\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}^{(n)}+\vec{F}\cdot\vec{\nabla}_{\vec{p}}\,\bar{\mathsf{f}}^{(n)}\bigg) = \sum_{n=0}^{\infty}\varepsilon^n\,\mathcal{C}^{(n)},$$

i.e. after identifying the factors multiplying the term ϵ^n

$$\frac{\partial\bar{\mathsf{f}}^{(n-1)}}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}^{(n-1)} + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}^{(n-1)} = \mathcal{C}^{(n)} \quad \text{for } n\geq 1. \tag{IV.46}$$

Invoking relation (IV.45e) to rewrite the right member of this equation and reorganizing the terms, one eventually obtains

$$\mathcal{C}_{\text{coll.}}\big[\bar{\mathsf{f}}^{(0)},\bar{\mathsf{f}}^{(n)}\big] = \frac{1}{2}\Bigg(\frac{\partial\bar{\mathsf{f}}^{(n-1)}}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}^{(n-1)} + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}^{(n-1)} - \sum_{k=1}^{n-1}\mathcal{C}_{\text{coll.}}\big[\bar{\mathsf{f}}^{(k)},\bar{\mathsf{f}}^{(n-k)}\big]\Bigg) \text{ for } n\geq 1. \tag{IV.47}$$

The right member only involves the functions $\bar{\mathsf{f}}^{(k)}$ with $k<n$, and accordingly is entirely determined when these functions are known. In turn, the term on the left hand side of Eq. (IV.47) is at fixed $\bar{\mathsf{f}}^{(0)}$ a linear functional in $\bar{\mathsf{f}}^{(n)}$. Inverting the corresponding operator thus allows one in principle to obtain $\bar{\mathsf{f}}^{(n)}$ building on the previous determination of the functions $\{\bar{\mathsf{f}}^{(k)}\}$ with $k<n$, thereby yielding a systematic, sequential method.

More precisely, it is customary to rewrite the successive unknown function as $\bar{\mathsf{f}}^{(n)} = \bar{\mathsf{f}}^{(0)}h^{(n)}$. Inserting this form in the operator (IV.44b) then yields

$$\mathcal{C}_{\text{coll.}}\big[\bar{\mathsf{f}}^{(0)},\bar{\mathsf{f}}^{(0)}h^{(n)}\big] = \frac{\bar{\mathsf{f}}^{(0)}(1)}{2}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\bar{\mathsf{f}}^{(0)}(2)\big[h^{(n)}(3)+h^{(n)}(4)-h^{(n)}(1)-h^{(n)}(2)\big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4),$$

where the identity $\bar{\mathsf{f}}^{(0)}(1)\,\bar{\mathsf{f}}^{(0)}(2) = \bar{\mathsf{f}}^{(0)}(3)\,\bar{\mathsf{f}}^{(0)}(4)$ has been used. Since both sides of Eq. (IV.47) can without difficulty be divided by the factor $\bar{\mathsf{f}}^{(0)}(1)$, the interest now lies in the linearized collision operator

$$\mathcal{C}^{\text{lin.}}_{\text{coll.}}[h] \equiv \frac{1}{2}\int_{\vec{p}_2}\int_{\vec{p}_3}\int_{\vec{p}_4}\bar{\mathsf{f}}^{(0)}(2)\big[h(3)+h(4)-h(1)-h(2)\big]\,\widetilde{w}(\vec{p}_1,\vec{p}_2\to\vec{p}_3,\vec{p}_4), \tag{IV.48}$$

where the functions h are assumed to be sufficiently regular to ensure the existence of the integral. By introducing an inner product

$$(h_1,h_2) \equiv \int_{\vec{p}}\bar{\mathsf{f}}^{(0)}h_1h_2,$$

the space of functions h becomes a Hilbert space $\mathscr{H}$. This simplifies the discussion of the properties of the operator $\mathcal{C}^{\text{lin.}}_{\text{coll.}}$, which is easily shown to be self-adjoint.

A problem with inverting the linearized collision operator so as to determine $\bar{\mathsf{f}}^{(n)}$, or equivalently $h^{(n)}$, from the $\{\bar{\mathsf{f}}^{(k)}\}$ (or $\{h^{(k)}\}$) with $k<n$ is that 0 is actually a fivefold degenerate eigenvalue—the corresponding eigenfunctions being the collisional invariant functions $\chi=1$, $\vec{p}$ and $\vec{p}^{\,2}/2m$. The solution is to invert $\mathcal{C}^{\text{lin.}}_{\text{coll.}}$ in the subspace of $\mathscr{H}$ orthogonal to that spanned by these eigenfunctions, which however means that each $h^{(n)}$ is only known up to a linear combination, with coefficients depending possibly on time and position, of the collisional invariants. The coefficients entering the expression of $h^{(n)}$ are actually fixed at the following step, when requiring that the term on the right hand side of Eq. (IV.47) for the determination of $\bar{\mathsf{f}}^{(n+1)}$ is in the proper space.

IV.4.2 c Orthogonal polynomial solutions

will be added later.

IV.4.3 Relaxation-time approximation

It may happen that the inter-particle interactions entering the collision integral of the Boltzmann equation are only imperfectly known, so that the approximation methods presented in the previous section IV.4.2 cannot be used. In order to still be able to derive qualitative behaviors in that situation, it is customary to perform a rather drastic approximation of the collision term, based on the physical significance of the latter, whose role is to let the distribution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ relax to a local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$, given by expression (IV.38), with local density, temperature and average particle velocity which are specific to the conditions imposed on the system.

In the so-called *relaxation-time approximation*, it is assumed that the approach to the local equilibrium distribution is exponential, with a characteristic time scale $\tau_{\mathrm{r}}(\vec{r},\vec{p})$. In that case, the Boltzmann equation is approximated by the linear partial differential equation

$$\boxed{\frac{\partial\bar{\mathsf{f}}(t,\vec{r},\vec{p})}{\partial t}+\vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})+\vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})=-\frac{\bar{\mathsf{f}}(t,\vec{r},\vec{p})-\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})}{\tau_{\mathrm{r}}(\vec{r},\vec{p})}.}\tag{IV.49}$$

This approximation thus leaves the integral aspect of the collision term, amounting to a “first linearization” of the Boltzmann equation.

Remarks:

∗ The *relaxation time* τ_{r} models the typical duration for reaching local equilibrium (see § IV.4.1 b). In Eq. (IV.49) it is introduced as a free parameter. It should nevertheless naturally be larger than the *mean free time*, i.e. the average “free-flight” time between two successive scatterings of a given particle, since one expects that equilibration requires several collisions. On the other hand, it should remain (much) smaller than “external” time scales imposed by macroscopic conditions on the systems.

∗ τ_{r} is often assumed to be uniform over the whole system—which is to a large extent justified as long as the spatial density does not vary too much—, yet still momentum dependent.

∗ When τ_{r} is also assumed to be independent of $\vec{p}$—which is far less obvious—, the approximation is sometimes referred to as the Bhatnagar[bb]–Gross[bc]–Krook[bd] (BGK) approximation [43].

∗ The approximation (IV.49) may be seen as a truncation prescription of the BBGKY hierarchy (III.14), irrespective of the Boltzmann equation.

If the departure from equilibrium remains always small, then the true solution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ to the Boltzmann equation never deviates much from a local equilibrium distribution $\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})$. One may then write

$$\bar{\mathsf{f}}(t,\vec{r},\vec{p})=\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})+\bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})\quad\text{with}\quad\big|\bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})\big|\ll\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})\quad\forall\,t,\vec{r},\vec{p}.\tag{IV.50}$$

The right-hand side of the simplified equation (IV.49) is then simply $-\bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})/\tau_{\mathrm{r}}(\vec{p})$, while on the left-hand side one will account for the assumption $|\bar{\mathsf{f}}^{(1)}|\ll\bar{\mathsf{f}}^{(0)}$ by linearizing the various terms with respect to the small perturbation $\bar{\mathsf{f}}^{(1)}$. This constitutes the “second linearization” of the Boltzmann equation, and will be illustrated on a few examples in the following section.

IV.5 Computation of transport coefficients

In this section, we show how the Boltzmann equation allows the computation of transport coefficients in a system which is slightly out of equilibrium, working within the relaxation-time approximation introduced in Sec. IV.4.3.

[bb] P. L. Bhatnagar, 1912–1976 [bc] E. P. Gross, 1926–1991 [bd] M. Krook, 1913–1985

IV.5.1 Electrical conductivity

As first example, consider a gas of charged particles of mass m, in the presence of a constant and uniform electric field $\vec{\mathscr{E}}$. These particles diffuse among a system of particles of a different species, which we shall not consider explicitly. As a result of the various possible types of (elastic) two-to-two collisions between the different particle species, the single-particle distribution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ obeys the Boltzmann equation (IV.15), with the external force $\vec{F} = q\vec{\mathscr{E}}$ in the drift term, where q is the electric charge of the particles under study.

We shall assume that we are in a regime where the Boltzmann equation can be approximated by Eq. (IV.49) while the single-particle distribution can be written in the form (IV.50). For reasons which will be made clearer below, we search for a stationary particle distribution. Given the uniformity and stationarity of the electric field, it is consistent to consider a local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$ involving time- and position-independent local quantities n, T, $\vec{\mathsf{v}}$. $\bar{\mathsf{f}}^{(0)}$ then only depends on momentum:

$$\bar{\mathsf{f}}^{(0)}(\vec{p}) = n\left(\frac{2\pi\hbar^2}{mk_BT}\right)^{3/2} \mathrm{e}^{-(\vec{p}-m\vec{\mathsf{v}})^2/2mk_BT}.$$

Assuming that the average velocity $\vec{\mathsf{v}}$ of the moving charges vanishes—which is only a matter of choosing the proper reference frame, since it is uniform across the system—, this local equilibrium distribution reduces to

$$\bar{\mathsf{f}}^{(0)}(\vec{p}) = n\left(\frac{2\pi\hbar^2}{mk_BT}\right)^{3/2} \mathrm{e}^{-\vec{p}^{\,2}/2mk_BT} \tag{IV.51}$$

i.e. actually to the Maxwell–Boltzmann distribution (IV.35). Given the homogeneity of the system, the relaxation time τ_{r} can be taken to be independent of position.

Inserting $\bar{\mathsf{f}}(\vec{r},\vec{p}) = \bar{\mathsf{f}}^{(0)}(\vec{p}) + \bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p})$ in Eq. (IV.49), the leading remaining term on the left-hand side is that involving $\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}^{(0)}(\vec{p})$: there is no term involving either the time or space derivative of $\bar{\mathsf{f}}^{(0)}$—since $\bar{\mathsf{f}}^{(0)}$ does not depend on either—, and the terms involving gradients of $\bar{\mathsf{f}}^{(1)}$ are of subleading order—physically, the corresponding time scales [cf. Eq. (III.16)] are much larger than $\tau_{\mathrm{r}}(\vec{p})$, since many more collisions are needed for smoothing out large-scale inhomogeneities than small-scale ones. All in all, Eqs. (IV.49) and (IV.50) thus give to leading order

$$q\vec{\mathscr{E}}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}^{(0)}(\vec{p}) = -\frac{\bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p})}{\tau_{\mathrm{r}}(\vec{p})}.$$

from where one reads off $\bar{\mathsf{f}}^{(1)}$ at once:

$$\bar{\mathsf{f}}^{(1)}(\vec{p}) = \frac{q}{mk_BT}\tau_{\mathrm{r}}(\vec{p})\,\bar{\mathsf{f}}^{(0)}(\vec{p})\,\vec{p}\cdot\vec{\mathscr{E}} \tag{IV.52}$$

In particular, $\bar{\mathsf{f}}^{(1)}$ is independent of $\vec{r}$.

Now, the electric current (density) is the sum over all carriers of their electric charge multiplied by their velocity

$$\vec{J}_{\mathrm{el.}}(t,\vec{r}) = \int_{\vec{p}} q\vec{v}\,\bar{\mathsf{f}}(t,\vec{r},\vec{p}) = \frac{q}{m}\int_{\vec{p}}\vec{p}\,\bar{\mathsf{f}}(t,\vec{r},\vec{p}).$$

Here the current is stationary and uniform, and one can replace $\bar{\mathsf{f}}$ by $\bar{\mathsf{f}}^{(0)} + \bar{\mathsf{f}}^{(1)}$, with Eqs. (IV.51) and (IV.52). The integral of the term involving the local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$ vanishes by symmetry.(55) There thus only remains the contribution of $\bar{\mathsf{f}}^{(1)}$, which gives for the i-th component of $\vec{J}_{\mathrm{el.}}$

$$J^i_{\mathrm{el.}} = \frac{q}{m}\int_{\vec{p}} p^i\,\bar{\mathsf{f}}^{(1)}(\vec{p}) = \sum_{j=1}^{3}\frac{q^2}{m^2k_BT}\int_{\vec{p}}\bar{\mathsf{f}}^{(0)}(\vec{p})\,\tau_{\mathrm{r}}(\vec{p})\,p^ip^j\mathscr{E}^j.$$

---

(55) Had we allowed for a non-vanishing average velocity $\vec{\mathsf{v}}$, the $\bar{\mathsf{f}}^{(0)}$-term would have given a contribution $nq\vec{\mathsf{v}}$ to $\vec{J}_{\mathrm{el.}}$.

This is the i-th component of the relation

$$\vec{J}_{\rm el.} = \boldsymbol{\sigma}_{\rm el.} \cdot \vec{\mathscr{E}}, \tag{IV.53a}$$

where the electrical-conductivity tensor $\boldsymbol{\sigma}_{\rm el.}$ has the components

$$\sigma^{ij}_{\rm el.} \equiv \frac{q^2}{m^2 k_B T} \int_{\vec{p}} \bar{\mathsf{f}}^{(0)}(\vec{p})\, \tau_{\rm r}(\vec{p})\, p^i p^j. \tag{IV.53b}$$

If the relaxation time $\tau_{\rm r}(\vec{p})$ only depends on the modulus $|\vec{p}|$ of momentum, not on its orientation, then the product $\bar{\mathsf{f}}^{(0)}(\vec{p})\, \tau_{\rm r}(\vec{p})$ is an even function of the (Cartesian) components p^i. The integrand in Eq. (IV.53b) is then odd in p^i (or p^j) when $i \neq j$, so that the non-diagonal elements of the tensor vanish. In turn, all three diagonal elements are equal, i.e. the tensor is proportional to the identity. Assuming that the relaxation time is independent of momentum (BGK approximation), one obtains

$$\sigma^{ii}_{\rm el.} = \frac{q^2 \tau_{\rm r}}{m^2 k_B T} \int_{\vec{p}} (p^i)^2\, \bar{\mathsf{f}}^{(0)}(\vec{p}) = \frac{q^2 \tau_{\rm r}}{m^2 k_B T}\, n m k_B T = \frac{n q^2 \tau_{\rm r}}{m}.$$

All in all, one thus finds Ohm's law [cf. Eq. (I.41b)]

$$\vec{J}_{\rm el.} = \sigma_{\rm el.} \vec{\mathscr{E}} \quad \text{with} \quad \sigma_{\rm el.} = \frac{n q^2 \tau_{\rm r}}{m}. \tag{IV.54}$$

Taking into account the dependence of the relaxation time on momentum gives a similar expression, with a different—yet in principle computable—expression of the electrical conductivity $\sigma_{\rm el.}$, or more generally of the electrical-conductivity tensor $\boldsymbol{\sigma}_{\rm el.}$ if $\tau_{\rm r}(\vec{p})$ depends on the direction of $\vec{p}$.

Remarks:

∗ Our interest in the electrical conductivity, which as recalled in the first remark on page 13 is defined in the stationary regime, justifies a posteriori the restriction to time-independent solutions.

∗ In writing down the local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$ as uniform, we ignored the electrostatic potential which gives rise to the electric field $\vec{\mathscr{E}}$. The reason is again our interest in the electrical conductivity $\sigma_{\rm el.}$, which is defined as the coefficient between $\vec{\mathscr{E}}$ and the electric current $\vec{J}_{\rm el.}$ in the absence of a gradient in the number of charge carriers, i.e. for a uniform particle number density—the reader should compare with § I.2.3 c.

∗ The electrical conductivity (IV.54) computed within the relaxation-time approximation is proportional to the relaxation time $\tau_{\rm r}$. Since the latter is supposedly governed by collisions, it should be roughly proportional to the mean-free time between two successive scatterings, or equivalently to the mean-free path $\ell_{\rm mfp}$. In this picture, a larger $\ell_{\rm mfp}$ results in a more efficient transport of charge carriers, which seems rather coherent with the longer way traveled by the moving charges in each step.

IV.5.2 Diffusion coefficient and heat conductivity

As second application of the ideas of Sec. IV.4.3, consider a Lorentz gas—i.e. a gas of light particles colliding on much heavier ones—, submitted to time-independent gradients in temperature and in chemical potential (or density)—which are imposed through the heavy, motionless collision partners.

The evolution equation of the single-particle distribution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ is again approximated as Eq. (IV.49), here without external force. As local equilibrium distribution, we take

$$\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}) = n(\vec{r}) \left(\frac{2\pi\hbar^2}{m k_B T(\vec{r})} \right)^{3/2} \mathrm{e}^{-\vec{p}^{\,2}/2mk_BT(\vec{r})},$$

where the local density $n(\vec{r})$ and temperature $T(\vec{r})$ are those imposed by the heavy partners. Instead of this "canonical" form of the Maxwell–Boltzmann velocity distribution, it will be more convenient to replace it by the "grand-canonical" expression(56)

$$\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}) = \exp\biggl[-\frac{\vec{p}^{\,2}/2m-\mu(\vec{r})}{k_BT(\vec{r})}\biggr], \tag{IV.55}$$

with the local chemical potential $\mu(\vec{r})$ instead of the number density. This will enable us to make more easily contact with the formalism of non-equilibrium thermodynamics in the entropy representation of Chapter I.

Looking for stationary solutions $\bar{\mathsf{f}}(\vec{r},\vec{p}) = \bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p})+\bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p})$ with $|\bar{\mathsf{f}}^{(1)}| \ll \bar{\mathsf{f}}^{(0)}$, Eq. (IV.49) reads to leading order

$$\frac{\partial\bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}) = -\frac{\bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p})}{\tau_{\rm r}(\vec{r},\vec{p})},$$

i.e.

$$\bar{\mathsf{f}}^{(1)}(\vec{r},\vec{p}) = -\frac{\tau_{\rm r}(\vec{r},\vec{p})}{m}\,\vec{p}\cdot\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}). \tag{IV.56}$$

IV.5.2 a Fluxes

With the help of the leading and subleading contributions (IV.55), (IV.56) to the single-particle distribution, we can first easily compute the fluxes of particle number and energy.

$$\vec{\mathcal{J}}_N(t,\vec{r}) = \int_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\vec{v} = \frac{1}{m}\int_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\,\vec{p},$$
$$\vec{\mathcal{J}}_E(t,\vec{r}) = \int_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\frac{\vec{p}^{\,2}}{2m}\,\vec{v} = \frac{1}{m}\int_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\frac{\vec{p}^{\,2}}{2m}\,\vec{p}.$$

The local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$ does not contribute for parity reasons—the integrands are odd in $\vec{p}$ and integrated over all momentum space.

To handle the contribution of $\bar{\mathsf{f}}^{(1)}$, we first use the chain rule to rewrite the gradient of $\bar{\mathsf{f}}^{(0)}$ with respect to position as

$$\vec{\nabla}_{\vec{r}}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}) = -\frac{1}{k_B}\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p})\biggl\{\frac{\vec{p}^{\,2}}{2m}\vec{\nabla}_{\vec{r}}\biggl[\frac{1}{T(\vec{r})}\biggr] + \vec{\nabla}_{\vec{r}}\biggl[-\frac{\mu(\vec{r})}{T(\vec{r})}\biggr]\biggr\}.$$

As in Sec. IV.5.1, inserting $\bar{\mathsf{f}}^{(1)}$, given by Eq. (IV.56), computed with this gradient in the flux densities $\vec{\mathcal{J}}_N$, $\vec{\mathcal{J}}_E$ leads to linear relations between the latter and the gradients $\vec{\nabla}_{\vec{r}}(1/T)$, $\vec{\nabla}_{\vec{r}}(-\mu/T)$, where the coefficients are in the general case tensors of rank 2. If the relaxation time $\tau_{\rm r}$ does not depend on the orientation of $\vec{p}$, as we assume from now on, the tensors are all diagonal and proportional to the unit tensor. The integrand of the diagonal coefficients of any of these tensors involves $(p^i)^2$, which has to be averaged—with a weight involving $\bar{\mathsf{f}}^{(0)}$, τ_r and various powers of $\vec{p}^{\,2}$—over all directions. Such an average is simply $\frac{1}{3}$ of the corresponding average of $\vec{p}^{\,2}$, so that one eventually obtains

$$\vec{\mathcal{J}}_N(\vec{r}) = \frac{1}{3m^2k_B}\int_{\vec{p}}\tau_{\rm r}(\vec{r},\vec{p})\,\vec{p}^{\,2}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p})\biggl\{\frac{\vec{p}^{\,2}}{2m}\vec{\nabla}_{\vec{r}}\biggl[\frac{1}{T(\vec{r})}\biggr] + \vec{\nabla}_{\vec{r}}\biggl[-\frac{\mu(\vec{r})}{T(\vec{r})}\biggr]\biggr\}, \tag{IV.57a}$$
$$\vec{\mathcal{J}}_E(\vec{r}) = \frac{1}{6m^3k_B}\int_{\vec{p}}\tau_{\rm r}(\vec{r},\vec{p})\,\vec{p}^{\,4}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p})\biggl\{\frac{\vec{p}^{\,2}}{2m}\vec{\nabla}_{\vec{r}}\biggl[\frac{1}{T(\vec{r})}\biggr] + \vec{\nabla}_{\vec{r}}\biggl[-\frac{\mu(\vec{r})}{T(\vec{r})}\biggr]\biggr\}. \tag{IV.57b}$$

Note that since T and μ only depend on position, not on momentum, we may drop the $\vec{r}$ subscripts from the $\vec{\nabla}$ operators without ambiguity.

---

(56) A simple mnemonic is to remember that it corresponds to the phase-space occupancy, as given by the Boltzmann factor, for a macrostate at temperature T and chemical potential μ.

IV.5.2 b Kinetic and transport coefficients

In the language of Chapter I, the relations (IV.57) express the fluxes $\vec{\mathcal{J}}_N$, $\vec{\mathcal{J}}_E$ as functions of the affinities $\vec{\nabla}(1/T)$ and $\vec{\nabla}(-\mu/T)$. Using the notation of Eq. (I.31), we may write these relations in the form

$$\vec{\mathcal{J}}_N = L_{NN}\vec{\nabla}\bigg(-\frac{\mu}{T}\bigg) + L_{NE}\vec{\nabla}\bigg(\frac{1}{T}\bigg),$$
$$\vec{\mathcal{J}}_E = L_{EN}\vec{\nabla}\bigg(-\frac{\mu}{T}\bigg) + L_{EE}\vec{\nabla}\bigg(\frac{1}{T}\bigg),$$

with position-dependent kinetic coefficients L_{NN}, L_{NE}, L_{EN} and L_{EE} that can be directly read off Eqs. (IV.57):

$$\begin{cases} L_{NN}(\vec{r}) = \dfrac{1}{3m^2k_B}\displaystyle\int_{\vec{p}} \tau_{\rm r}(\vec{r},\vec{p})\,\vec{p}^{\,2}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}),\\ L_{EN}(\vec{r}) = L_{NE}(\vec{r}) = \dfrac{1}{6m^3k_B}\displaystyle\int_{\vec{p}} \tau_{\rm r}(\vec{r},\vec{p})\,\vec{p}^{\,4}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}),\\ L_{EE}(\vec{r}) = \dfrac{1}{12m^4k_B}\displaystyle\int_{\vec{p}} \tau_{\rm r}(\vec{r},\vec{p})\,\vec{p}^{\,6}\,\bar{\mathsf{f}}^{(0)}(\vec{r},\vec{p}). \end{cases} \tag{IV.58}$$

In particular, we see that the Onsager reciprocal relation $L_{EN} = L_{NE}$ is fulfilled.

If we assume that the relaxation time is independent of position and momentum, $\tau_{\rm r}(\vec{r},\vec{p}) = \tau_{\rm r}$, the integrals can be performed with the help of formula (A.1c) with $a = 1/2mk_BT(\vec{r})$ and $2n = 4$, 6 and 8, and lead to

$$\begin{cases} L_{NN}(\vec{r}) = \dfrac{n(\vec{r})\tau_{\rm r}T(\vec{r})}{m},\\ L_{EN}(\vec{r}) = L_{NE}(\vec{r}) = \dfrac{5}{2}\dfrac{n(\vec{r})\tau_{\rm r}T(\vec{r})}{m}k_BT(\vec{r}),\\ L_{EE}(\vec{r}) = \dfrac{35}{4}\dfrac{n(\vec{r})\tau_{\rm r}T(\vec{r})}{m}\big[k_BT(\vec{r})\big]^2. \end{cases}$$

From these kinetic coefficients, one deduces expressions for various transport coefficients using relations which were derived in Sec. I.2.3.

Thus, the diffusion coefficient of the particles of the Lorentz gas is related through Eq. (I.39c) to the kinetic coefficient L_{NN}

$$D = \frac{1}{T}\bigg(\frac{\partial\mu}{\partial n}\bigg)_T L_{NN} = \frac{\tau_{\rm r}k_BT}{m},$$

where the second identity holds when the relaxation time is independent of position and momentum. In turn, inserting the above expressions of the kinetic coefficients in the heat conductivity (I.55) yields

$$\kappa = \frac{L_{NN}L_{EE} - L_{NE}L_{EN}}{T^2\,L_{NN}} = \frac{5}{2}\frac{nk_B\tau_{\rm r}}{m}k_BT.$$

For both transport coefficients, we now have expressions in terms of microscopic quantities.

IV.6 From Boltzmann to hydrodynamics

In Sec. IV.3.1, we have seen that if $\chi(t,\vec{r},\vec{p})$ is an additive collisional invariant, then the integral over momentum space of its product with the collision term on the right-hand side of the Boltzmann equation is vanishing, Eq. (IV.19). This property can then be used to derive local conservation laws.

Here, we wish to exploit this idea again, rewriting the conservation laws in an alternative way (Sec. IV.6.1). We then apply the new forms of the balance equations to solutions $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$ to the Boltzmann equation written as in Eq. (IV.50) as the sum of a local equilibrium distribution and a (hopefully small) deviation. This leads to relations between the fields (number density, mean velocity, temperature) entering the expression of the local equilibrium solution, which are exactly the laws governing the (thermo)dynamics of a perfect (Sec. IV.6.2) or a Newtonian (Sec. IV.6.3) fluid. In the latter case, we can compute the various transport coefficients in terms of microscopic quantities.

IV.6.1 Conservation laws revisited

If $\chi(t,\vec{r},\vec{p})$ denotes a quantity, carried by the colliding particles, which is locally conserved in elastic collisions, then

$$\int_{\vec{p}} \chi(t,\vec{r},\vec{p}) \left(\frac{\partial\bar{\mathsf{f}}}{\partial t}\right)_{\rm coll.}\!(t,\vec{r},\vec{p}) = 0,$$

where $(\partial\bar{\mathsf{f}}/\partial t)_{\rm coll.}$ denotes the collision integral of the Boltzmann equation, computed with *any* single-particle distribution $\bar{\mathsf{f}}(t,\vec{r},\vec{p})$—that is, $\bar{\mathsf{f}}$ need not be a solution to the Boltzmann equation.

Now, if $\bar{\mathsf{f}}$ is a solution to the Boltzmann equation, then the collision term equals the left-hand side of Eq. (IV.8), resulting in the identity

$$\int_{\vec{p}} \chi(t,\vec{r},\vec{p}) \left[\frac{\partial\bar{\mathsf{f}}(t,\vec{r},\vec{p})}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}}\bar{\mathsf{f}}(t,\vec{r},\vec{p}) + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\bar{\mathsf{f}}(t,\vec{r},\vec{p})\right] = 0,$$

Dropping the dependence of the functions on their variables and performing some straightforward algebra, this can be rewritten as

$$\int_{\vec{p}} \frac{\partial(\chi\bar{\mathsf{f}})}{\partial t} - \int_{\vec{p}} \bar{\mathsf{f}}\,\frac{\partial\chi}{\partial t} + \int_{\vec{p}} \vec{\nabla}_{\vec{r}}\cdot\big(\chi\bar{\mathsf{f}}\vec{v}\big) - \int_{\vec{p}} \bar{\mathsf{f}}\,\vec{\nabla}_{\vec{r}}\cdot\big(\chi\vec{v}\big) + \int_{\vec{p}} \vec{\nabla}_{\vec{p}}\cdot\big(\chi\bar{\mathsf{f}}\vec{F}\big) - \int_{\vec{p}} \bar{\mathsf{f}}\,\vec{\nabla}_{\vec{p}}\cdot\big(\chi\vec{F}\big) = 0.$$

In the fourth term, the velocity is independent of the position, and can thus be taken out of the gradient, $\vec{\nabla}_{\vec{r}}\cdot\big(\chi\vec{v}\big) = \vec{v}\cdot\vec{\nabla}_{\vec{r}}\chi$. Similarly, if we from now on restrict ourselves to momentum-independent forces, then in the sixth term one may write $\vec{\nabla}_{\vec{p}}\cdot\big(\chi\vec{F}\big) = \vec{F}\cdot\vec{\nabla}_{\vec{p}}\chi$. Eventually, the fifth term, which is trivially integrated, is zero, since the single-particle density vanishes when $|\vec{p}|\to\infty$.

Exchanging the order of the integration over $\vec{p}$ with the derivation with respect to t or $\vec{r}$ in the first and third terms, and introducing the notation

$$\langle\,\cdot\,\rangle_{\vec{p}} \equiv \frac{1}{n(t,\vec{r})}\int_{\vec{p}} (\,\cdot\,)\,\bar{\mathsf{f}}(t,\vec{r},\vec{p}),$$

which clearly represents an $\bar{\mathsf{f}}$-weighted average over momenta, the above identity becomes

$$\frac{\partial(n\langle\chi\rangle_{\vec{p}})}{\partial t} - n\left\langle\frac{\partial\chi}{\partial t}\right\rangle_{\vec{p}} + \vec{\nabla}_{\vec{r}}\cdot\big\langle n\chi\vec{v}\,\big\rangle_{\vec{p}} - n\big\langle\vec{v}\cdot\vec{\nabla}_{\vec{r}}\chi\big\rangle_{\vec{p}} - n\vec{F}\cdot\big\langle\vec{\nabla}_{\vec{p}}\chi\big\rangle_{\vec{p}} = 0. \qquad \text{(IV.59)}$$

Note that since the number density n is independent of velocity, it can be moved inside or outside of averages over $\vec{p}$ at will.

We can now re-express the balance equations (IV.20)–(IV.22) in an equivalent, yet slightly different manner.

IV.6.1 a Mass conservation

First, instead of particle number, we consider mass, setting $\chi = m$ in formula (IV.59). Since this is a constant, the second, fourth and fifth term drop out, while the average $\langle m\rangle_{\vec{p}}$ in the first term is trivial. The remaining first and third terms then give

$$\frac{\partial}{\partial t}\big[mn(t,\vec{r})\big] + \vec{\nabla}_{\vec{r}}\cdot\big[mn(t,\vec{r})\langle\vec{v}\rangle_{\vec{p}}\big] = 0.$$

Introducing the mass density

$$\rho(t,\vec{r}) = mn(t,\vec{r}) \qquad \text{(IV.60a)}$$

and the average velocity [cf. Eq. (IV.22a)]

$$\vec{\mathrm{v}}(t,\vec{r}) \equiv \langle\vec{v}\rangle_{\vec{p}}, \tag{IV.60b}$$

this becomes

$$\boxed{\frac{\partial\rho(t,\vec{r})}{\partial t} + \vec{\nabla}\cdot\big[\rho(t,\vec{r})\,\vec{\mathrm{v}}(t,\vec{r})\big] = 0,} \tag{IV.61}$$

which expresses the local conservation of mass.(57)

IV.6.1 b Momentum balance

Choosing now for χ the i-th component of linear momentum $p^i = mv^i$, the second and fourth terms in the balance equation (IV.59) vanish, leaving

$$\frac{\partial}{\partial t}\big[mn\langle v^i\rangle_{\vec{p}}\big] + \vec{\nabla}_{\vec{r}}\cdot\big[mn\langle v^i\vec{v}\rangle_{\vec{p}}\big] - nF^i = 0. \tag{IV.62}$$

The term within square brackets in the time derivative is clearly $\rho(t,\vec{r})\,\mathrm{v}^i(t,\vec{r})$. The mass density also appears in the argument of the spatial-divergence term. The average $\langle v^i\vec{v}\rangle_{\vec{v}}$ can be transformed by noting the identity

$$\langle(v^i-\mathrm{v}^i)(v^j-\mathrm{v}^j)\rangle_{\vec{p}} = \langle v^iv^j\rangle_{\vec{p}} - \mathrm{v}^i\mathrm{v}^j$$

and introducing the second-rank *stress tensor* $\boldsymbol{\pi}$, whose components are given by

$$\pi^{ij}(t,\vec{r}) = \rho(t,\vec{r})\Big\langle\big[v^i-\mathrm{v}^i(t,\vec{r})\big]\big[v^j-\mathrm{v}^j(t,\vec{r})\big]\Big\rangle_{\vec{p}}. \tag{IV.63}$$

The balance equation (IV.62) then becomes

$$\boxed{\frac{\partial}{\partial t}\big[\rho(t,\vec{r})\mathrm{v}^i(t,\vec{r})\big] + \sum_{j=1}^{3}\frac{\partial}{\partial x^j}\big[\rho(t,\vec{r})\mathrm{v}^i(t,\vec{r})\mathrm{v}^j(t,\vec{r}) + \pi^{ij}(t,\vec{r})\big] = \frac{1}{m}\,\rho(t,\vec{r})F^i(t,\vec{r}).} \tag{IV.64a}$$

Using the mass conservation equation (IV.61), this can be rewritten as(57)

$$\boxed{\frac{\partial\mathrm{v}^i(t,\vec{r})}{\partial t} + \vec{\mathrm{v}}(t,\vec{r})\cdot\vec{\nabla}\mathrm{v}^i(t,\vec{r}) = \frac{1}{m}\,F^i(t,\vec{r}) - \frac{1}{\rho(t,\vec{r})}\sum_{j=1}^{3}\frac{\partial\pi^{ij}(t,\vec{r})}{\partial x^j}.} \tag{IV.64b}$$

IV.6.1 c Balance equation for internal energy

Consider eventually $\chi(t,\vec{r},\vec{p}) = \big[\vec{p}-m\vec{\mathrm{v}}(t,\vec{r})\big]^2/2m = \frac{1}{2}m\big[\vec{v}-\vec{\mathrm{v}}(t,\vec{r})\big]^2$, which represents the kinetic energy of the particles in a frame locally comoving with their average velocity. Differentiation with respect to time or momentum followed by an average over $\vec{p}$ easily allow one to check that the second and fifth terms in Eq. (IV.59) are zero. There remains

$$\frac{1}{2}\frac{\partial}{\partial t}\big\langle mn\big(\vec{v}-\vec{\mathrm{v}}\big)^2\big\rangle_{\vec{p}} + \frac{1}{2}\vec{\nabla}_{\vec{r}}\cdot\big\langle mn\big(\vec{v}-\vec{\mathrm{v}}\big)^2\vec{v}\,\big\rangle_{\vec{p}} - \frac{1}{2}mn\big\langle\vec{v}\cdot\vec{\nabla}_{\vec{r}}\big(\vec{v}-\vec{\mathrm{v}}\big)^2\big\rangle_{\vec{p}} = 0. \tag{IV.65}$$

In the third term, one can first write $\frac{1}{2}\vec{\nabla}_{\vec{r}}\big(\vec{v}-\vec{\mathrm{v}}\big)^2 = \big[\big(\vec{v}-\vec{\mathrm{v}}\big)\cdot\vec{\nabla}_{\vec{r}}\big]\vec{\mathrm{v}} - \big(\vec{v}-\vec{\mathrm{v}}\big)\times\big(\vec{\nabla}_{\vec{r}}\times\vec{\mathrm{v}}\big)$. One then checks that the third term equals the negative of the sum over all indices i and j of π^{ij} times the derivative $\partial\mathrm{v}^j/\partial x^i$.

Let then

$$e(t,\vec{r}) = \frac{1}{2}\rho(t,\vec{r})\Big\langle\big[\vec{v}-\vec{\mathrm{v}}(t,\vec{r})\big]^2\Big\rangle_{\vec{p}} \tag{IV.66a}$$

---

(57) The subscript on the $\vec{\nabla}$ operator has been suppressed, as it is obvious that it denotes the gradient with respect to position.

denote the local density of internal energy and

$$\vec{\mathcal{J}}_U(t,\vec{r}) = \frac{1}{2}\rho(t,\vec{r})\Big\langle\big[\vec{v}-\vec{\mathsf{v}}(t,\vec{r})\big]^2\big[\vec{v}-\vec{\mathsf{v}}(t,\vec{r})\big]\Big\rangle_{\vec{p}} \tag{IV.66b}$$

be the internal energy (or heat) flux in the local rest frame. These quantities allow one to rewrite the first and second term in Eq. (IV.65) as the time derivative of internal-energy density and the divergence of $\vec{\mathcal{J}}_U + e\vec{\mathsf{v}}$, respectively.

All in all, one finds(57)

$$\boxed{\frac{\partial e(t,\vec{r})}{\partial t} + \vec{\nabla}\cdot\Big[\vec{\mathcal{J}}_U(t,\vec{r}) + e(t,\vec{r})\,\vec{\mathsf{v}}(t,\vec{r})\Big] = -\sum_{i,j=1}^{3}\pi^{ij}(t,\vec{r})\frac{\partial \mathsf{v}_j(t,\vec{r})}{\partial x^i}.} \tag{IV.67}$$

This expresses the local balance of energy.

The balance equations (IV.61), (IV.64b) and (IV.67) are the same, up to the external force, as the laws of hydrodynamics (I.68), (I.69) and (I.70), which were derived at the macroscopic level. Here, we have expressions of the stress tensor and the internal-energy flux in terms of microscopic quantities, which will allow us to compute them provided we have the form of the single-particle density in the Boltzmann gas.

IV.6.2 Zeroth-order approximation: the Boltzmann gas as a perfect fluid

We first start by assuming that the single-particle distribution in the Boltzmann gas is given by a local equilibrium distribution $\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})$ of the type (IV.38), characterized by a local number density $n(t,\vec{r})$, a local temperature $T(t,\vec{r})$ and a local average velocity $\vec{\mathsf{v}}(t,\vec{r})$.

IV.6.2 a Stress tensor and internal-energy flux density

Inserting $\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})$ in the expressions (IV.63) and (IV.66b), we obtain the corresponding stress tensor $\pi_{ij}^{(0)}$ and flux density of internal energy in the local rest frame $\vec{\mathcal{J}}_U^{(0)}$.

Let $\vec{w} \equiv \vec{v} - \vec{\mathsf{v}}(t,\vec{r})$ denote the velocity of a particle as measured with respect to that reference frame. The stress tensor reads component-wise(58)

$$\pi_{ij}^{(0)}(t,\vec{r}) = \rho(t,\vec{r})\left[\frac{m}{2\pi k_BT(t,\vec{r})}\right]^{3/2}\int w_iw_j\exp\left\{-\frac{m\vec{w}^{\,2}}{2k_BT(t,\vec{r})}\right\}\mathrm{d}^3\vec{w}.$$

When $i \neq j$, the integrand is odd in w_i, and thus the integral vanishes: the off-diagonal elements of the stress tensor are zero. Using formula (A.1b) with $n = 2$, one finds that all three diagonal elements are equal to $\mathcal{P}(t,\vec{r}) \equiv n(t,\vec{r})k_BT(t,\vec{r})$. All in all

$$\pi_{ij}^{(0)} = \mathcal{P}(t,\vec{r})\delta_{ij} = n(t,\vec{r})k_BT(t,\vec{r})\delta_{ij}. \tag{IV.68}$$

Interpreting $\mathcal{P}(t,\vec{r})$ as the local pressure—which can be justified by investigating the force on a surface element, which is then normal—, one recognizes the mechanical equation of state of a classical ideal gas.

Calculating the internal energy density (IV.66a), one at once finds $e = \frac{3}{2}nk_BT = \frac{3}{2}\mathcal{P}$, i.e. the thermal equation of state of a classical ideal gas.

For the internal-energy flux, Eqs. (IV.38) and (IV.66b) give

$$\vec{\mathcal{J}}_U^{(0)}(t,\vec{r}) = \frac{1}{2}\rho(t,\vec{r})\left[\frac{m}{2\pi k_BT(t,\vec{r})}\right]^{3/2}\int \vec{w}^{\,2}\vec{w}\exp\left\{-\frac{m\vec{w}^{\,2}}{2k_BT(t,\vec{r})}\right\}\mathrm{d}^3\vec{w} = \vec{0} \tag{IV.69}$$

---

(58) Throughout this section we use the straightforward change of variable $\displaystyle\int_{\vec{p}}(\,\cdot\,) = \int(\,\cdot\,)\frac{\mathrm{d}^3\vec{p}}{(2\pi\hbar)^3} = \frac{m^3}{(2\pi\hbar)^3}\int(\,\cdot\,)\,\mathrm{d}^3\vec{w}$.

where the rightmost identity is due to the oddness of the integrand. That is, there is no diffusive transport of internal energy to this order of approximation.

Since $\boldsymbol{\pi}^{(0)}$ is diagonal, there is no transport of linear momentum by (shear) viscosity either: the Boltzmann gas behaves like a non-dissipative fluid. As a consequence, the balance equations (IV.61), (IV.64b) and (IV.67) will represent the laws governing the dynamics of such a perfect fluid.

IV.6.2 b Dynamics of a perfect fluid

Inserting the stress tensor (IV.68) and the (trivial!) heat flux density (IV.69) in Eqs. (IV.64b) and (IV.67), one finds

$$\boxed{\frac{\partial\vec{\mathsf{v}}(t,\vec{r})}{\partial t}+\Big[\vec{\mathsf{v}}(t,\vec{r})\cdot\vec{\nabla}\Big]\vec{\mathsf{v}}(t,\vec{r})=\frac{1}{m}\,\vec{F}(t,\vec{r})-\frac{1}{\rho(t,\vec{r})}\,\vec{\nabla}\mathcal{P}(t,\vec{r}),}\tag{IV.70a}$$

which is the *Euler equation* for the dynamics of a perfect fluid, while the local balance of internal energy reads

$$\boxed{\frac{\partial e(t,\vec{r})}{\partial t}+\vec{\nabla}\cdot\Big[e(t,\vec{r})\,\vec{\mathsf{v}}(t,\vec{r})\Big]=-\mathcal{P}(t,\vec{r})\,\vec{\nabla}\cdot\vec{\mathsf{v}}(t,\vec{r}),}\tag{IV.70b}$$

where in the right-side one can make the substitution $\mathcal{P}(t,\vec{r})=\frac{2}{3}e(t,\vec{r})$.

IV.6.3 First-order approximation: the Boltzmann gas as a Newtonian fluid

Adopting now the relaxation-time approximation (IV.49) and assuming that the single-particle distribution can be written as $\bar{\mathsf{f}}=\bar{\mathsf{f}}^{(0)}+\bar{\mathsf{f}}^{(1)}$ [Eq. (IV.50)], we can find $\bar{\mathsf{f}}^{(1)}$, and deduce the modifications of the stress tensor and flux of internal energy, which give us new dynamical equations.

IV.6.3 a Single-particle distribution to subleading order

In the framework of relaxation-time approximation with a momentum-independent relaxation time, and making the decomposition $\bar{\mathsf{f}}=\bar{\mathsf{f}}^{(0)}+\bar{\mathsf{f}}^{(1)}$ with $\big|\bar{\mathsf{f}}^{(1)}\big|\ll\bar{\mathsf{f}}^{(0)}$, one finds after dropping out the terms of higher order the relation

$$\bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})=-\tau_{\rm r}\bigg(\frac{\partial}{\partial t}+\vec{v}\cdot\vec{\nabla}_{\vec{r}}+\vec{F}\cdot\vec{\nabla}_{\vec{p}}\bigg)\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p}),\tag{IV.71}$$

which allows the calculation of $\bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})$ by differentiating the expression (IV.38) of the local equilibrium distribution $\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p})$.

Assuming for simplicity that there is not external force $\vec{F}$ and focusing on stationary solutions, one finds after performing all computations(59)

$$\bar{\mathsf{f}}^{(1)}=-\tau_{\rm r}\Bigg[\frac{\vec{w}\cdot\vec{\nabla}T}{T}\bigg(\frac{1}{k_BT}\frac{m}{2}\vec{w}^{\,2}-\frac{5}{2}\bigg)+\frac{m}{k_BT}\sum_{i,j}w^iw^j\frac{\partial\mathsf{v}_i}{\partial x^j}-\frac{1}{3}\frac{m\vec{w}^{\,2}}{k_BT}\,\vec{\nabla}\cdot\vec{\mathsf{v}}\Bigg]\bar{\mathsf{f}}^{(0)},$$

where as above $\vec{w}\equiv\vec{v}-\vec{\mathsf{v}}(t,\vec{r})$, while all gradients are with respect to spatial coordinates. This can also be rewritten as

$$\bar{\mathsf{f}}^{(1)}=-\tau_{\rm r}\Bigg[\frac{\vec{w}\cdot\vec{\nabla}T}{T}\bigg(\frac{1}{k_BT}\frac{m}{2}\vec{w}^{\,2}-\frac{5}{2}\bigg)+\frac{m}{2k_BT}\sum_{i,j=1}^{3}\bigg(\frac{\partial\mathsf{v}_i}{\partial x^j}+\frac{\partial\mathsf{v}_j}{\partial x^i}\bigg)\bigg(w^iw^j-\frac{1}{3}\vec{w}^{\,2}\delta^{ij}\bigg)\Bigg]\bar{\mathsf{f}}^{(0)}.\tag{IV.72}$$

IV.6.3 b Stress tensor and internal-energy flux density

Substituting $\bar{\mathsf{f}}=\bar{\mathsf{f}}^{(0)}+\bar{\mathsf{f}}^{(1)}$ in Eq. (IV.63) yields the expression of the stress tensor to first order in $\tau_{\rm r}$. With the same local equilibrium distribution $\bar{\mathsf{f}}^{(0)}$ as in Sec. IV.6.2 and $\bar{\mathsf{f}}^{(1)}$ given by Eq. (IV.72),

---

(59) The detailed derivation—which only necessitates some careful bookkeeping in computing the various partial derivatives, but presents no real difficulty, and does not provide much physical insight—can be found e.g. in Huang [33], chapter 5.5.

one finds

$$\pi_{ij}(t,\vec{r}) = m\int_{\vec{p}} w_i w_j \big[\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p}) + \bar{\mathsf{f}}^{(1)}(t,\vec{r},\vec{p})\big] = \pi_{ij}^{(0)}(t,\vec{r}) + \pi_{ij}^{(1)}(t,\vec{r}),$$

with obvious notations, where $\pi_{ij}^{(0)}$ is again given by Eq. (IV.68), $\pi_{ij}^{(0)} = n k_B T\,\delta_{ij}$.

Inspecting the correction to the single-particle distribution (IV.72), the first term within the square brackets is odd in $\vec{w}$ and thus will not contribute to $\pi_{ij}^{(1)}$, which leaves

$$\pi_{ij}^{(1)}(t,\vec{r}) = -\frac{m^2\tau_{\rm r}}{2k_BT}\sum_{k,l=1}^{3}\bigg(\frac{\partial \mathsf{v}_k}{\partial x^l}+\frac{\partial \mathsf{v}_l}{\partial x^k}\bigg)\int_{\vec{p}} w_i w_j \bigg(w^k w^l - \frac{1}{3}\vec{w}^2\delta^{kl}\bigg)\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p}).$$

In the integral, only the indices k, l such that every component of $\vec{w}$ appears with an even power contribute: either $k=i$ and $j=l$, or $k=j$ and $l=i$, or $i=j$ (diagonal terms) and $k=l$. Using then formula (A.1b) with $n=2$ or 3, one finds

$$\pi_{ij}^{(1)}(t,\vec{r}) = -\eta(t,\vec{r})\bigg[\frac{\partial \mathsf{v}_i(t,\vec{r})}{\partial x^j}+\frac{\partial \mathsf{v}_j(t,\vec{r})}{\partial x^i}-\frac{2}{3}\delta_{ij}\vec{\nabla}\cdot\vec{\mathsf{v}}(t,\vec{r})\bigg] \tag{IV.73a}$$

with

$$\eta(t,\vec{r}) = n(t,\vec{r})k_BT(t,\vec{r})\tau_{\rm r}, \tag{IV.73b}$$

and all in all

$$\pi_{ij} = \mathcal{P}\,\delta_{ij} - \eta\bigg[\frac{\partial \mathsf{v}_i}{\partial x^j}+\frac{\partial \mathsf{v}_j}{\partial x^i}-\frac{2}{3}(\vec{\nabla}\cdot\vec{\mathsf{v}})\,\delta_{ij}\bigg]. \tag{IV.73c}$$

Comparing with Eqs. (I.71b) and (I.71d), one recognizes the stress tensor of a Newtonian fluid, whose shear viscosity η is given by Eq. (IV.73b), while its bulk viscosity ζ vanishes.

Inserting now $\bar{\mathsf{f}} = \bar{\mathsf{f}}^{(0)} + \bar{\mathsf{f}}^{(1)}$ in the internal energy flux (IV.66b), $\bar{\mathsf{f}}^{(0)}$ will as in Eq. (IV.69) not contribute, neither will the second term within the square brackets in Eq. (IV.72). There remains

$$\begin{aligned}\vec{\mathcal{J}}_U(t,\vec{r}) &= \vec{\mathcal{J}}_U^{(0)}(t,\vec{r}) + \vec{\mathcal{J}}_U^{(1)}(t,\vec{r})\\ &= -\frac{m^2\tau_{\rm r}}{2T(t,\vec{r})}\sum_{j=1}^{3}\frac{\partial T(t,\vec{r})}{\partial x_j}\int_{\vec{p}}\bigg(\frac{1}{k_BT}\frac{m}{2}\vec{w}^2-\frac{5}{2}\bigg)w_j\vec{w}^2\vec{w}\,\bar{\mathsf{f}}^{(0)}(t,\vec{r},\vec{p}).\end{aligned}$$

Performing the integration, one eventually finds

$$\vec{\mathcal{J}}_U(t,\vec{r}) = \vec{\mathcal{J}}_U^{(1)}(t,\vec{r}) = -\kappa(t,\vec{r})\vec{\nabla}T(t,\vec{r}) \tag{IV.74a}$$

with

$$\kappa(t,\vec{r}) = \frac{5}{2}\frac{n(t,\vec{r})k_B^2T(t,\vec{r})\tau_{\rm r}}{m}. \tag{IV.74b}$$

One recognizes Fourier's law, with κ the heat conductivity.

Remark: The transport coefficients η and κ, Eqs. (IV.73b) and (IV.74b) are both proportional to the relaxation time $\tau_{\rm r}$. As mentioned in Sec. IV.4.3, this time is (at least) of the order of the mean free time $\tau_{\rm mfp}$ between two successive collisions of a particle, say $\tau_{\rm r}\sim\tau_{\rm mfp}$. The latter, divided by some typical particle velocity, gives the mean free path $\ell_{\rm mfp}$, i.e. the typical length traveled by a particle between two successive collisions. In turn, $\ell_{\rm mfp}$ is inversely proportional to the particle density and to the total interaction cross-section, $\ell_{\rm mfp}\sim 1/n\sigma_{\rm tot.}$. As a consequence, η and κ in a Boltzmann gas are in first approximation independent of density—yet the latter should be small enough that only two-to-two collisions take place—and inversely proportional to the cross-section: the more ideal a gas is (small $\sigma_{\rm tot}$), the more dissipative (large transport coefficients) it is. An ideal gas is thus not a perfect fluid!

IV.6.3 c Dynamics of a Newtonian fluid

Eventually, one can substitute the stress tensor (IV.73c) and the internal-energy flux (IV.74a) in the balance equation for linear momentum [Eq. (IV.64b)]. Straightforward calculations give

$$\boxed{\frac{\partial\vec{\mathrm{v}}}{\partial t}+\big(\vec{\mathrm{v}}\cdot\vec{\nabla}\big)\vec{\mathrm{v}}=\frac{1}{m}\vec{F}-\frac{1}{\rho}\vec{\nabla}\mathcal{P}+\frac{\eta}{\rho}\bigg[\triangle\vec{\mathrm{v}}+\frac{1}{3}\vec{\nabla}\big(\vec{\nabla}\cdot\vec{\mathrm{v}}\big)\bigg],}\tag{IV.75}$$

which is the *Navier–Stokes equation* [Eq. (I.77)] for a Newtonian fluid with vanishing bulk viscosity. Inserting the stress tensor and heat flux in the balance equation for internal energy, one accordingly recovers Eq. (I.78) with $\zeta = 0$.

Bibliography

- Huang, *Statistical Mechanics* [33], chapter 3–5.
- Landau & Lifshitz, *Course of theoretical physics. Vol. 10: Physical kinetics* [5], chapter I.
- Le Bellac, Mortessagne & Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* [9], chapter 8.
- Pottier, *Nonequilibrium statistical physics* [6], chapters 5–7.
- Reif [35], chapter 14.
- Zwanzig, *Nonequilibrium statistical mechanics* [7], chapter 5.

Appendix to Chapter IV

IV.A Derivation of the Boltzmann equation from the BBGKY hierarchy

In Sec. IV.2, we obtained the right hand side of the kinetic Boltzmann equation, the collision term, by writing it down as a sum—rather, a difference—of two contributions describing the influence of particle scatterings using heuristic arguments. Accordingly, the relationship to the first equation (III.14a) of the Bogolioubov–Born–Green–Kirkwood–Yvon hierarchy derived in Chapter III is far from being transparent.

In this Appendix, we go back to the hierarchy and take as starting point its first two equations (III.14a)–(III.14b). Applying to them the physical assumptions on the system introduced in Sec. IV.1—together with microreversibility and molecular chaos—, we show that they indeed lead to the Boltzmann equation in its form (IV.15b).

IV.A.1 BBGKY hierarchy in a weakly interacting system

Let us start by recalling the first two equations of the BBGKY hierarchy deduced from the Liouville evolution equation, namely

$$\left(\frac{\partial}{\partial t}+\vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1}+\vec{F}_1\cdot\vec{\nabla}_{\vec{p}_1}\right)f_1(t,\vec{r}_1,\vec{p}_1)=-\int\vec{K}_{12}\cdot\vec{\nabla}_{\vec{p}_1}f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2)\,\mathrm{d}^3\vec{r}_2\,\mathrm{d}^3\vec{p}_2 \tag{IV.76a}$$

and

$$\left[\frac{\partial}{\partial t}+\vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1}+\vec{v}_2\cdot\vec{\nabla}_{\vec{r}_2}+\vec{F}_1\cdot\vec{\nabla}_{\vec{p}_1}+\vec{F}_2\cdot\vec{\nabla}_{\vec{p}_2}+\vec{K}_{12}\cdot\big(\vec{\nabla}_{\vec{p}_1}-\vec{\nabla}_{\vec{p}_2}\big)\right]f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2)=\\ -\int\big(\vec{K}_{13}\cdot\vec{\nabla}_{\vec{p}_1}+\vec{K}_{23}\cdot\vec{\nabla}_{\vec{p}_2}\big)f_3(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2,\vec{r}_3,\vec{p}_3)\,\mathrm{d}^3\vec{r}_3\,\mathrm{d}^3\vec{p}_3. \tag{IV.76b}$$

In those equations, $\vec{F}_i$ and $\vec{F}_{ij}$ are the forces acting on particle i respectively due to the external potential and the action of particle $j\neq i$, see Eqs. (III.10). In § III.2.2 b–III.2.2 c, we already introduced the three time scales involved in Eq. (IV.76a) and on the left hand side of Eq. (IV.76b), namely

$$\tau_{\mathrm{s}}\sim\big(\vec{v}\cdot\vec{\nabla}_{\vec{r}}\big)^{-1},\qquad\tau_{\mathrm{e}}\sim\big(\vec{F}\cdot\vec{\nabla}_{\vec{p}}\big)^{-1},\qquad\tau_{\mathrm{c}}\sim\big(\vec{K}\cdot\vec{\nabla}_{\vec{p}}\big)^{-1}, \tag{IV.77}$$

where the first two ones are typically much larger than the third, which may be seen as the typical time over which two particles interact with each other—i.e. remain within a distance from each other of the order of the interaction range r_0.

An extra time scale follows from the right member of Eq. (IV.76b), when compared to the left member of the equation, namely

$$\tau_{\text{r.h.s.}}\sim\left[\frac{1}{f_2}\int\vec{K}\cdot\vec{\nabla}_{\vec{p}}f_3\,\mathrm{d}^3\vec{r}_3\,\mathrm{d}^3\vec{p}_3\right]^{-1}. \tag{IV.78}$$

Using Eq. (IV.77), the product $\vec{K}\cdot\vec{\nabla}_{\vec{p}}$ is of order τ_{c}^{-1}. Since $\vec{K}$ depends on the distance between particle 3 and the other particle it is interacting with, the spatial volume over which the integrand takes significant values is of order r_0^3. Eventually, going back to the definitions of the reduced phase-space densities shows that integrating f_3 over the whole momentum and spatial volumes available for the third particle yields

$$\int f_3\,\mathrm{d}^3\vec{r}_3\,\mathrm{d}^3\vec{p}_3\sim Nf_2,$$

so that the integral of f_3 over $\vec{p}_3$ only is of order

$$\int f_3 \, \mathrm{d}^3\vec{p}_3 \sim n f_2,$$

with n the particle number density. All in all, one obtains

$$\tau_{\text{r.h.s.}} \sim \frac{\tau_{\text{c}}}{n r_0^3}. \tag{IV.79}$$

The case of a weakly interacting system considered in this Chapter is that in which the diluteness condition $n r_0^3 \ll 1$ holds, resulting in $\tau_{\text{r.h.s.}} \gg \tau_{\text{c}}$. That is, the right member of Eq. (IV.76b) is necessarily much smaller than the terms on the left hand side, which justifies setting it equal to zero

$$\left[\frac{\partial}{\partial t} + \vec{v}_1 \cdot \vec{\nabla}_{\vec{r}_1} + \vec{v}_2 \cdot \vec{\nabla}_{\vec{r}_2} + \vec{F}_1 \cdot \vec{\nabla}_{\vec{p}_1} + \vec{F}_2 \cdot \vec{\nabla}_{\vec{p}_2} + \vec{K}_{12} \cdot \big(\vec{\nabla}_{\vec{p}_1} - \vec{\nabla}_{\vec{p}_2}\big)\right] f_2(t, \vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = 0, \tag{IV.80}$$

thereby effectively truncating (and closing) the BBGKY hierarchy.

Using the same order-of-magnitude approximations and arguments, one checks that the term on the right hand side of the k-th equation of the BBGKY hierarchy is of order $f_k/\tau_{\text{r.h.s.}}$. As in the case $k = 2$ discussed above, the left member of the evolution equation for $k \geq 3$ involves a term of order f_k/τ_{c}, much larger (by a factor $1/n r_0^3$) that the contribution of the right hand side. Accordingly, in each of the higher equations of the hierarchy the right member can be set to zero, so that these equations all decouple from each other. That is, the diluteness assumption alone already ensures that every f_k with $k \neq 2$ is evolving independently of any other reduced density.

On the other hand, the equation for f_1, Eq. (IV.76a) does not involve τ_{c} in its left member, so that there is no rationale for dropping the right hand side. Accordingly, f_1 is still affected by f_2.

Remark: The time scale $\tau_{\text{r.h.s.}}$ is actually of the same order as the mean free time τ_{mfp} between two successive collisions of a given particle.

IV.A.2 Transition to the Boltzmann equation

As stated above, the equation (IV.76a) involves the time scales τ_{e}, τ_{s} and $\tau_{\text{r.h.s.}} \sim \tau_{\text{c}}/n r_0^3$, while Eq. (IV.80) includes in addition a term with the characteristic time scale τ_{c}, much smaller than the other three. Accordingly, f_2 will evolve much more rapidly than f_1. One may thus assume that on the time scales relevant for the evolution of the single-particle density, the two-particle one has already reached some (quasi)stationary regime, i.e. $\partial f_2/\partial t \sim 0$ in Eq. (IV.80).

Consider now the other terms on the left hand side of Eq. (IV.80). First, we may write

$$\vec{F}_1 \cdot \vec{\nabla}_{\vec{p}_1} + \vec{F}_2 \cdot \vec{\nabla}_{\vec{p}_2} = \frac{1}{2}\big(\vec{F}_1 + \vec{F}_2\big) \cdot \big(\vec{\nabla}_{\vec{p}_1} + \vec{\nabla}_{\vec{p}_2}\big) + \frac{1}{2}\big(\vec{F}_1 - \vec{F}_2\big) \cdot \big(\vec{\nabla}_{\vec{p}_1} - \vec{\nabla}_{\vec{p}_2}\big). \tag{IV.81}$$

Let us perform a change of variables from $\vec{p}_1$, $\vec{p}_2$ to the total momentum and half relative momentum of the particle pair

$$\vec{P} = \vec{p}_1 + \vec{p}_2 \quad , \quad \vec{q} = \frac{1}{2}\big(\vec{p}_1 - \vec{p}_2\big), \tag{IV.82a}$$

and accordingly the change of spatial variables

$$\vec{R} = \frac{1}{2}\big(\vec{r}_1 + \vec{r}_2\big) \quad , \quad \vec{\rho} = \vec{r}_2 - \vec{r}_1. \tag{IV.82b}$$

One checks that the first term on the right hand side of Eq. (IV.81) involves $\vec{\nabla}_{\vec{P}} = \frac{1}{2}\big(\vec{\nabla}_{\vec{p}_1} + \vec{\nabla}_{\vec{p}_2}\big)$, i.e. will describe the dependence of f_2 on the motion of the center of mass $\vec{R}$ of the particle pair. This involves the "external" time scale τ_{e} and is to a large extent independent of the dependence of f_2 on the relative coordinates $\vec{\rho}$, $\vec{q}$, which involves more rapid time scales. Accordingly, we shall drop the corresponding term.

In agreement with this focus on small distances and time scales, the external forces $\vec{F}_1$, $\vec{F}_2$ in the second contribution to the right hand side of Eq. (IV.81) are evaluated at two positions separated by at most a few r_0. Remembering that the external potential is assumed to vary slowly in space, $\vec{F}_1 \simeq \vec{F}_2$, or equivalently $\vec{F}_1 - \vec{F}_2 \simeq 0$, so that we shall also neglect this second term. All in all, the equation (IV.80) governing f_2 in the stationary regime relevant for the evolution of f_1 thus becomes

$$\Big[\vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1} + \vec{v}_2\cdot\vec{\nabla}_{\vec{r}_2} + \vec{K}_{12}\cdot\big(\vec{\nabla}_{\vec{p}_1} - \vec{\nabla}_{\vec{p}_2}\big)\Big] f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2) \simeq 0.$$

Rewriting the first two terms in the angular brackets as

$$\vec{v}_1\cdot\vec{\nabla}_{\vec{r}_1} + \vec{v}_2\cdot\vec{\nabla}_{\vec{r}_2} = \frac{1}{2}\big(\vec{v}_1+\vec{v}_2\big)\cdot\big(\vec{\nabla}_{\vec{r}_1}+\vec{\nabla}_{\vec{r}_2}\big) + \frac{1}{2}\big(\vec{v}_1-\vec{v}_2\big)\cdot\big(\vec{\nabla}_{\vec{r}_1}-\vec{\nabla}_{\vec{r}_2}\big) = \frac{1}{2}\big(\vec{v}_1+\vec{v}_2\big)\cdot\vec{\nabla}_{\vec{R}} + \big(\vec{v}_2-\vec{v}_1\big)\cdot\vec{\nabla}_{\vec{\rho}}$$

and neglecting again the term involving the center-of-mass motion eventually yields

$$-\vec{K}_{12}\cdot\big(\vec{\nabla}_{\vec{p}_1} - \vec{\nabla}_{\vec{p}_2}\big) f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{\rho},\vec{p}_2) \simeq \big(\vec{v}_2-\vec{v}_1\big)\cdot\vec{\nabla}_{\vec{\rho}} f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{\rho},\vec{p}_2). \tag{IV.83}$$

If one now integrates this equation over all positions and momenta of particle 2, i.e. over all values of $\vec{\rho}$ (at fixed $\vec{r}_1$) and $\vec{p}_2$, the term in $-\vec{\nabla}_{\vec{p}_2} f_2$ on the left hand side is a total derivative, whose integral over $\vec{p}_2$ will simply vanish. That is, one has

$$-\int \vec{K}_{12}\cdot\vec{\nabla}_{\vec{p}_1} f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{\rho},\vec{p}_2)\,\mathrm{d}^3\vec{\rho}\,\mathrm{d}^3\vec{p}_2 \simeq \int \big(\vec{v}_2-\vec{v}_1\big)\cdot\vec{\nabla}_{\vec{\rho}} f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{\rho},\vec{p}_2)\,\mathrm{d}^3\vec{\rho}\,\mathrm{d}^3\vec{p}_2. \tag{IV.84}$$

One recognizes on the left hand side the right member—up to a trivial change of integration variable—of the first equation (IV.76a) of the BBGKY hierarchy.

The integrand of the right member in Eq. (IV.84) is the (partial) derivative of f_2 with respect to $\vec{\rho}$ along the direction of the relative velocity $\vec{v}_2 - \vec{v}_1$, i.e. of the relative motion of the particle pair. Let us choose a Cartesian coordinate system in which the z-axis is along that direction, with z negative resp. positive before resp. after the collision of the particles. Let the *impact vector* $\vec{b}$ denote the projection of $\vec{\rho}$ on the plane orthogonal to the z-axis, so that $\vec{\rho} = \vec{b} + z\,\vec{e}_z$, where $\vec{e}_z$ denotes the unit vector along the z-axis, and $\mathrm{d}^3\vec{\rho} = \mathrm{d}^2\vec{b}\,\mathrm{d}z$. We wish to perform the integral over z only.

A priori, z takes all values from $-\infty$ to ∞. Yet in practice, the $\vec{\rho}$-dependence of f_2 is only significant with a region of the order of a few interaction ranges, i.e. for $|\vec{\rho}| \lesssim r_0$, after which it reaches its asymptotic value. That is, the limits of f_2 for z going to $\pm\infty$ are already reached at points $z^\pm \approx \pm r_0$. Denoting by $(\partial f_1/\partial t)_{\text{coll.}}$ the integral in Eq. (IV.84), one thus has

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll.}} \simeq \int \big|\vec{v}_2-\vec{v}_1\big| \big[f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{b}+z^+\,\vec{e}_z,\vec{p}_2) - f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{b}+z^-\,\vec{e}_z,\vec{p}_2)\big]\,\mathrm{d}^2\vec{b}\,\mathrm{d}^3\vec{p}_2.$$

The first f_2 in the angular brackets corresponds to the situation after a collision, while the second, with a minus sign, describes the two-particle density before the particles have scattered on each other. Invoking the assumption of molecular chaos will allow us to replace the latter by a product of single-particle densities; yet this cannot be done for the former, since just after a collision two particles are no longer uncorrelated. Using the determinism of the trajectories of classical particles, f_2 after a collision can be related to f_2 before the scattering, yet for different momenta $\vec{p}_3$, $\vec{p}_4$ instead of $\vec{p}_1$, $\vec{p}_2$:

$$f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{b}+z^+\,\vec{e}_z,\vec{p}_2) = f_2(t,\vec{r}_1,\vec{p}_3,\vec{r}_1+\vec{b}+z^-\,\vec{e}_z,\vec{p}_4),$$

leading to

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll.}} \simeq \int \big|\vec{v}_2-\vec{v}_1\big| \big[f_2(t,\vec{r}_1,\vec{p}_3,\vec{r}_1+\vec{b}+z^-\,\vec{e}_z,\vec{p}_4) - f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{b}+z^-\,\vec{e}_z,\vec{p}_2)\big]\,\mathrm{d}^2\vec{b}\,\mathrm{d}^3\vec{p}_2, \tag{IV.85}$$

where $\vec{p}_3$ and $\vec{p}_4$ are in fact both function of $\vec{p}_1$, $\vec{p}_2$ and the impact vector $\vec{b}$. Taking into account kinetic energy and momentum conservation, only two out of the six degrees of freedom of $\vec{p}_3$, $\vec{p}_4$

are actually independent: as the reader knows from elementary kinematics, one may choose two angles(60) θ, φ or equivalently a solid angle Ω. At fixed $\vec{p}_1$ and $\vec{p}_2$, the latter is a one-to-one function of the impact vector, $\Omega(\vec{b})$. Considering the inverse function $\vec{b}(\Omega)$, one can perform a change of integration variable from $\vec{b}$ to Ω in Eq. (IV.85). The Jacobian of the transformation is precisely the differential cross section $\mathrm{d}^2\sigma/\mathrm{d}^2\Omega$, resulting in

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll.}} \simeq \int |\vec{v}_2 - \vec{v}_1| \left[f_2(t,\vec{r}_1,\vec{p}_3,\vec{r}_1+\vec{b}+z^-\,\vec{\mathrm{e}}_z,\vec{p}_4) - f_2(t,\vec{r}_1,\vec{p}_1,\vec{r}_1+\vec{b}+z^-\,\vec{\mathrm{e}}_z,\vec{p}_2)\right] \frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega}\,\mathrm{d}^2\Omega\,\mathrm{d}^3\vec{p}_2.$$

Eventually, we can on the one hand look at a coarse-grained version of position space, which results in the replacement of $f_2(t,\vec{r}_1,\vec{p}_j,\vec{r}_1+\vec{b}+z^-\,\vec{\mathrm{e}}_z,\vec{p}_k)$ by $\bar{\mathsf{f}}_2(t,\vec{r}_1,\vec{p}_j,\vec{r}_1,\vec{p}_k)/(2\pi\hbar)^6$ for both terms in the angular brackets on the right hand side, and that of f_1 by $\bar{\mathsf{f}}_1/(2\pi\hbar)^3$ in the left member. On the other hand, since both two-particle densities are now those before a collision, we can invoke the molecular-chaos assumption, leading to

$$\left(\frac{\partial \bar{\mathsf{f}}_1}{\partial t}\right)_{\text{coll.}} \simeq \int |\vec{v}_2 - \vec{v}_1| \Big[\bar{\mathsf{f}}_1(t,\vec{r}_1,\vec{p}_3)\,\bar{\mathsf{f}}_1(t,\vec{r}_1,\vec{p}_4) - \bar{\mathsf{f}}_1(t,\vec{r}_1,\vec{p}_1)\,\bar{\mathsf{f}}_1(\vec{r}_1,\vec{p}_2)\Big] \frac{\mathrm{d}^2\sigma}{\mathrm{d}^2\Omega}\,\mathrm{d}^2\Omega\,\frac{\mathrm{d}^3\vec{p}_2}{(2\pi\hbar)^3}, \qquad \text{(IV.86)}$$

which is exactly the collision term as given in Eq. (IV.15b).

---

(60) usually those giving the orientation of a momentum transfer like $\vec{p}_4 - \vec{p}_2$.

CHAPTER V

Brownian motion

In this chapter, we study the very general paradigm provided by Brownian motion. Originally, this motion is that a "heavy" particle, called *Brownian particle*, immersed in a fluid of much lighter particles—in Robert Brown's(be) original observations, this was some pollen grain in water. Due to the successive collisions with the fluid constituents, the Brownian particle is constantly moving, going in always changing, apparently random directions, even if the fluid itself is at rest: the position $\vec{x}(t)$ and the velocity $\vec{v}(t)$ of the Brownian particle are then naturally modeled as stochastic processes, driven by a fluctuating force.

The interest of this rather specific physical problem lies in the fact that the dynamical equation governing the motion of the Brownian particle actually also applies to many stochastic collective properties of a macroscopic system as they approach their equilibrium values. Accordingly, the techniques used for solving the initial question extend to a much wider class of problems in nonequilibrium statistical physics and even beyond. This notwithstanding, we shall throughout this chapter retain the terminology of the original physical problem.

In Sec. V.1, we introduce the approach pioneered by Paul Langevin,(bf) which describes the dynamics of the Brownian particle velocity on time scales larger than the typical autocorrelation time of the fluctuating force acting on the particle by explicitly solving the evolution equation for given initial conditions. We then adopt in Sec. V.2 an alternative description, in which we rather focus on the time evolution of the probability distribution of the velocity. That approach is quite generic and can be used for any Markov process, so that we discuss a straightforward extension in an appendix (V.A). Finally, we investigate a generalization of the Langevin model, in which the friction force exerted by the fluid on the Brownian motion is non-local in time (Sec. V.3).

For the sake of simplicity, we consider in this chapter classical one-dimensional Brownian motion only. The generalization to motion in two or more dimensions is straightforward. In the following chapter (Sec. VI.4.3), we shall introduce quantum-mechanical models analogous to classical Brownian motion, in that the spectral properties of some of their operators are similar to those of the Langevin model of Sec. V.1 or the generalization of Sec. V.3.

V.1 Langevin dynamics

Following P. Langevin's modeling, the dynamics of a Brownian particle much heavier than the constituents of the medium in which it evolves can be viewed as resulting from the influence of two complementary forces, namely an instantaneous friction force and a fluctuating force. After writing down the corresponding dynamical equation for the velocity of the Brownian particle (Sec. V.1.1), we study its solution for given initial conditions (Sec. V.1.2), as well as the resulting time evolution of the displacement from the initial position (Sec. V.1.3). We then turn in Sec. V.1.4 to the dynamics of the fluctuations of the velocity for a Brownian particle at equilibrium with its environment. Eventually, anticipating on applications of the Brownian-motion paradigm to other problems, we introduce the spectral function associated to the Langevin model at equilibrium (Sec. V.1.5), as well as the linear response of the model to an external force (Sec. V.1.6).

(be)R. BROWN, 1773–1858 (bf)P. LANGEVIN, 1872–1946

V.1.1 Langevin model

Let M denote the mass of the Brownian particle and $v(t)$ its velocity.

V.1.1 a Langevin equation

The classical model introduced by P. Langevin [44] consists in splitting the total force exerted by the fluid particles on the Brownian particle into two contributions:

- First, a Brownian particle in motion with a velocity v with respect to the fluid sees more fluid particles coming from the direction in which it is moving as from the other direction. The larger v is, the more pronounced the imbalance is.

  To account for this effect, one introduces a friction force opposed to the instantaneous direction of motion—i.e. to the velocity *at the same instant*—and increasing with velocity. The simplest possibility is that of a force proportional to $v(t)$, which will be denoted as $-M\gamma v(t)$ with $\gamma > 0$.

  > This actually corresponds to the viscous force exerted by a Newtonian fluid on an immersed body, in which case the "friction coefficient" $M\gamma$ is proportional to the shear viscosity η of the fluid.

- The fluid particles also exert a *fluctuating force* $F_{\rm L}(t)$, due to their random individual collisions with the Brownian particle. This *Langevin force*, also referred to as *noise term*, will be assumed to be independent of the kinematic variables (position and velocity) of the Brownian particle.

Since both friction and noise terms introduced by this decomposition are actually two aspects of a single underlying phenomenon—the microscopic scatterings of fluid particles on the Brownian particle—, one can reasonably anticipate the existence of a relationship between them, i.e. between γ and a characteristic property of $F_{\rm L}(t)$, as we shall indeed find in § V.1.2 b.

Assuming for the moment that there is no additional force acting on the Brownian particle, i.e. that it is "free",(61) the equation of motion reads

$$\boxed{M\frac{{\rm d}v(t)}{{\rm d}t} = -M\gamma v(t) + F_{\rm L}(t) \quad \text{with} \quad v(t) = \frac{{\rm d}x(t)}{{\rm d}t}.} \tag{V.1}$$

This *Langevin equation* is an instance of a linear *stochastic differential equation*, i.e. an equation including a randomly varying term—here $F_{\rm L}(t)$—with given statistical properties—which we shall specify in the next paragraph in the case of the Langevin force. The solution $v(t)$ to such an equation for a given initial condition is itself a stochastic process.

> Accordingly, one should distinguish—although we shall rather sloppily use the same notations—between the stochastic processes $F_{\rm L}(t)$, $v(t)$ and below $x(t)$, and their respective realizations. If $F_{\rm L}(t)$ is a realization of the corresponding stochastic process, then Eq. (V.1) is an ordinary (linear) differential equation for $v(t)$, including a perfectly deterministic term $F_{\rm L}$. Its solution $v(t)$ for a given initial condition is a well-determined function in the usual sense.
>
> The reader should keep in mind this dual meaning of the notations when going through the following Secs. V.1.2–V.1.6.

Remarks:

∗ Strictly speaking, the *classical* collisions of the fluid particles on the Brownian particle are not random, but entirely governed by the deterministic Liouville equation for the total system. The randomness of the macroscopically-perceived Langevin force comes from the fact that it is in practice impossible to fully characterize the microstate of the fluid, which has to be described statistically.

---

(61) This assumption will be relaxed in Sec. V.1.6.

∗ As mentioned at the end of Sec. I.2.1, the relaxation of a thermodynamic extensive variable towards its equilibrium value can be described, provided the system is near equilibrium, by a first order linear differential equation. Such an extensive variable is in fact the expectation value of the sum over many particles of a microscopic quantity, so that it can fluctuate around its average. These fluctuations can be modeled by adding a fluctuating (generalized) force in the evolution equation (I.33), which then becomes of the Langevin type (V.1):

$$\frac{\mathrm{d}\Delta\mathcal{X}_a(t)}{\mathrm{d}t} = -\sum_c \lambda_{ac}\Delta\mathcal{X}_c(t) + F_{a,\mathrm{L}}, \tag{V.2}$$

with $F_{a,\mathrm{L}}$ a fluctuating term.

V.1.1 b Properties of the noise term

The fluid in which the Brownian particle is immersed is assumed to be in a stationary state, for instance in thermodynamic equilibrium—in which case it is also in mechanical equilibrium, and thus admits a global rest frame, with respect to which $v(t)$ is measured. Accordingly, the Langevin force acting upon the particle is described by a stationary process, that is, the single-time average $\langle F_\mathrm{L}(t)\rangle$ is time-independent, while the two-point average $\langle F_\mathrm{L}(t)F_\mathrm{L}(t')\rangle$ only depends on the difference $t'-t$.

In order for the particle to remain (on average) motionless when it is at rest with respect to the fluid, the single-time average of the Langevin force should actually vanish. Since we assumed $F_\mathrm{L}(t)$ to be independent of the particle velocity, this remains true even when the Brownian particle is moving:

$$\big\langle F_\mathrm{L}(t)\big\rangle = 0. \tag{V.3a}$$

In consequence, the autocorrelation function (C.5) simplifies to

$$\kappa(\tau) = \big\langle F_\mathrm{L}(t)F_\mathrm{L}(t+\tau)\big\rangle. \tag{V.3b}$$

As always for stationary processes, $\kappa(\tau)$ only depends on $|\tau|$. $\kappa(\tau)$ is assumed to be integrable, with

$$\int_{-\infty}^{\infty}\kappa(\tau)\,\mathrm{d}\tau \equiv 2D_vM^2, \tag{V.3c}$$

which defines the parameter D_v.

Let τ_c be the *autocorrelation time* over which $\kappa(\tau)$ decreases. τ_c is typically of the order of the time interval between two collisions of the fluid particles on the Brownian particle. If τ_c happens to be much smaller than all other time scales in the system, then the autocorrelation function can meaningfully be approximated by a Dirac distribution

$$\kappa(\tau) = 2D_vM^2\delta(\tau). \tag{V.3d}$$

More generally, one may write

$$\kappa(\tau) = 2D_vM^2\delta_{\tau_c}(\tau), \tag{V.3e}$$

where δ_{τ_c} is an even function, peaked around the origin with a typical width of order τ_c, and whose integral over $\mathbb{R}$ equals 1.

Remarks:

∗ Throughout this chapter, expectation values—as e.g. in Eq. (V.3a) or (V.3b)—represent averages over different microscopic configurations of the "fluid" with the same macroscopic properties.

∗ In the case of multidimensional Brownian motion, one usually assumes that the correlation matrix $\kappa^{ij}(\tau)$ of the Cartesian components of the fluctuating force is diagonal, which can be justified in the case the underlying microscopic dynamics involve interactions depending only on inter-particle distances (see § VI.4.3 b).

∗ Equations (V.3) constitute "minimal" assumptions, which will allow us hereafter to compute the first and second moments of the stochastic processes $v(t)$ and $x(t)$, but do not fully specify the

Langevin force $F_{\rm L}(t)$. A possibility for determining entirely the statistical properties of $F_{\rm L}(t)$ could be to assume that it is *Gaussian*, in addition to stationary.

> Since $F_{\rm L}(t)$ actually results from summing a large number of stochastic processes—the microscopic forces due to individual collisions with the fluid particles—with the same probability distribution, this assumption of Gaussianity simply reflects the central limit theorem (Appendix B.5).

V.1.2 Relaxation of the velocity

We now wish to solve the Langevin equation (V.1), assuming that at the initial time $t = t_0$, the velocity of the Brownian particle is fixed, $v(t_0) = v_0$.

With that initial condition, one finds at once that the solution to the Langevin equation for $t > t_0$ reads[62]

$$v(t) = v_0\,{\rm e}^{-\gamma(t-t_0)} + \frac{1}{M}\int_{t_0}^{t} F_{\rm L}(t')\,{\rm e}^{-\gamma(t-t')}\,{\rm d}t' \quad \text{for } t > t_0. \tag{V.4}$$

Since $F_{\rm L}(t')$ is a stochastic process, so is $v(t)$. The first moments of its distribution can easily be computed.

Remark: The integral on the right-hand side of the previous equation has to be taken with a grain of salt, as it is not clear whether $F_{\rm L}$ is integrable.

V.1.2 a Average velocity

Averaging Eq. (V.4) over an ensemble of realizations, one finds thanks to property (V.3a)

$$\langle v(t)\rangle = v_0\,{\rm e}^{-\gamma(t-t_0)} \quad \text{for } t > t_0. \tag{V.5}$$

That is, the average velocity relaxes exponentially to 0 with a characteristic *relaxation time*

$$\tau_{\rm r} \equiv \frac{1}{\gamma}. \tag{V.6}$$

Since its average depends on time, $v(t)$ is not a stationary process.

V.1.2 b Variance of the velocity. Fluctuation–dissipation theorem

Recognizing the average velocity (V.5) in the first term on the right-hand side of Eq. (V.4), one also obtains at once the variance

$$\sigma_v(t)^2 \equiv \Big\langle \big[v(t) - \langle v(t)\rangle\big]^2 \Big\rangle = \frac{1}{M^2}\int_{t_0}^{t}\!\int_{t_0}^{t} \big\langle F_{\rm L}(t')F_{\rm L}(t'')\big\rangle\,{\rm e}^{-\gamma(t-t')}\,{\rm e}^{-\gamma(t-t'')}\,{\rm d}t'\,{\rm d}t'' \quad \text{for } t > t_0. \tag{V.7}$$

If the simplified form (V.3d) of the autocorrelation function of the Langevin force holds—which for the sake of consistency necessitates at least $\tau_{\rm c} \ll \tau_{\rm r}$—, this variance becomes

$$\sigma_v(t)^2 = 2D_v\int_{t_0}^{t} {\rm e}^{-2\gamma(t-t')}\,{\rm d}t' = \frac{D_v}{\gamma}\big(1 - {\rm e}^{-2\gamma(t-t_0)}\big) \quad \text{for } t > t_0. \tag{V.8}$$

σ_v^2 thus vanishes at $t = t_0$—the initial condition is the same for all realizations—, then grows, at first almost linearly

$$\sigma_v(t)^2 \simeq 2D_v(t - t_0) \quad \text{for } 0 \le t - t_0 \ll \tau_{\rm r}, \tag{V.9}$$

before saturating at large times

$$\sigma_v(t)^2 \simeq \frac{D_v}{\gamma} \quad \text{for } t - t_0 \gg \tau_{\rm r}. \tag{V.10}$$

Equation (V.9) suggests that D_v is a *diffusion coefficient* in velocity space.

---

(62) Remember that this expression, as well as many other ones in this section, holds both for realizations of the stochastic processes at play—in which case the meaning is clear—and by extension for the stochastic processes themselves.

Remark: The above results remain valid even if the simplified form (V.3d) does not hold, provided the discussion is restricted to times t siginificantly larger than the autocorrelation time τ_c of the Langevin force.

Using definition (V.3b), the right-hand side of Eq. (V.7) can be recast as

$$\frac{\mathrm{e}^{-2\gamma t}}{M^2}\int_{t_0}^{t}\int_{t_0}^{t}\kappa(t'-t'')\,\mathrm{e}^{\gamma(t'+t'')}\,\mathrm{d}t'\,\mathrm{d}t'' = \frac{\mathrm{e}^{-2\gamma t}}{2M^2}\int_{t_0-t}^{t-t_0}\kappa(\tau)\,\mathrm{d}\tau\int_{2t_0}^{2t}\mathrm{e}^{\gamma(t'+t'')}\,\mathrm{d}(t'+t'').$$

The integral over $t'+t''$ is straightforward, while for that over τ, one may for $t-t_0 \gg \tau_c$ extend the boundaries to $-\infty$ and $+\infty$ without changing much the result. Invoking then the normalization (V.3c), one recovers the variance (V.8). □

From Eq. (V.5), the average velocity vanishes at large times, so that the variance (V.10) equals the mean square velocity. That is, the average kinetic energy of the Brownian particle tends at large times towards a fixed value

$$\langle E(t)\rangle \simeq \frac{MD_v}{2\gamma} \quad \text{for } t-t_0 \gg \tau_r. \tag{V.11}$$

In that limit, the Brownian particle is in equilibrium with the surrounding fluid. If the latter is in thermodynamic equilibrium at temperature T—one then refers to it as a *thermal bath* or *thermal reservoir*—then the average energy of the Brownian particle is according to the equipartition theorem equal to $\frac{1}{2}k_BT$, which yields

$$\boxed{D_v = \frac{k_BT}{M}\gamma.} \tag{V.12}$$

This identity relates a quantity associated with fluctuations—the diffusion coefficient D_v, see Eq. (V.9)—with a coefficient modeling dissipation, namely γ. This is thus an example of the more general *fluctuation–dissipation theorem* discussed in Sec. VI.3.4.

Since D_v characterizes the statistical properties of the stochastic Langevin force [Eq. (V.3c)], Eq. (V.12) actually relates the latter to the friction force.

V.1.3 Evolution of the position of the Brownian particle. Diffusion

Until now, we have only investigated the velocity of the Brownian particle. Instead, one could study its position $x(t)$, or equivalently its displacement from an initial position $x(t_0) = x_0$ at time $t = t_0$.

Integrating the velocity (V.4) from the initial instant until time t yields the formal solution

$$x(t) = x_0 + \frac{v_0}{\gamma}\big(1-\mathrm{e}^{-\gamma(t-t_0)}\big) + \frac{1}{M}\int_{t_0}^{t} F_{\mathrm{L}}(t')\,\frac{1-\mathrm{e}^{-\gamma(t-t')}}{\gamma}\,\mathrm{d}t' \quad \text{for } t>t_0. \tag{V.13}$$

$x(t)$, and in turn the displacement $x(t)-x_0$, is also a stochastic process, whose first two moments we shall now compute.

V.1.3 a Average displacement

First, the average position is simply

$$\langle x(t)\rangle = x_0 + \frac{v_0}{\gamma}\big(1-\mathrm{e}^{-\gamma(t-t_0)}\big) \quad \text{for } t>t_0, \tag{V.14}$$

where the vanishing of the expectation value of the Langevin force was used.
For $t-t_0 \ll \tau_r$, $\langle x(t)\rangle \simeq x_0 + v_0(t-t_0)$, i.e. the motion is approximately linear. In the opposite limit $t-t_0 \gg \tau_r$, the mean displacement $\langle x(t)\rangle - x_0$ tends exponentially towards the asymptotic value v_0/γ.

V.1.3 b Variance of the displacement

The first two terms in the right member of Eq. (V.13) are exactly the average position (V.14), that is, the last term is precisely $x(t) - \langle x(t)\rangle$, or equivalently $[x(t)-x_0] - \langle x(t) - x_0\rangle$. The variance of the position is thus equal to the variance of the displacement, and is given by

$$\sigma_x(t)^2 = \frac{1}{M^2\gamma^2}\int_{t_0}^{t}\!\int_{t_0}^{t} \langle F_{\rm L}(t')F_{\rm L}(t'')\rangle\big[1-{\rm e}^{-\gamma(t-t')}\big]\big[1-{\rm e}^{-\gamma(t-t'')}\big]\,{\rm d}t'\,{\rm d}t'' \quad \text{for } t > t_0. \tag{V.15}$$

When the autocorrelation function of the Langevin force can be approximated by the simplified form (V.3d), this yields

$$\sigma_x(t)^2 = \frac{2D_v}{\gamma^2}\int_{t_0}^{t}\big[1-{\rm e}^{-\gamma(t-t')}\big]^2\,{\rm d}t' = \frac{2D_v}{\gamma^2}\left(t - \frac{2-2{\rm e}^{-\gamma(t-t_0)}}{\gamma} + \frac{1-{\rm e}^{-2\gamma(t-t_0)}}{2\gamma}\right) \quad \text{for } t > t_0. \tag{V.16}$$

This variance vanishes at $t = t_0$—the initial condition is known with certainty—, grows as t^3 for times $0 \leq t - t_0 \ll \tau_{\rm r}$, then linearly at large times

$$\sigma_x(t)^2 \simeq \frac{2D_v}{\gamma^2}(t-t_0) \quad \text{for } t - t_0 \gg \tau_{\rm r}. \tag{V.17}$$

Since Eq. (V.16) also represents the variance of the displacement, one finds under consideration of Eq. (V.13)

$$\begin{aligned}\Big\langle \big[x(t)-x_0\big]^2\Big\rangle &= \sigma_x(t)^2 + \langle x(t) - x_0\rangle^2 \\ &= \sigma_x(t)^2 + \frac{v_0^2}{\gamma^2}\big(1-{\rm e}^{-\gamma(t-t_0)}\big)^2 \simeq \frac{2D_v}{\gamma^2}(t-t_0) \quad \text{for } t-t_0 \gg \tau_{\rm r}.\end{aligned} \tag{V.18}$$

The last two equations show that the position of the Brownian particle behaves as the solution of a diffusion equation at large times, with a diffusion coefficient (in position space)

$$D = \frac{D_v}{\gamma^2}, \tag{V.19}$$

(cf. § I.2.3 b).

V.1.3 c Viscous limit. Einstein relation

In the limit $M \to 0$, $\gamma \to \infty$ at constant product $\eta_v \equiv M\gamma$, which physically amounts to neglecting the effect of inertia ($M\,{\rm d}v/{\rm d}t$) compared to that of friction ($-\eta_v v$)—hence the denomination "viscous limit"—, the Langevin equation (V.1) becomes

$$\boxed{\eta_v \frac{{\rm d}x(t)}{{\rm d}t} = F_{\rm L}(t).} \tag{V.20a}$$

In that limit, the autocorrelation function of the fluctuating force in the limit of negligibly small autocorrelation time is denoted as

$$\kappa(\tau) = 2D\eta_v^2\,\delta(\tau), \tag{V.20b}$$

which defines the coefficient D.

In this approach, the displacement can directly be computed by integrating Eq. (V.20a) with the initial condition $x(t_0) = x_0$, yielding

$$x(t) - x_0 = \frac{1}{\eta_v}\int_{t_0}^{t} F_{\rm L}(t')\,{\rm d}t' \quad \text{for } t > t_0. \tag{V.21}$$

With the autocorrelation function (V.20b), this gives at once

$$\Big\langle \big[x(t)-x_0\big]^2\Big\rangle = 2D(t-t_0) \quad \text{for } t \geq t_0, \tag{V.22}$$

which now holds at any time $t \geq t_0$, not only in the large time limit as in Eq. (V.18). That is, the motion of the Brownian particle is now a diffusive motion at all times.

Combining now Eq, (V.19) with the fluctuation–dissipation relation (V.12) and the relation $\eta_v = m\gamma$, one obtains

$$D = \frac{k_B T}{\eta_v}. \tag{V.23}$$

If the Brownian particles are charged, with electric charge q, then the friction coefficient η_v is related to the electrical mobility $\mu_{\text{el.}}$ by $\mu_{\text{el.}} = q/\eta_v$ (see Sec. V.1.6 below), so that Eq. (V.23) becomes the *Einstein relation*

$$D = \frac{k_B T}{q}\,\mu_{\text{el.}} \tag{V.24}$$

[cf. Eq. (I.48)].

V.1.4 Autocorrelation function of the velocity at equilibrium

In this subsection and the following one, we assume that the Brownian particle is in equilibrium with the surrounding environment. This amounts to considering that a large amount of time has passed since the instant t_0 at which the initial condition was fixed, or equivalently that t_0 is far back in the past, $t_0 \to -\infty$.

Taking the latter limit in Eq. (V.4), the velocity at time t reads

$$v(t) = \frac{1}{M}\int_{-\infty}^{t} F_{\text{L}}(t'')\,\mathrm{e}^{-\gamma(t-t'')}\,\mathrm{d}t'', \tag{V.25}$$

where we have renamed the integration variable t'' for later convenience. As could be anticipated, v_0 no longer appears in this expression: the initial condition has been "forgotten".

One easily sees that the average value $\langle v(t)\rangle$ vanishes, and is thus in particular time-independent. More generally, one can check with a straightforward change of variable that $v(t)$ at equilibrium is a stationary stochastic process, thanks to the assumed stationarity of $F_{\text{L}}(t)$. We shall now compute the autocorrelation function of $v(t)$, which characterizes its fluctuations.

Starting from the velocity (V.25), one first finds the correlation function between the velocity and the fluctuating force

$$\langle F_{\text{L}}(t_1)v(t_2)\rangle = \frac{1}{M}\int_{-\infty}^{t_2} \langle F_{\text{L}}(t_1)F_{\text{L}}(t'')\rangle \mathrm{e}^{-\gamma(t_2-t'')}\,\mathrm{d}t''.$$

In the case where the simplified form (V.3d) of the autocorrelation function of the Langevin force holds, this becomes

$$\big\langle F_{\text{L}}(t_1)v(t_2)\big\rangle = 2D_vM\int_{-\infty}^{t_2}\delta(t_1-t'')\,\mathrm{e}^{-\gamma(t_2-t'')}\,\mathrm{d}t'' = \begin{cases} 2D_vM\,\mathrm{e}^{-\gamma(t_2-t_1)} & \text{for } t_1<t_2, \\ 0 & \text{for } t_1>t_2. \end{cases} \tag{V.26}$$

That is, the velocity of the Brownian particle at time t is only correlated to past values of the Langevin force, and the correlation dies out on a typical time scale of order $\gamma^{-1} = \tau_{\text{r}}$.

The autocorrelation function for the velocity is then easily deduced from

$$\big\langle v(t)v(t')\big\rangle = \frac{1}{M}\int_{-\infty}^{t} \big\langle F_{\text{L}}(t'')v(t')\big\rangle \mathrm{e}^{-\gamma(t-t'')}\,\mathrm{d}t'',$$

which follows from Eq. (V.25). In the regime where the approximation (V.26) is valid, that is neglecting the autocorrelation time τ_{c} of the Langevin force, this yields

$$\big\langle v(t)v(t')\big\rangle = \frac{D_v}{\gamma}\,\mathrm{e}^{-\gamma|t-t'|}. \tag{V.27}$$

This autocorrelation function only depends on the modulus of the time difference, as expected for a stationary stochastic process, and decreases exponentially with an autocorrelation time given by the relaxation time $\tau_{\rm r}$. Note that for $t' = t$, we recover the large-time limit (V.10) of the variance of the velocity.

If the environment is in thermal equilibrium at temperature T, relation (V.12) gives

$$\big\langle v(t)v(t')\big\rangle = \frac{k_BT}{M}\,{\rm e}^{-\gamma|t-t'|}. \tag{V.28}$$

Remark: Inspecting the average velocity (V.5) and autocorrelation function (V.27), one sees that they obey the same first-order linear differential equation, with the same characteristic relaxation time scale $\tau_{\rm r}$.

V.1.5 Harmonic analysis

In the regime in which the Brownian particle is in equilibrium with the fluid, the velocity $v(t)$ becomes a stationary stochastic process, as is the fluctuating force $F_{\rm L}(t)$ itself. One can thus apply to them the concepts introduced in Appendix C.3, and in particular introduce their Fourier transforms(63)

$$\tilde{F}_{\rm L}(\omega) \equiv \int F_{\rm L}(t)\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t, \qquad \tilde{v}(\omega) \equiv \int v(t)\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t. \tag{V.29}$$

In Fourier space, the Langevin equation (V.1) leads to the relation

$$\tilde{v}(\omega) = \frac{1}{M}\frac{1}{\gamma - {\rm i}\omega}\,\tilde{F}_{\rm L}(\omega). \tag{V.30}$$

One also introduces the respective *spectral densities* of the stochastic processes(63)

$$S_F(\omega) \equiv \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\Big\langle\big|\tilde{F}_{\rm L}(\omega)\big|^2\Big\rangle, \qquad S_v(\omega) \equiv \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\Big\langle\big|\tilde{v}(\omega)\big|^2\Big\rangle. \tag{V.31}$$

For these spectral densities, Eq. (V.30) yields at once the relation

$$S_v(\omega) = \frac{1}{M^2}\frac{1}{\gamma^2+\omega^2}\,S_F(\omega). \tag{V.32}$$

The spectral density of the velocity if thus simply related to that of the force, for which we shall consider two possibilities.

V.1.5 a White noise

A first possible ansatz for $S_F(\omega)$, compatible with the assumptions in § V.1.1 b, is that of a frequency-independent spectral density, i.e. of *white noise*

$$S_F(\omega) = S_F. \tag{V.33a}$$

According to the Wiener–Khinchin theorem (C.46), the autocorrelation function of the fluctuating force is then the Fourier transform of a constant, i.e. a Dirac distribution

$$\big\langle F_{\rm L}(t)F_{\rm L}(t+\tau)\big\rangle = \int S_F\,{\rm e}^{-{\rm i}\omega\tau}\,\frac{{\rm d}\omega}{2\pi} = S_F\,\delta(\tau). \tag{V.33b}$$

This thus constitutes the case in which Eq. (V.3d) holds, with $S_F = 2D_vM^2$.

With this simple form for $S_F(\omega)$, the spectral density (V.32) of the velocity is given by the Lorentzian distribution

$$S_v(\omega) = \frac{2D_v}{\gamma^2+\omega^2},$$

---

(63) Remember that, formally, one defines the transforms considering first the restrictions of the processes to a finite-size time interval of width $\mathcal{T}$, and at the end of calculations one takes the large-$\mathcal{T}$ limit. Here we drop the subscript $\mathcal{T}$ designating these restrictions to simplify the notations.

which after an inverse Fourier transformation yields for the autocorrelation function

$$\langle v(t)v(t+\tau)\rangle = \frac{D_v}{\gamma}\,\mathrm{e}^{-\gamma|\tau|}, \tag{V.34}$$

in agreement with what was already found in Eq. (V.27).

V.1.5 b Colored noise

While a frequency-independent white noise spectrum of Langevin-force fluctuations amounts to a vanishingly small autocorrelation time τ_c, a very wide—but not everywhere constant—spectrum will give rise to a finite τ_c. One then talks of *colored noise*.

Assume for instance that the density spectrum of the fluctuating force is given by a Lorentzian distribution centered on $\omega = 0$, with a large typical width ω_c, where the precise meaning of “large” will be specified later:

$$S_F(\omega) = S_F\,\frac{\omega_c^2}{\omega_c^2+\omega^2}. \tag{V.35a}$$

Since $S_F(\omega=0)$ equals the integral of the autocorrelation function $\kappa(\tau)$, condition (V.3c) leads to the identity $S_F = 2D_vM^2$. With the Wiener–Khinchin theorem (C.46) and relation (V.32), this corresponds to an autocorrelation function of the fluctuating force given by

$$\langle F_{\rm L}(t)F_{\rm L}(t+\tau)\rangle = \int 2D_vM^2\,\frac{\omega_c^2}{\omega_c^2+\omega^2}\,\mathrm{e}^{-\mathrm{i}\omega\tau}\,\frac{\mathrm{d}\omega}{2\pi} = D_vM^2\omega_c\,\mathrm{e}^{-\omega_c|\tau|}, \tag{V.35b}$$

i.e. the autocorrelation time of the Langevin force is $\tau_c = \omega_c^{-1}$.

Using Eq. (V.32), the spectral density of the velocity is

$$S_v(\omega) = 2D_v\omega_c^2\frac{1}{\gamma^2+\omega^2}\frac{1}{\omega_c^2+\omega^2} = \frac{2D_v\omega_c^2}{\omega_c^2-\gamma^2}\left(\frac{1}{\gamma^2+\omega^2}-\frac{1}{\omega_c^2+\omega^2}\right).$$

The autocorrelation function of the velocity then reads

$$\langle v(t)v(t+\tau)\rangle = \frac{D_v}{\gamma}\,\frac{\omega_c^2}{\omega_c^2-\gamma^2}\left(\mathrm{e}^{-\gamma|\tau|}-\frac{\gamma}{\omega_c}\mathrm{e}^{-\omega_c|\tau|}\right). \tag{V.36}$$

At small $|\tau| \ll \omega_c^{-1} \ll \gamma^{-1}$, this becomes

$$\langle v(t)v(t+\tau)\rangle \sim \frac{D_v}{\omega_c+\gamma}\left(\frac{\omega_c}{\gamma}-\frac{\gamma\omega_c}{2}\,\tau^2\right),$$

i.e. it departs quadratically from its value at $\tau = 0$. In particular, the singularity of the derivative at $\tau = 0$ which appears when τ_c is neglected [cf. Eq. (V.34)] has been smoothed out.

Remark: The autocorrelation function (V.36) actually involves two time scales, namely $\tau_c = \omega_c^{-1}$ and $\tau_r = \gamma^{-1}$. The Langevin model only makes sense if $\tau_c \ll \tau_r$, i.e. $\gamma \ll \omega_c$, in which case the second term in the brackets in the autocorrelation function is negligible, and the only remaining time scale for the fluctuations of velocity is τ_r. Velocity if thus a “slow” stochastic variable, compared to the more quickly evolving Langevin force. Physically, many collisions with lighter particles are necessary to change the velocity of the Brownian particle.

V.1.6 Response to an external force

Let us finally momentarily assume that the Brownian particle is submitted to an additional external force $F_{\rm ext.}(t)$, independent of its position and velocity. The equation of motion describing the Brownian particle dynamics then becomes

$$M\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -M\gamma v(t) + F_{\rm L}(t) + F_{\rm ext.}(t). \tag{V.37}$$

Averaging over many realizations, one obtains

$$M\frac{\mathrm{d}\langle v(t)\rangle}{\mathrm{d}t} = -M\gamma\langle v(t)\rangle + F_{\text{ext.}}(t), \tag{V.38}$$

where we have used property (V.3a) of the Langevin noise. This is now a linear ordinary differential equation, which is most easily solved by introducing the Fourier transforms

$$\langle\tilde{v}(\omega)\rangle \equiv \int \langle v(t)\rangle\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t, \qquad \tilde{F}_{\text{ext.}}(\omega) \equiv \int F_{\text{ext.}}(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t.$$

In Fourier space, Eq. (V.38) becomes $-\mathrm{i}\omega M\langle\tilde{v}(\omega)\rangle = -M\gamma\langle\tilde{v}(\omega)\rangle + \tilde{F}_{\text{ext.}}(\omega)$, i.e.

$$\langle\tilde{v}(\omega)\rangle = Y(\omega)\,\tilde{F}_{\text{ext.}}(\omega), \tag{V.39a}$$

with

$$Y(\omega) \equiv \frac{1}{M}\frac{1}{\gamma - \mathrm{i}\omega} \tag{V.39b}$$

the (complex) *admittance* of the Langevin model. That is, the (sample) average velocity of the Brownian particle responds *linearly* to the external force.

In the Hamilton function for the Brownian particle, the external force $F_{\text{ext.}}$ couples to the particle position x. Thus, the admittance (V.39b) represents, in the language which will be developed in Chapter VI, the generalized susceptibility $\tilde{\chi}_{vx}$ that characterizes the linear response of the velocity to a perturbation of the position. Accordingly, Eq. (V.40) below is nothing but relation (VI.65) with $K_{v\dot{x}} = K_{vv}$, in the classical case.

Consider now the autocorrelation function at equilibrium (V.27). Setting $t' = 0$ and assuming that the environment is at thermodynamic equilibrium with temperature T, in which case relation (V.12) holds, one finds

$$\big\langle v(t)v(0)\big\rangle = \frac{k_BT}{M}\,\mathrm{e}^{-\gamma|t|}.$$

The Fourier–Laplace transform of this autocorrelation function reads

$$\int_0^\infty \langle v(t)v(0)\rangle\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t = \frac{k_BT}{M}\frac{1}{\gamma - \mathrm{i}\omega},$$

that is, given expression (V.39b) of the admittance

$$\boxed{Y(\omega) = \frac{1}{k_BT}\int_0^\infty \big\langle v(t)v(0)\big\rangle \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t.} \tag{V.40}$$

This result relating the admittance to the autocorrelation function is again a form of the fluctuation–dissipation theorem.

If the Brownian particle carries an electric charge q, then one may as external force consider an electrostatic force $F_{\text{ext.}}(t) = q\mathscr{E}(t)$. Inserting this form in Eq. (V.38), one sees that the average velocity of the Brownian particle in the stationary regime is $\langle v\rangle = q\mathscr{E}/M\gamma$. Defining the electrical mobility as $\mu_{\text{el.}} \equiv \langle v\rangle/\mathscr{E}$, one finds

$$\mu_{\text{el.}} = \frac{q}{M\gamma} = q\,Y(\omega\!=\!0),$$

where the stationary regime obviously corresponds to the vanishing-frequency limit.

V.2 Fokker–Planck equation

In this Section, we analyze the Langevin model of Sec. V.1 by adopting a different view of the dynamics of a Brownian particle in an environment. Instead of focusing on the solution $v(t)$ of the Langevin equation for a given initial condition, we rather investigate the dynamics of the *velocity probability density* $f(t,v)$, such that $f(t,v)\,\mathrm{d}v$ is the probability that at time t the Brownian particle velocity lies between v and $v+\mathrm{d}v$.

We first argue in Sec. V.2.1 that on time scales larger than the autocorrelation time $\tau_{\rm c}$ of the fluctuating force, the velocity is a Markov process. The density $f(t,v)$ thus obeys the usual consistency equation involving the transition probability, which is recast in Sec. V.2.2 in the form of a partial differential equation of first order in t, yet involving an infinite number of successive derivatives with respect to v. Truncating this equation at second order yields the Fokker–Planck equation (Sec. V.2.3), whose solutions we examine in Sec. V.2.4. Eventually, we repeat the same analysis in the case of the position of the Brownian particle and its probability density (Sec. V.2.5).

V.2.1 Velocity of a Brownian particle as a Markov process

Assume first that the spectral density of the Langevin force is a white noise, i.e. that its autocorrelation function is proportional to a Dirac distribution, Eq. (V.3d), or equivalently, that the autocorrelation time $\tau_{\rm c}$ vanishes. In that case, we have seen [Eq. (V.26)] that the velocity at a given instant t and the fluctuating force at a later time t' are uncorrelated, $\langle v(t)F_{\rm L}(t')\rangle = 0$ for $t' > t$. That is, the Langevin force at time t' has no memory of the past of t'.

Now, if the Langevin force is a Gaussian stochastic process, then so is the velocity of the Brownian particle. The covariance $\langle v(t)F_{\rm L}(t')\rangle = 0$ for $t' > t$ then means that $v(t)$ and $F_{\rm L}(t')$ are statistically independent for $t' > t$.

If $F_{\rm L}(t)$ is a Gaussian process, then its Fourier transform $\tilde{F}_{\rm L}(\omega)$ is a Gaussian random variable. In turn, Eq. (V.30) shows that $\tilde{v}(\omega)$ is also Gaussian—the proportionality factor $1/[M(\gamma-\mathrm{i}\omega)]$ is a "deterministic" function of ω. After a last inverse Fourier transform, $v(t)$ is a Gaussian random process, entirely determined by its first two moments.

Since the Langevin equation (V.1) is of first order, with the source $F_{\rm L}(t)$, the velocity shift between t and $t+\Delta t$ only depends on the velocity at time t and the force in the interval $[t, t+\Delta t]$, yet is totally independent of v and $F_{\rm L}$ at times prior to t, so that $v(t)$ is a Markov process.

If on the other hand $F_{\rm L}(t)$ and thus $v(t)$ is not Gaussian, or if $\tau_{\rm c}$ is finite, then the velocity is strictly speaking no longer a Markov process. Restricting oneself to the change on time scales much larger than $\tau_{\rm c}$—and assuming from now on that $F_{\rm L}(t)$ and $v(t)$ are Gaussian—, $v(t)$ can be approximated as Markovian. That is, we shall in the remainder of this Chapter consider the evolution of the Brownian particle velocity on a coarse-grained version of time, and "infinitesimal" time steps Δt will actually always be much larger than $\tau_{\rm c}$, although we shall consider the formal limit $\Delta t \to 0$.

Remark: From the physical point of view, the coarse-graining of time actually corresponds to the experimental case, in which observations are not performed continuously—in the mathematical sense—, but rather at successive instants, between which the Brownian particle has actually undergone many collisions with its environment.

Since the velocity $v(t)$ of the Brownian particle is assumed to be a Markov process, it is entirely described by its probability density, which will be denoted by $f(t,v)$ instead of the notation $p_1(t,v)$ used in Appendix C.2.5, and by the transition probability $p_{1|1}(t_2,v_2\,|\,t_1,v_1)$. These obey the consistency condition (C.24), which for the evolution between times t and $t+\Delta t$ reads

$$f(t+\Delta t, v) = \int p_{1|1}(t+\Delta t, v\,|\,t, v')\, f(t,v')\,\mathrm{d}v', \tag{V.41a}$$

where $\Delta t \gg \tau_{\rm c}$.

Physically, the collisions with the much lighter constituents of the environment lead on short time scales—i.e. for Δt much smaller than the relaxation time $\tau_{\rm r} = \gamma^{-1}$—only to small shifts of the velocity v of the Brownian particle. That is, the modulus of $w = v - v'$ is much smaller than v. In order to later exploit this property, let us rewrite Eq. (V.41a) as

$$f(t+\Delta t, v) = \int p_{1|1}(t+\Delta t, v \,|\, t, v-w)\, f(t, v-w)\, {\rm d}w, \tag{V.41b}$$

where we now integrate over the change in velocity.

V.2.2 Kramers–Moyal expansion

We shall now assume that the transition probability $p_{1|1}(t+\Delta t, v \,|\, t, v')$ and the probability density $f(t, v')$ are continuous functions of t and Δt, and that their product is analytic in the velocity variables, which will allow us to derive a partial differential equation obeyed by f.
Note that the calculations in this subsection hold more generally for any Markovian stochastic process with the necessary regularity properties; the specific case of the velocity in the Langevin model will be studied in further detail in the next subsection.

Under the above assumptions, the integrand in the evolution equation (V.41b) can be expanded in Taylor series as

$$\begin{aligned} p_{1|1}(t+\Delta t, v \,|\, t, v-w)\, f(t, v-w) &= p_{1|1}(t+\Delta t, v+w-w \,|\, t, v-w)\, f(t, v-w) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, w^n\, \frac{{\rm d}^n}{{\rm d}v^n} \big[p_{1|1}(t+\Delta t, v+w \,|\, t, v)\, f(t,v)\big]. \end{aligned}$$

Introducing for $n \in \mathbb{N}$ the *jump moments*

$$\mathscr{M}_n(t, t+\Delta t, v) \equiv \int w^n\, p_{1|1}(t+\Delta t, v+w \,|\, t, v)\, {\rm d}w = \int (v'-v)^n\, p_{1|1}(t+\Delta t, v' \,|\, t, v)\, {\rm d}v', \tag{V.42}$$

and exchanging the order of integration over w and partial differentiation with respect to v, the evolution equation (V.41) can be rewritten as

$$f(t+\Delta t, v) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial v^n} \big[\mathscr{M}_n(t, t+\Delta t, v)\, f(t,v)\big]. \tag{V.43}$$

Definition (V.42) shows that $\mathscr{M}_0(t, t+\Delta t, v) = 1$ for arbitrary t and Δt—which actually only states that the integral over all possible final states of the transition probability of a Markov process is 1.

For $n \geq 1$, the "initial condition" $p_{1|1}(t, v' \,|\, t, v) = \delta(v'-v)$ and the assumed continuity in Δt mean that $\mathscr{M}_n(t, t+\Delta t, v)$ tends to 0 in the limit $\Delta t \to 0$. Assume now—this will be shown explicitly in the next subsection in the cases $n = 1$ and 2 for the jump moments of the velocity of a Brownian particle—that the jump moments with $n \geq 1$ are to leading order linear in Δt at small Δt:

$$\mathscr{M}_n(t, t+\Delta t, v) \underset{\Delta t \to 0}{\sim} \mathcal{M}_n(t, v)\, \Delta t + o(\Delta t), \tag{V.44}$$

where $o(\Delta t)/\Delta t$ tends towards 0 when $\Delta t \to 0$. Subtracting then from both sides of Eq. (V.43) the term with $n = 0$, dividing by Δt, and finally taking the formal limit $\Delta t \to 0$ leads to(64)

$$\boxed{\frac{\partial f(t,v)}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial v^n} \big[\mathcal{M}_n(t,v)\, f(t,v)\big].} \tag{V.45}$$

---

(64) As in the study of the Boltzmann kinetic equation (Chapter IV), we take the mathematical limit of infinitesimal small Δt, notwithstanding the fact that physically it should be larger than $\tau_{\rm c}$.

This equation is the so-called *Kramers*[(bg)]*–Moyal*[(bh)] *expansion*, which may be written for any Markovian stochastic process fulfilling the regularity hypotheses we have made.

In many situations, the first two jump moments yield a suitable description, and one truncates the expansion at second order, neglecting the terms with $n \geq 3$. This approximation yields the *Fokker*[(bi)]*–Planck*[(bj)] *equation*

$$\boxed{\frac{\partial f(t,v)}{\partial t} = -\frac{\partial}{\partial v}\big[\mathcal{M}_1(t,v)\,f(t,v)\big] + \frac{1}{2}\frac{\partial^2}{\partial v^2}\big[\mathcal{M}_2(t,v)\,f(t,v)\big].} \tag{V.46}$$

The first resp. second term on the right hand side is referred to as *drift* resp. *diffusive* term, and accordingly $\mathcal{M}_1(t,v)$ resp. $\mathcal{M}_2(t,v)$ as drift resp. diffusion coefficient.

Remarks:

∗ To give an interpretation of the jump moments, let us introduce the notation

$$\big\langle g\big(v(t)\big)\,\big|\,v(t_0)\!=\!v_0\big\rangle_v \equiv \int g(v)\,p_{1|1}(t,v\,|\,t_0,v_0)\,\mathrm{d}v,$$

which denotes the average value at time t of the function $g(v)$ of the stochastic process $v(t)$, under the condition that at some earlier instant t_0 the latter takes the value v_0. Comparing with definition (V.42), the jump moment can be rewritten as

$$\mathscr{M}_n(t,t+\Delta t,v) = \Big\langle \big[v(t+\Delta t) - v\big]^n\,\Big|\,v(t)\!=\!v\Big\rangle_v\,. \tag{V.47}$$

That is, $\mathscr{M}_n(t,t+\Delta t,v)$ represents the n-th moment of the probability distribution for the change in velocity between t and $t+\Delta t$, starting from velocity v at time t.

Hereafter, we shall use the fact that such moments can actually be computed in two equivalent ways: either, as in the above two equations, by using the conditional probability $p_{1|1}(t+\Delta t,v'\,|\,t,v)$ and integrating over v'; or by following explicitly trajectories in velocity space that start with the fixed velocity v at time t, and computing the average velocity at a later time as in Sec. V.1.2, from which the average velocity shift easily follows.

∗ If the Markov process under consideration is stationary, the jump moments are independent of time. As we shall see below, the reciprocal does not hold.

∗ The Kramers–Moyal expansion (V.45) is sometimes referred to as *generalized Fokker–Planck equation.*

V.2.3 Fokker–Planck equation for the Langevin model

We now apply the formalism developed in the previous subsection to the specific case of the Langevin model.

V.2.3 a Jump moments for the Langevin model

Let us compute the first two jump moments of the velocity in the Langevin model. Integrating the Langevin equation (V.1) between t and $t+\Delta t$, one finds

$$v(t+\Delta t) = v(t) - \gamma\int_t^{t+\Delta t} v(t')\,\mathrm{d}t' + \frac{1}{M}\int_t^{t+\Delta t} F_{\rm L}(t')\,\mathrm{d}t'. \tag{V.48}$$

Considering now that $v(t)$ is fixed and equal to v, and subtracting it from both sides of the equations, one obtains the velocity change between t and $t+\Delta t$ for a given realization of the Langevin force. Averaging over the possible realizations of the latter, one finds the average velocity shift between t

---

(bg) H. Kramers, 1894–1952 (bh) J. E. Moyal, 1910–1998 (bi) A. Fokker, 1887–1972 (bj) M. Planck, 1858–1947 and $t+\Delta t$ under the condition that $v(t)=v$, i.e. according to Eq. (V.47) precisely the first jump moment

$$\mathscr{M}_1(t,t+\Delta t,v) = -\gamma\int_t^{t+\Delta t} \big\langle v(t')\,\big|\,v(t)\!=\!v\big\rangle\,\mathrm{d}t' + \frac{1}{M}\int_t^{t+\Delta t} \big\langle F_{\mathrm{L}}(t')\,\big|\,v(t)\!=\!v\big\rangle\,\mathrm{d}t',$$

where the fact that the averages over realizations of the Langevin force are conditional ones has explicitly been specified. Thanks to the absence of correlation between $F_{\mathrm{L}}(t')$ and $v(t)$ when $t'>t$, see Eq. (V.26), the condition on $v(t)$ actually plays no role in the expectation value of the Langevin force, which vanishes. In turn, a Taylor expansion of the integrand of the first integral yields

$$\mathscr{M}_1(t,t+\Delta t,v) = -\gamma v\Delta t + \mathcal{O}\big((\gamma\Delta t)^2\big). \tag{V.49a}$$

For time steps $\Delta t \ll \tau_{\mathrm{r}}$, the term of order $(\gamma\Delta t)^2$ is much smaller than the linear term and we may write

$$\mathscr{M}_1(t,t+\Delta t,v) \underset{\Delta t\ll\tau_{\mathrm{r}}}{\simeq} \mathcal{M}_1(t,v)\Delta t + o(\gamma\Delta t) \quad \text{with} \quad \mathcal{M}_1(t,v)\equiv -\gamma v, \tag{V.49b}$$

so that Eq. (V.44) holds here.

Equations (V.47) and (V.48) also give the higher jump moments, in particular the second one, which follows from

$$\big[v(t+\Delta t)-v(t)\big]^2 = \gamma^2\bigg[\int_t^{t+\Delta t} v(t')\,\mathrm{d}t'\bigg]^2 - \frac{2\gamma}{M}\int_t^{t+\Delta t}\!\int_t^{t+\Delta t} v(t')F_{\mathrm{L}}(t'')\,\mathrm{d}t'\,\mathrm{d}t'' + \frac{1}{M^2}\int_t^{t+\Delta t}\!\int_t^{t+\Delta t} F_{\mathrm{L}}(t')F_{\mathrm{L}}(t'')\,\mathrm{d}t'\,\mathrm{d}t''.$$

Fixing the initial value $v(t)$ to v and averaging over an ensemble of realizations of the environment amounts to performing the conditional averaging with $p_{1|1}(\,\cdot\,|\,t,v)$. In that average, the first term on the right-hand side is of order $(\gamma\Delta t)^2$. Since $\Delta t \gg \tau_{\mathrm{c}}$, we can use approximation (V.26) for the integrand of the second term, which again leads to a quadratic term in $\gamma\Delta t$. Eventually, the integrand of the third term can be approximated by $2D_vM^2\delta(t''-t')$ [Eq. (V.3d)], which gives

$$\mathscr{M}_2(t,t+\Delta t,v) = 2D_v\Delta t + \mathcal{O}\big((\Delta t)^2\big), \tag{V.50a}$$

that is, a second jump moment

$$\mathscr{M}_2(t,t+\Delta t,v) \underset{\Delta t\ll\tau_{\mathrm{r}}}{\simeq} \mathcal{M}_2(t,v)\Delta t + o(\Delta t) \quad \text{with} \quad \mathcal{M}_2(t,v)\equiv 2D_v. \tag{V.50b}$$

Here again Eq. (V.44) holds.

V.2.3 b Fokker–Planck equation

Inserting the jump moments (V.49b) and (V.50b) in the general relation (V.46), one obtains the Fokker–Planck equation for the Langevin model

$$\boxed{\frac{\partial f(t,v)}{\partial t} = \gamma\frac{\partial}{\partial v}\big[vf(t,v)\big] + D_v\frac{\partial^2 f(t,v)}{\partial v^2}.} \tag{V.51}$$

We thus recover the interpretation of D_v as a diffusion coefficient in velocity space.

Remarks:

∗ Interestingly, the jump moments $\mathscr{M}_1$, $\mathscr{M}_2$ for the velocity of the Langevin model are not explicitly time-dependent but only depend on Δt, even though the velocity is not a stationary process as long as equilibrium has not been reached.

$*$ If the Langevin force is a Gaussian stochastic process, so is the velocity, and the transition probability $p_{1|1}(t+\Delta t, v' \,|\, t, v)$ is also Gaussian.(65) The transition probability is thus entirely determined by its first two moments, which are precisely the jump moments $\mathscr{M}_1$, $\mathscr{M}_2$, and we may write

$$p_{1|1}(t+\Delta t, v' \,|\, t, v) = \frac{1}{\sqrt{4\pi D_v \Delta t}} \exp\left\{-\frac{[v'-(1-\gamma\Delta t)v]^2}{4D_v\Delta t}\right\} \quad \text{for } \tau_{\rm c} \ll \Delta t \ll \tau_{\rm r}, \tag{V.52}$$

where we have used the fact that $\mathscr{M}_2$ is also the variance, since it is much larger than $\mathscr{M}_1^2$.

$*$ If the Langevin force and the velocity are not Gaussian processes, then one may still argue that the transition probability, as function of the velocity shift $v'-v$ at fixed v, is given by a Gaussian distribution when $\Delta t \gg \tau_{\rm c}$. In such a time interval, many statistically independent collisions between the Brownian particle and its environment take place, which lead to as many statistically independent tiny velocity shifts: according to the central limit theorem, the resulting total velocity shift over Δt, which is the sum of these tiny shifts, is Gaussian distributed.

V.2.4 Solution of the Fokker–Planck equation

The Fokker–Planck equation (V.51) is a linear partial differential equation with non-constant coefficients relating the time derivative of the velocity density to its first two "spatial" derivatives—or, equivalently, an equation with constant coefficients involving time derivative, the first two spatial derivatives, and the function itself. Accordingly, it has the form of a generalized diffusion equation in velocity space, with a diffusion coefficient D_v—we recover the interpretation of that coefficient found in Sec. V.1.2—, and a "drift term" $\gamma\partial[vf(t,v)]/\partial v$—so that γ is referred to as *drift coefficient*.

Defining a probability current (in velocity space) as

$$\mathscr{J}_v(t,v) \equiv -\gamma v f(t,v) - D_v \frac{\partial f(t,v)}{\partial v}, \tag{V.53a}$$

the Fokker–Planck equation can be recast in the form of a continuity equation

$$\frac{\partial f(t,v)}{\partial t} + \frac{\partial \mathscr{J}_v(t,v)}{\partial v} = 0 \tag{V.53b}$$

for the probability density.

V.2.4 a Stationary solution

One can first investigate the stationary (or steady-state) solutions $f_{\rm st.}(v)$ to the Fokker–Planck equation. According to Eq. (V.53b), these solutions make the probability current (V.53a) constant. To be normalizable, a solution $f_{\rm st.}(v)$ should decrease faster than $1/|v|$ when $|v|$ tends to ∞. The only possibility is when $\mathscr{J}_v(t,v) = 0$.(66) The corresponding stationary solution is then simply

$$f_{\rm st.}(v) = \sqrt{\frac{\gamma}{2\pi D_v}}\, {\rm e}^{-\gamma v^2/2D_v}. \tag{V.54}$$

If the environment of the Brownian particle is in thermal equilibrium at temperature T, then the fluctuation–dissipation relation $D_v/\gamma = k_BT/M$ [Eq. (V.12)] shows that the steady-state solution to the Fokker–Planck equation is the Maxwell–Boltzmann distribution. The Brownian particle is thus "thermalized".

---

(65) According to Bayes' theorem (C.13), it equals the ratio of two Gaussian distributions.

(66) For a generic stochastic process $Y(t)$, whose realizations take their values in a bounded real interval $[a,b]$, the existence and number of stationary solutions of the corresponding Fokker–Planck equation (V.46) depend on the choice of boundary conditions imposed at a and b: vanishing $\mathscr{J}_Y \equiv \mathcal{M}_1 p_{Y,1} - \frac{1}{2}\partial(\mathcal{M}_2 p_{Y,1})/\partial y$ for $y=a$ and $y=b$—i.e. so-called *reflecting boundary conditions*—, vanishing $p_{Y,1}(y)$... The stationary solutions also depend on the dimension of the stochastic process—in two or more dimensions, non-vanishing probability currents exist without involving a flow of probability towards infinitely large values.

Remark: If the drift coefficient γ were negative, the Fokker–Planck equation (V.51) would have no stationary solution.

V.2.4 b Fundamental solution

As next step, one can search for the fundamental solutions—also called Green's functions—of the Fokker–Planck equation, namely the solutions to equation (V.51) obeying the initial condition $f(0,v) = \delta(v-v_0)$ for an arbitrary $v_0 \in \mathbb{R}$.

One can show that this fundamental solution is given by

$$\boxed{f(t,v) = \sqrt{\frac{\gamma}{2\pi D_v(1-\mathrm{e}^{-2\gamma t})}}\,\exp\!\left[-\frac{\gamma}{2D_v}\frac{(v-v_0\,\mathrm{e}^{-\gamma t})^2}{1-\mathrm{e}^{-2\gamma t}}\right] \quad \text{for } t>0.} \tag{V.55}$$

- At a given instant $t>0$, this distribution is Gaussian, with average value and variance

  $$\langle v(t)\rangle = v_0\,\mathrm{e}^{-\gamma t}, \qquad \sigma_v(t)^2 = \frac{D_v}{\gamma}\big(1-\mathrm{e}^{-2\gamma t}\big),$$

  in agreement with expressions (V.5) and (V.8), with $t_0 = 0$, found in the case $t - t_0 \gg \tau_c$.

- When t becomes much larger than τ_r, the fundamental solution (V.55) tends to the stationary solution (V.54).

In agreement with the consistency condition (C.24), the fundamental solution (V.55) equals the transition probability $p_{1|1}(t,v\,|\,t_0=0,v_0)$. At small $t \ll \tau_r$—which is then rewritten as Δt—, one actually recovers Eq. (V.52). More generally, one recognizes in $p_{1|1}(t,v\,|\,0,v_0)$ given by Eq. (V.55) the transition probability (C.31b) of the Ornstein–Uhlenbeck process [45].

Remark: The Ornstein–Uhlenbeck process is actually defined by both its transition probability and its time-independent single-time probability. The latter condition is not fulfilled by the velocity of a Brownian particle in general—the velocity is not a stationary process—, yet is obeyed in the equilibrium regime $t \gg \tau_r$, in which case the terms $\mathrm{e}^{-\gamma t}$ and $\mathrm{e}^{-2\gamma t}$ in Eq. (V.55) vanish.

V.2.5 Position of a Brownian particle as a Markov process

We may now repeat the study of the previous subsections for the case of the position $x(t)$ of a free Brownian particle.

The first important point is that the evolution equation for the position deduced from the Langevin equation (V.1) is of second order. As a consequence, the displacement $x(t+\Delta t)-x(t)$ in a small time step depends not only on $x(t)$ and $F_L(t')$ for $t' \geq t$ only, but also on the velocity $v(t)$, which in turn depends on the past of t. That is, the position is in general not a Markov process, even if the Langevin force is a Gaussian process with a vanishing autocorrelation time.

To recover the Markovian character, one has to consider time steps $\Delta t \gg \tau_r$, i.e. a coarser graining than for velocity. Over such a time interval, the velocity of the Brownian undergoes many random changes, and $x(t+\Delta t)-x(t)$ will be independent of the position at previous times.

On such time scales, the acceleration term in the Langevin equation plays no role, and one recovers the first-order equation (V.20a) valid in the viscous limit (cf. § V.1.3 c)

$$\eta_v \frac{\mathrm{d}x(t)}{\mathrm{d}t} = F_L(t).$$

Additionally, the condition $\Delta t \gg \tau_r$ automatically leads to $\Delta t \gg \tau_c$, so that the autocorrelation time of the Langevin force can be neglected:

$$\langle F_L(t)F_L(t+\tau)\rangle = 2D\eta_v^2\,\delta(\tau)$$

[cf. Eq. (V.20b)]. In analogy to the finding in Sec. V.1.4, this leads to $\langle x(t)\, F_{\rm L}(t')\rangle = 0$ for $t' > t$, which if $F_{\rm L}(t)$, and thus $x(t)$, is Gaussian, guarantees their statistical independence.

Repeating the derivation of Sec. V.2.3 with $v(t)$, M, γ and $D_v M^2$ respectively replaced by $x(t)$, η_v, 0 and $D\eta_v^2$, one finds that the jumps moments for the position are $\mathcal{M}_1(t,x) = 0$ and $\mathcal{M}_2(t,x) = 2D$. The Fokker–Planck equation for the evolution of the probability density $f(t,x)$ of the position thus reads

$$\boxed{\frac{\partial f(t,x)}{\partial t} = D\frac{\partial^2 f(t,x)}{\partial x^2}.} \tag{V.56}$$

That is, $f(t,x)$ obeys the ordinary diffusion equation, with the fundamental solution corresponding to the initial condition $f(0,x) = \delta(x - x_0)$ given by

$$f(t,x) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{(x-x_0)^2}{4Dt}\right] \quad \text{for } t > 0. \tag{V.57}$$

From this probability density, one recovers the large-time limit of the variance found in Eq. (V.22).

Again, the fundamental solution (V.57) also equals the transition probability of the Markov process $x(t)$. Together with the initial condition $f(t\!=\!0, x) = \delta(x - x_0)$, they exactly match the definition of the Wiener process, Eq. (C.26).

V.2.6 Diffusion in phase space

Let us again assume as in Sec. V.1.6 that the Brownian particle is subject to a deterministic external force $\vec{F}_{\rm ext.}$ besides the fluctuating force. The corresponding Langevin equation (V.37) and the normal relation between velocity and position can be recast as

$$\frac{\rm d}{{\rm d}t}\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ -\gamma v(t) + \dfrac{1}{M}F_{\rm ext.} \end{pmatrix} + \begin{pmatrix} 0 \\ \dfrac{1}{M}F_{\rm L}(t) \end{pmatrix}. \tag{V.58}$$

This system may be viewed as a stochastic first-order differential equation for the two-dimensional process $\mathbf{Y}(t) \equiv \big(x(t), v(t)\big)$, with a “drift term” resp. “noise term” given by the first resp. second term on the right hand side. Here the noise term affecting the first component $x(t)$ is actually vanishing: this obviously involves that the integral of the corresponding autocorrelation function—which may in analogy with Eqs. (V.3b)–(V.3c) be denoted by $2D_x$—is vanishing, i.e. $D_x = 0$. There is also no friction term proportional to $x(t)$ in the first equation, only a “forcing term” $v(t)$.

If the external force is independent of velocity, one easily sees by repeating the same reasoning as in Sec. V.2.1 that $\mathbf{Y}(t)$ is Markovian, since the terms on the right hand side of the equation only depend on the physical quantities at time t, not before. Repeating the same steps as in Sec. V.2.3, one can derive the corresponding Fokker–Planck equation,(67) which reads

$$\boxed{\frac{\partial f(t,x,v)}{\partial t} = -v\frac{\partial f(t,x,v)}{\partial x} + \gamma\frac{\partial}{\partial v}\big[v f(t,x,v)\big] - \frac{F_{\rm ext.}}{M}\frac{\partial f(t,x,v)}{\partial v} + D_v\frac{\partial^2 f(t,x,v)}{\partial v^2},} \tag{V.59}$$

where there is no second derivative with respect to x since the corresponding diffusion coefficient D_x vanishes. This constitutes the so-called *Klein*(bk)*–Kramers* or *Kramers–Chandrasekhar*(bl) *equation*.

Remark: The careful reader may complain that the position $x(t)$ was argued in Sec. V.2.5 to be a non-Markovian stochastic process on time scales comparable to or smaller than γ^{-1}, while this

---

(67) See the appendix V.A to this Chapter for a short derivation of the multivariate Fokker–Planck equation.

(bk) O. Klein, 1894–1977 (bl) S. Chandrasekhar, 1910–1995 seems to be ignored here. The reason is that the new stochastic process $\mathbf{Y}(t)$ involves not only $x(t)$, but also $v(t)$, which was the obstacle to the Markov property in the previous section. Introducing explicitly $v(t)$ into $\mathbf{Y}(t)$ amounts to include the information on the past of $x(t)$ in the new process. As a result, we now deal with a first-order equation, instead of a second-order one.

Invoking relation (V.12) between the coefficients D_v and γ,(68) and performing some trivial rewriting, it becomes

$$\frac{\partial f(t,x,v)}{\partial t}+v\frac{\partial f(t,x,v)}{\partial x}+\frac{F_{\rm ext.}}{M}\frac{\partial f(t,x,v)}{\partial v}=\gamma\frac{\partial}{\partial v}\left[vf(t,x,v)+\frac{k_BT}{M}\frac{\partial f(t,x,v)}{\partial v}\right]. \tag{V.60}$$

In this form, one recognizes on the left hand side that of the first equation (III.14a) of the BBGKY hierarchy—here in position-velocity space. The Klein–Kramers–Chandrasekhar equation thus represents a possible truncation of that hierarchy.

V.3 Generalized Langevin dynamics

In Sec. V.1, we have seen that the Langevin equation (V.1), with an instantaneous friction term $-M\gamma v(t)$, leads at large times, when the Brownian particle is in equilibrium with the surrounding fluid, to the velocity [cf. Eq. (V.25)]

$$v(t)=\frac{1}{M}\int_{-\infty}^{t}F_{\rm L}(t')\,{\rm e}^{-\gamma(t-t')}\,{\rm d}t'=\int_{-\infty}^{\infty}\chi(t-t')\,F_{\rm L}(t')\,{\rm d}t'\quad\text{with}\quad\chi(\tau)\equiv\frac{1}{M}\,{\rm e}^{-\gamma\tau}\Theta(\tau),$$

where $\Theta(\tau)$ denotes the Heaviside function. $\chi(\tau)$ is the *response function*, whose Fourier transform is precisely the admittance $Y(\omega)$, Eq. (V.39b).

The velocity thus responds instantaneously to the Langevin force and is independent of its own past values, which is physically unrealistic. A more realistic model consists in introducing a time-non-local friction term, which leads to a generalized Langevin equation (Sec. V.3.1), whose spectral properties we analyse in Sec. V.3.2. This generalized form of the Langevin equation is actually that which naturally emerges from a more microscopic description, in particular for the dynamics of a free classical particle interacting with an infinite bath of degrees of freedom in thermal equilibrium (Sec. V.3.3).

V.3.1 Generalized Langevin equation

To account for retardation effects in the motion of the Brownian particle, one can replace Eq. (V.1) by the linear integro-differential equation

$$\boxed{M\frac{{\rm d}v(t)}{{\rm d}t}=-M\int_{-\infty}^{t}\gamma(t-t')\,v(t')\,{\rm d}t'+F_{\rm L}(t)\quad\text{with}\quad v(t)=\frac{{\rm d}x(t)}{{\rm d}t},} \tag{V.61a}$$

called *generalized Langevin equation*, with a memory kernel $\gamma(\tau)$ for the friction force, where τ denotes the retardation $t-t'$. While only values for $\tau\geq 0$ play a role in this equation, yet it is convenient to view γ as an even function of $\tau\in\mathbb{R}$, whose integral over $\mathbb{R}$ equals a number denoted as 2γ, and to recast Eq. (V.61a) as

$$M\frac{{\rm d}v(t)}{{\rm d}t}=-M\int_{-\infty}^{\infty}\gamma(t-t')\,\Theta(t-t')v(t')\,{\rm d}t'+F_{\rm L}(t). \tag{V.61b}$$

As before, the Langevin force $F_{\rm L}(t)$ is a stationary stochastic process with vanishing average. Since retardation effects are taken into account in the evolution of the velocity, they should for

---

(68) This relation also holds between the corresponding coefficients in the first equation of system (V.58), since both are zero—there is no noise term, and $x(t)$ does not appear on the right hand side.

consistency also be present in the fluctuating force, whose autocorrelation time $\tau_{\rm c}$ will thus be non-zero.

V.3.2 Spectral analysis

Since the initial time in the friction term is at $t_0 \to -\infty$, the Brownian particle is at every instant in equilibrium with its environment. Its velocity is thus a stationary stochastic process, to which one can apply the concepts of Appendix C.3. We consider the case of an environment in thermal equilibrium at temperature T.

Generalizing the analysis of Sec. V.1.6, one easily finds that the stationary response to an external force $F_{\rm ext.}(t)$ independent of position and velocity reads in Fourier space

$$\langle\tilde{v}(\omega)\rangle = Y(\omega)\,\tilde{F}_{\rm ext.}(\omega), \tag{V.62a}$$

with

$$Y(\omega) \equiv \frac{1}{M}\frac{1}{\tilde{\gamma}(\omega)-{\rm i}\omega} \tag{V.62b}$$

the complex admittance, where $\tilde{\gamma}(\omega)$ is given by

$$\tilde{\gamma}(\omega) = \int_0^\infty \gamma(t)\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t = \int \gamma(t)\,\Theta(t)\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t. \tag{V.62c}$$

Note that $\tilde{\gamma}(\omega\!=\!0) = \gamma$.

Fourier-transforming the generalized Langevin equation (V.61a), one finds that the spectral densities of the velocity and the fluctuating force are related by [cf. Eq. (V.32)]

$$S_v(\omega) = \frac{1}{M^2}\frac{1}{\big|\tilde{\gamma}(\omega)-{\rm i}\omega\big|^2}\,S_F(\omega). \tag{V.63}$$

Assuming(69) that relation (V.40) between the admittance and the autocorrelation of velocity still holds here, one finds, since the velocity is real-valued, the identity

$${\rm Re}\,Y(\omega) = \frac{1}{2k_BT}\int_{-\infty}^{\infty}\langle v(t)\,v(0)\rangle\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t. \tag{V.64}$$

Using Eq. (V.62b), this yields

$$\int\langle v(t)\,v(0)\rangle\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t = 2k_BT\,{\rm Re}\,Y(\omega) = \frac{2k_BT}{M}\frac{{\rm Re}\,\tilde{\gamma}(\omega)}{\big|\tilde{\gamma}(\omega)-{\rm i}\omega\big|^2}.$$

Now, invoking the Wiener–Khinchin theorem, the Fourier transform of the autocorrelation function is exactly the spectral density, i.e.

$$S_v(\omega) = \frac{2k_BT}{M}\frac{{\rm Re}\,\tilde{\gamma}(\omega)}{\big|\tilde{\gamma}(\omega)-{\rm i}\omega\big|^2}. \tag{V.65}$$

Comparing with relation (V.63), there comes

$$S_F(\omega) = 2Mk_BT\,{\rm Re}\,\tilde{\gamma}(\omega). \tag{V.66}$$

From the Wiener–Khinchin theorem, the spectral density of the Langevin force fluctuations is the Fourier transform of the corresponding autocorrelation function, which gives

$$\kappa(\tau) = Mk_BT\gamma(\tau), \tag{V.67a}$$

---

(69) This will be demonstrated in Sec.VI.3.4 in next Chapter.

or equivalently

$$\tilde{\gamma}(\omega) = \frac{1}{M k_B T} \int_0^\infty \langle F_{\rm L}(t)\, F_{\rm L}(t+\tau)\rangle\, {\rm e}^{{\rm i}\omega\tau}\, {\rm d}\tau. \tag{V.67b}$$

Equation (V.67a) shows that the memory kernel for the friction term has the same characteristic time scale τ_c as the fluctuations of the Langevin force—hence the necessity of considering a finite τ_c of the Langevin noise when allowing for retardation effects in the friction term.

V.3.3 Caldeira–Leggett model

In this subsection, we introduce a simple microscopical model for the fluid in which the Brownian particle is immersed, which leads to a friction force proportional to the velocity.

V.3.3 a Caldeira–Leggett Hamiltonian

Consider a "Brownian" particle of mass M, with position and momentum $x(t)$ and $p(t)$ respectively, interacting with a "bath" of N mutually independent harmonic oscillators with respective masses m_j, positions $x_j(t)$ and momenta $p_j(t)$. The coupling between the particle and each of the oscillators is assumed to be bilinear in their positions, with a coupling strength C_j. Additionally, we also allow for the Brownian particle to be in a position-dependent potential $V_0(x)$.

Under these assumptions, the Hamilton function of the system consisting of the Brownian particle and the oscillators reads

$$H = \frac{p^2}{2M} + V_0(x) + \sum_{j=1}^{N}\left(\frac{p_j^2}{2m_j} + \frac{1}{2}m_j\omega_j^2 x_j^2\right) - \sum_{j=1}^{N} C_j\, x_j\, x. \tag{V.68a}$$

It is convenient to rewrite the potential V_0 as

$$V_0(x) = V(x) + \left(\sum_{j=1}^{N}\frac{C_j^2}{2m_j\omega_j^2}\right)x^2,$$

where the second term in the right member clearly vanishes when the Brownian particle does not couple to the oscillators. The Hamilton function (V.68a) can then be recast as

$$\boxed{H = \frac{p^2}{2M} + V(x) + \sum_{j=1}^{N}\left[\frac{p_j^2}{2m_j} + \frac{1}{2}m_j\omega_j^2\left(x_j - \frac{C_j}{m_j\omega_j^2}\,x\right)^2\right].} \tag{V.68b}$$

This Hamilton function—or its quantum-mechanical counterpart, which we shall meet again in Sec. VI.4.3 d—is known as the *Caldeira*[bm]*–Leggett*[bn] *Hamiltonian* [46].

For a physical interpretation, it is interesting to rescale the characteristics of the bath oscillators, performing the change of variables

$$m_j \to m_j' = \frac{m_j}{\lambda_j^2}, \qquad x_j \to x_j' = \lambda_j x_j, \qquad p_j \to p_j' = \frac{p_j}{\lambda_j}, \qquad j = 1,\dots,N$$

with $\lambda_j \equiv \dfrac{m_j\omega_j^2}{C_j}$ dimensionless constants. The Hamiltonian (V.68b) then becomes

$$H = \frac{p^2}{2M} + V(x) + \sum_{j=1}^{N}\left[\frac{p_j'^2}{2m_j'} + \frac{1}{2}m_j'\omega_j^2\left(x_j' - x\right)^2\right], \tag{V.69}$$

i.e. the interaction term between the Brownian particle and each oscillator only depends on their relative distance. The Caldeira–Leggett Hamiltonian can thus be interpreted as that of a particle of mass m, moving in the potential V with (light) particles of masses m_j' attached to it by springs of respective spring constants $m_j'\omega_j^2$ [47].

[bm] A. Caldeira, born 1950 [bn] A. J. Leggett, born 1938

Coming back to the form (V.68b) of the Hamilton function, the corresponding equations of motion, derived from the Hamilton equations (II.1), read

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{1}{M}\,p(t), \qquad \frac{\mathrm{d}p(t)}{\mathrm{d}t} = \sum_{j=1}^{N} C_j\left[x_j(t) - \frac{C_j}{m_j\omega_j^2}\,x(t)\right] - \frac{\mathrm{d}V\big(x(t)\big)}{\mathrm{d}x}, \tag{V.70a}$$

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = \frac{1}{M}\,p_j(t), \qquad \frac{\mathrm{d}p_j(t)}{\mathrm{d}t} = -m\omega_j^2 x_j(t) + C_j x(t), \tag{V.70b}$$

which can naturally be recast as second-order differential equations for the positions $x(t)$, $x_j(t)$.

V.3.3 b Free particle

Let us now assume that the Brownian particle is "free", in the sense that the potential V in the Hamiltonian (V.68b) vanishes, $V(x) = 0$. In that case, the last term on the right-hand side of second equation of motion for the Brownian particle, Eq. (V.70a), vanishes.

Integrating formally the equations of motion (V.70b) for each oscillator between an initial time t_0 and time t, one obtains

$$x_j(t) = x_j(t_0)\cos\omega_j(t-t_0) + \frac{p_j(t_0)}{m_j\omega_j}\sin\omega_j(t-t_0) + C_j\int_{t_0}^{t} x(t')\,\frac{\sin\omega_j(t-t')}{m_j\omega_j}\,\mathrm{d}t'.$$

Performing an integration by parts and rearranging the terms, one finds

$$x_j(t) - \frac{C_j}{m_j\omega_j^2}\,x(t) = \left[x_j(t_0) - \frac{C_j}{m_j\omega_j^2}\,x(t_0)\right]\cos\omega_j(t-t_0) + \frac{p_j(t_0)}{m_j\omega_j}\sin\omega_j(t-t_0) - C_j\int_{t_0}^{t}\frac{p(t')}{M}\,\frac{\cos\omega_j(t-t')}{m_j\omega_j^2}\,\mathrm{d}t'.$$

This can then be inserted in the right member of the second equation of motion in Eq. (V.70a). Combining with the first equation of motion giving $p(t)$ in function of $\mathrm{d}x(t)/\mathrm{d}t$, one obtains

$$\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} + \int_{t_0}^{t}\left[\frac{1}{M}\sum_{j=1}^{N}\frac{C_j^2}{m_j\omega_j^2}\cos\omega_j(t-t')\right]\frac{\mathrm{d}x(t')}{\mathrm{d}t}\,\mathrm{d}t' = \frac{1}{M}\sum_{j=1}^{N} C_j\left[x_j(t_0)\cos\omega_j(t-t_0) + \frac{p_j(t_0)}{m_j\omega_j}\sin\omega_j(t-t_0)\right] - \left[\frac{1}{M}\sum_{j=1}^{N}\frac{C_j^2}{m_j\omega_j^2}\cos\omega_j(t-t_0)\right]x(t_0). \tag{V.71}$$

Introducing the quantities

$$\gamma(t) \equiv \frac{1}{M}\sum_{j=1}^{N}\frac{C_j^2}{m_j\omega_j^2}\cos\omega_j t \tag{V.72a}$$

and

$$F_{\mathrm{L}}(t) \equiv \sum_{j=1}^{N} C_j\left[x_j(t_0)\cos\omega_j(t-t_0) + \frac{p_j(t_0)}{m_j\omega_j}\sin\omega_j(t-t_0)\right], \tag{V.72b}$$

which both only involve characteristics of the bath, Eq. (V.71) becomes

$$M\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} + M\int_{t_0}^{t}\gamma(t-t')\,\frac{\mathrm{d}x(t')}{\mathrm{d}t}\,\mathrm{d}t' = -M\gamma(t-t_0)x(t_0) + F_{\mathrm{L}}(t). \tag{V.72c}$$

This evolution equation for the position—or equivalently the velocity, since $x(t_0)$ in the right-hand side is only a number—of the Brownian particle is exact, and follows from the Hamiltonian equations of motion without any approximation. It is obviously very reminiscent of the generalized Langevin equation (V.61a), up to a few points, namely the lower bound of the integral, the term $-M\gamma(t-t_0)x(t_0)$, and the question whether $F_{\mathrm{L}}(t)$ as defined by relation (V.72b) is a Langevin force.

The first two differences between Eq. (V.61a) and Eq. (V.72c), rewritten in terms of the velocity, are easily dealt with, by sending the arbitrary initial time t_0 to $-\infty$: anticipating on what we shall find below, the memory kernel $\gamma(t)$ vanishes at infinity for the usual choices for the distribution of the bath oscillator frequencies, which suppresses the contribution $-M\gamma(t-t_0)x(t_0)$.

In turn, the characteristics of the force $F_{\rm L}(t)$, Eq. (V.72b), depend on the initial positions and momenta $\{x_j(t_0), p_j(t_0)\}$ of the bath oscillators at time t_0. Strictly speaking, if the latter are exactly known, then $F_{\rm L}(t)$ is a deterministic force, rather than a fluctuating one. When the number N of oscillators becomes large, the deterministic character of the force becomes elusive, since in practice one cannot know the variables $\{x_j(t_0), p_j(t_0)\}$ with infinite accuracy (see the discussion in Sec. II.1). In practice, it is then more fruitful to consider the phase-space density $\rho_N\big(t_0, \{q_j\}, \{p_j\}\big)$, Eq. (II.3). Assuming that at t_0 the bath oscillators are in thermal equilibrium at temperature T, ρ_N at that instant is given by the canonical distribution

$$\rho_N\big(t_0, \{q_j\}, \{p_j\}\big) = \frac{1}{Z_N(T)} \exp\left[-\frac{1}{k_BT}\sum_{j=1}^{N}\left(\frac{p_j^2}{2m_j} + \frac{1}{2}m_j\omega_j^2x_j^2\right)\right].$$

Computing average values with this Gaussian distribution,(70) one finds that the force $F_{\rm L}(t)$ as given by Eq. (V.72b) is a stationary Gaussian random process, with vanishing average value $\langle F_{\rm L}(t)\rangle = 0$ and the autocorrelation function

$$\langle F_{\rm L}(t)F_{\rm L}(t+\tau)\rangle = Mk_BT\gamma(\tau).$$

That is, $F_{\rm L}(t)$ has the properties of a Langevin force as discussed in § V.1.1 b.

Remark: For $V(x) = 0$, the Hamilton function (V.69) is invariant under global translations of all (Brownian and light) particles, since it only depends on relative distances. As a consequence, the corresponding total momentum $p + \sum_j p'_j$ is conserved.

V.3.3 c Limiting case of a continuous bath

As long as the number N of bath oscillators is finite, the Caldeira–Leggett Hamiltonian (V.68b) with $V(x) = 0$ [or with a harmonic potential $V(x) \propto x^2$] strictly speaking leads to a periodic dynamical evolution. As thus, it cannot provide an underlying microscopic model for Brownian motion.(71) The latter can however emerge if one considers the limit of an infinite number of bath degrees of freedom,(72) in particular if the oscillator frequencies span a continuous interval.

To provide an appropriate description for both finite- and infinite-N cases, it is convenient to introduce the *spectral density of the coupling to the bath*

$$J(\omega) \equiv \frac{\pi}{2}\sum_j \frac{C_j^2}{m_j\omega_j}\,\delta(\omega-\omega_j). \tag{V.73}$$

With its help, the memory kernel (V.72a) can be recast as

$$\gamma(t) = \frac{2}{\pi}\int \frac{J(\omega)}{M\omega}\cos\omega t\,{\rm d}\omega. \tag{V.74}$$

---

(70) As noted in the remark at the end of Sec. V.1.1 b, this is indeed the meaning of expectation values in this chapter, since we are averaging over all microscopic configurations $\{x_j(t_0), p_j(t_0)\}$ compatible with a given macroscopic temperature.

(71) ... valid on any time scale. Physically, if $N \gg 1$, the Poincaré(bo) recurrence time of the system will in general be very large. On a time scale much smaller than this recurrence time, the periodicity of the problem can be ignored, and the dynamics is well described by the generalized Langevin model.

(72) The frequencies of the bath oscillators should not stand in simple relation to each other—as for instance if they were all multiples of a single frequency.

(bo) H. Poincaré, 1854–1912

If N is finite, then $J(\omega)$ is a discrete sum of δ-distributions. Let ε denote the typical spacing between two successive frequencies ω_j of the bath oscillators. For evolutions on time scales much smaller than $1/\varepsilon$, the discreteness of the set of frequencies may be ignored.(71) Consider the continuous function $J_c(\omega)$, such that on every interval $\mathcal{I}_\omega \equiv [\omega, \omega + \mathrm{d}\omega]$ of width $\mathrm{d}\omega \gg \varepsilon$, with $\mathrm{d}\omega$ small enough that $J_c(\omega)$ does not vary significantly over $\mathcal{I}_\omega$, one has

$$J_c(\omega)\,\mathrm{d}\omega = \sum_{\omega_j \in \mathcal{I}_\omega} \frac{\pi}{2} \frac{C_j^2}{m_j \omega_j}. \tag{V.75}$$

One can then replace $J(\omega)$ by $J_c(\omega)$, for instance in Eq. (V.74), which amounts to considering a continuous spectrum of bath frequencies.

The simplest possible choice for $J_c(\omega)$ consists in assuming that it is proportional to the frequency for positive values of ω. To be more realistic, one also introduces an upper cutoff frequency ω_c, above which J_c vanishes:

$$J_c(\omega) = \begin{cases} M\gamma\omega & \text{for } 0 \leq \omega \leq \omega_c, \\ 0 & \text{otherwise.} \end{cases} \tag{V.76}$$

This choice leads at once with Eq. (V.74) to $\gamma(t) = 2\gamma\delta_{\omega_c}(t)$, where

$$\delta_{\omega_c}(t) = \frac{1}{\pi} \frac{\sin \omega_c t}{t} \tag{V.77}$$

is a function that tends to $\delta(t)$ as $\omega_c \to +\infty$ and only takes significant values on a range of typical width ω_c^{-1} around $t = 0$. ω_c^{-1} is thus the characteristic time scale of the memory kernel $\gamma(t)$.

Remarks:

$*$ In the limit $\omega_c \to +\infty$, i.e. of an instantaneous memory kernel $\gamma(t) = 2\gamma\delta(t)$, the evolution equation (V.72c) reduces to the Langevin equation [cf. (V.1)] $M\ddot{x}(t) + M\gamma\dot{x}(t) = F_{\mathrm{L}}(t)$. As this is also the equation governing the electric charge in a RL circuit, the choice $J_c(\omega) \propto \omega$ at low frequencies is referred to as "ohmic bath". In turn, a harmonic bath characterized by $J_c(\omega) \propto \omega^\eta$ with $\eta < 1$ (resp. $\eta > 1$) is referred to as sub-ohmic (resp. super-ohmic).(73)

$*$ Instead of a step function $\Theta(\omega_c - \omega)$ as in Eq. (V.76), one may also use a smoother cutoff function to handle the ultraviolet modes in the bath, without affecting the physical results significantly.

Bibliography for Chapter V

- Landau & Lifshitz, *Course of theoretical physics. Vol. 10: Physical kinetics* [5], chapter II § 21.
- Le Bellac, Mortessagne & Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* [9], chapter 9.4–9.5.
- Pottier, *Nonequilibrium statistical physics* [6], chapter 10–11.
- Reif, *Fundamentals of statistical and thermal physics* [35], chapter 15.6–15.12.
- van Kampen, *Stochastic processes in physics and chemistry* [49], chapter VIII.
- Zwanzig, *Nonequilibrium statistical mechanics* [7], chapters 1 & 2.

---

(73) See Ref. [48] for a study in the non-ohmic case.

Appendix to Chapter V

V.A Fokker–Planck equation for a multidimensional Markov process

The results of Secs. V.2.2 are readily extended to the case of a multidimensional Markovian stochastic process $\mathbf{Y}(t) = \big\{Y_1(t), \ldots, Y_D(t)\big\}$ with single-time density $p_{Y,1}(t, \{y_i\}) \equiv f(t, \{y_i\})$ and probability transition $p_{Y,1|1}(t', \{y'_i\} \,|\, t, \{y_i\}) \equiv p_{1|1}(t', \{y'_i\} \,|\, t, \{y_i\})$. In this appendix we state the main results, skipping the detailed derivations. As usual, we indifferently write $\mathbf{y}(t)$ or $\{y_i(t)\}$ for a realization of the stochastic process.

V.A.1 Kramers–Moyal expansion

Generalizing definition (V.42), one defines *jump moments* of order $n \in \mathbb{N}$ as

$$\mathscr{M}^{(n)}_{i_1,\ldots,i_n}(t, t+\Delta t, \mathbf{y}) \equiv \int (y'_{i_1} - y_{i_1}) \cdots (y'_{i_n} - y_{i_n})\, p_{1|1}(t+\Delta t, \mathbf{y}' \,|\, t, \mathbf{y})\, \mathrm{d}y'_1 \cdots \mathrm{d}y'_D, \tag{V.78a}$$

where each i_j is an integer between 1 and D. As in (V.47), this jump moment can be recast as

$$\mathscr{M}^{(n)}_{i_1,\ldots,i_n}(t, t+\Delta t, \mathbf{y}) = \big\langle \big(y_{i_1}(t+\Delta t) - y_{i_1}\big) \cdots \big(y_{i_n}(t+\Delta t) - y_{i_n}\big) \,\big|\, \mathbf{y}(t) = \mathbf{y} \big\rangle_{\mathbf{y}}. \tag{V.78b}$$

Under the assumption that each of these jump moments is at most linear in Δt in the limit $\Delta t \to 0$, i.e.

$$\mathscr{M}^{(n)}_{i_1,\ldots,i_n}(t, t+\Delta t, \mathbf{y}) \underset{\Delta t \to 0}{\sim} \mathcal{M}^{(n)}_{i_1,\ldots,i_n}(t, \mathbf{y})\, \Delta t + o(\Delta t), \tag{V.79}$$

where $o(\Delta t)/\Delta t$ tends towards 0 when $\Delta t \to 0$, then a Taylor expansion similar to that of Sec. V.2.2 leads to the multivariate Kramers–Moyal expansion

$$\boxed{\frac{\partial f(t, \mathbf{y})}{\partial t} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \sum_{i_1,\ldots,i_n} \frac{\partial^n}{\partial y_{i_1} \cdots \partial y_{i_n}} \big[\mathcal{M}^{(n)}_{i_1,\ldots,i_n}(t, \mathbf{y}) f(t, \mathbf{y})\big],} \tag{V.80}$$

involving all possible partial derivatives with respect to the $\{y_i\}$ of the single-time density $f(t, \mathbf{y})$.

V.A.2 Fokker–Planck equation

Truncating the above Kramers–Moyal expansion after the second order yields the corresponding *Fokker–Planck equation*

$$\boxed{\frac{\partial f(t, \mathbf{y})}{\partial t} = -\sum_{i=1}^{D} \frac{\partial}{\partial y_i} \big[\mathcal{M}^{(1)}_i(t, \mathbf{y}) f(t, \mathbf{y}) + \frac{1}{2} \sum_{i,j=1}^{D} \frac{\partial^2}{\partial y_i\, \partial y_j} \big[\mathcal{M}^{(2)}_{i,j}(t, \mathbf{y}) f(t, \mathbf{y})\big],} \tag{V.81}$$

Going back to the form (V.78b) of the jump moments, one sees that the second-order ("diffusion") coefficients $\mathcal{M}^{(2)}_{i,j}(t, \mathbf{y})$ in this equation directly reflect the second-order moments of the conditional probability for the change in $\mathbf{y}$ at time t. As such, the off-diagonal ones with $i \neq j$ are related to the correlation (in a loose sense) between the transition probabilities for the change in y_i and y_j: if these transition probabilities are independent—for instance because y_i and y_j evolve under different influences—, then the off-diagonal second-order terms in the Fokker–Planck equation (V.81) vanish.

CHAPTER VI

Linear response theory

When the Hamiltonian of a system—be it described classically or quantum mechanically—is known, the time evolution of its observables is fixed and obeys the appropriate deterministic equation (II.16) or (II.38). Consider for instance a classical system, governed by a (time-independent) Hamilton function H_0, which encodes all interactions between the system particles. If $O(t)$ is one of the observables of the system, there is the same amount of information in its values at two successive times t and t'.

On the other hand, if H_0 is unknown, measuring both $O(t)$ and $O(t')$ and correlating their values is likely to improve the knowledge on the system. For the correlation to be meaningful—on the theoretical side, one pair of successive values represent only one glimpse into a realization of a stochastic process; and on the other side, experimental uncertainties cannot be discarded—, the measurements have to be repeated many times in similar conditions(74) and their results averaged over. In this way one builds a *time correlation function* $\langle O(t)O(t')\rangle$. Technically, the "similar conditions" amount to an identical macrostate of the system, the state of choice being that of thermodynamic equilibrium, which leads to correlators $\langle O(t)O(t')\rangle_{\text{eq.}}$.

The procedure can be repeated for other observables, and one can even correlate the values taken at two instants by two different observables. This potentially leads to plenty of time-correlation functions, each of which encode some information about the system. More precisely, correlators built at thermodynamic equilibrium allow one to access the coefficients that characterize out-of-equilibrium states of the system. Since many kinds of departure from equilibrium are possible, one has to consider several correlation functions to describe them—in contrast to the equilibrium state, whose properties are entirely contained in the relevant partition function.

Restricting the discussion to near-equilibrium macrostates, the deviation of their properties from the equilibrium ones can be approximated as being linear in some appropriate small perturbation(s). This is similar to the assumed linearity of the fluxes in the affinities of Sec. I.2. Accordingly, each of the transport coefficients introduced in that chapter can be expressed in terms of the integral of a given correlation function, as we shall illustrate in Sec. VI.4. Before coming to that point, we first introduce time-correlation functions for homogeneous quantum-mechanical systems in Sec. VI.1. We then discuss the meaning of these functions (Sec. VI.2) and consider some of their more formal aspects (Sec. VI.3), which will in particular allow us to show the Onsager relations which have been introduces as postulated in § I.2.2 b. Eventually, we discuss in two appendices the important generalization to non-uniform systems as well as the classical theory of linear response.

Before going any further, let us emphasize that the formalism of linear response developed hereafter, even though limited to small departures from equilibrium, is not a phenomenological description, but a theory, based on exact quantum mechanical equations—which are dealt with perturbatively. As thus, the results discussed in Sec. VI.3 constitute stringent constraints for the parameters used in models.

---

(74) That is, for identically prepared systems.

VI.1 Time correlation functions

In this section, we introduce various functions relating the values taken by two observables A and B of a macroscopic system at thermodynamic equilibrium. We mostly consider the case of a quantum mechanical system, specified in Sec. VI.1.1, in which case the observables are Hermitian operators $\hat{A}$ and $\hat{B}$. For the sake of simplicity, we first consider observables associated to physical quantities which are uniform across the system under study, so that they do not depend on position, only on time. The generalization to non-uniform phenomena will be shortly presented in Sec. VI.A.1, and linear response in classical systems discussed in Sec. VI.B.1.

Since the operators $\hat{A}$ and $\hat{B}$ generally do not commute with each other, their are several possible choices of correlation functions, which we define in Secs. VI.1.2–VI.1.5. At the same time, we introduce their respective Fourier transforms and we indicate a few straightforward properties. However, we postpone the discussion of the physical content of each correlation function to next section, while their at times important mathematical properties will be studied at greater length in Sec. VI.3.

VI.1.1 Assumptions and notations

Consider an isolated quantum-mechanical system, governed by the Hamilton operator $\hat{H}_0$—acting on a Hilbert space which we need not specify—, whose eigenvalues and eigenstates are respectively denoted as $\{E_n\}$ and $\{\,|\phi_n\rangle\}$:

$$\hat{H}_0\,|\phi_n\rangle = E_n\,|\phi_n\rangle. \tag{VI.1}$$

The system is assumed to be initially in thermodynamic equilibrium at temperature T. The associated density operator $\hat{\rho}_{\text{eq.}}$ thus reads

$$\hat{\rho}_{\text{eq.}} = \frac{1}{Z(\beta)}\,\mathrm{e}^{-\beta\hat{H}_0}, \quad \text{with} \quad Z(\beta) = \mathrm{Tr}\,\mathrm{e}^{-\beta\hat{H}_0} \text{ and } \beta = \frac{1}{k_BT}. \tag{VI.2a}$$

This canonical density operator is quite obviously diagonal in the basis $\{\,|\phi_n\rangle\}$

$$\langle\phi_n|\,\hat{\rho}_{\text{eq.}}|\phi_{n'}\rangle = \frac{1}{Z(\beta)}\,\mathrm{e}^{-\beta E_n}\,\delta_{nn'} \equiv \pi_n\,\delta_{nn'}, \tag{VI.2b}$$

where the diagonal elements π_n represent the equilibrium populations of the energy eigenstates.

Let $\hat{O}$ denote a time-independent operator on the Hilbert space of the system. In the Heisenberg representation with respect to $\hat{H}_0$, it is represented by the operator [this is relation (II.39), here with $t_0 = 0$(75)]

$$\hat{O}_{\text{I}}(t) = \mathrm{e}^{\mathrm{i}\hat{H}_0t/\hbar}\,\hat{O}\,\mathrm{e}^{-\mathrm{i}\hat{H}_0t/\hbar}, \tag{VI.3}$$

where instead of H we used the subscript I, for "interaction picture", anticipating the fact that we shall often consider perturbations of the system. The expectation value of the observable in the equilibrium (macro)state reads

$$\big\langle\hat{O}_{\text{I}}(t)\big\rangle_{\text{eq.}} = \mathrm{Tr}\big[\hat{\rho}_{\text{eq.}}\hat{O}_{\text{I}}(t)\big] = \mathrm{Tr}\big[\hat{\rho}_{\text{eq.}}\hat{O}\big], \tag{VI.4}$$

where the second identity follows from the invariance of the trace under cyclic permutations and the commutativity of $\hat{\rho}_{\text{eq.}}$ and $\hat{H}_0$. That is, $\langle\hat{O}_{\text{I}}(t)\rangle_{\text{eq.}}$ is actually independent of time, which is to be expected since this is an expectation value at equilibrium.

Remark: Equation (VI.3) gives $\hat{O}_{\text{I}}(t\!=\!0) = \hat{O}$, which will allow us to write $\hat{O}$ instead of $\hat{O}_{\text{I}}(0)$.(75)

---

(75) The reader may check that adopting another choice for t_0 does not make any difference, except that it changes the reference point where $\hat{O}_{\text{I}}(t)$ coincides with the Schrödinger-representation $\hat{O}$.

Let $O_{nn'} \equiv \langle\phi_n|\,\hat{O}\,|\phi_{n'}\rangle$ denote the matrix elements of the observable $\hat{O}$ in the basis $\{\,|\phi_n\rangle\}$. Since the latter is formed of eigenstates of the Hamilton operator, one readily finds that the matrix elements of $\hat{O}_{\rm I}(t)$ are

$$[O_{\rm I}(t)]_{nn'} = O_{nn'}\,{\rm e}^{{\rm i}(E_n-E_{n'})t/\hbar} = O_{nn'}\,{\rm e}^{{\rm i}\omega_{nn'}t}, \tag{VI.5}$$

where in the second identity we have introduced the Bohr[(bp)] frequencies

$$\omega_{nn'} \equiv \frac{E_n - E_{n'}}{\hbar} \tag{VI.6}$$

of the system.

VI.1.2 Linear response function and generalized susceptibility

The system initially at equilibrium is submitted to a small excitation, which we shall also refer to as perturbation, described by a time-dependent additional term in the Hamiltonian in Schrödinger representation

$$\hat{W}(t) = -f(t)\,\hat{A}, \tag{VI.7}$$

where $\hat{A}$ is an observable of the system and $f(t)$ a given classical function of t, which is assumed to vanish for $t \to -\infty$. $f(t)$ is sometimes referred to as the *generalized force* conjugate to $\hat{A}$—which is then the corresponding “generalized displacement”.

At $t \to -\infty$, the system is thus in the macroscopic state (VI.2), and the excitation (VI.7) drives it out of equilibrium; if the where the perturbation is weak, the resulting departure will remain small. The various observables $\hat{B}$ of the system then acquire expectation values $\langle\hat{B}_{\rm I}(t)\rangle_{\rm n.eq.}$ which will in general differ from their respective equilibrium values $\langle\hat{B}_{\rm I}(t)\rangle_{\rm eq.}$ as given by Eq. (VI.4).

Remarks:

∗ In this section, since $\hat{A}$ and $\hat{B}$ are observables, they are Hermitian operators. To ensure the hermiticity of the Hamiltonian, $f(t)$ should be real-valued. In Secs. VI.1.3–VI.1.5, we shall more generally consider time-correlation functions of operators that need not necessarily be Hermitian.

∗ Excitations which can be described by an extra term in the Hamiltonian, as we consider here, are often referred to as *mechanical*, in opposition to *thermal* disturbances — as e.g. a temperature gradient, which cannot trivially be rendered by a shift in the Hamiltonian. The former are driven by external forces, which an experimenter may control, while the former rather arise from internal, “thermodynamical” forces. It is sometimes also possible to deal with thermal excitations by engineering theoretical forces which allows one to use the formalism developed for mechanical perturbations.

VI.1.2 a Linear response function

To describe the linear response of the system to the perturbation (VI.7), one introduces the *(linear) response function* χ_{BA} such that

$$\boxed{\big\langle\hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} = \big\langle\hat{B}\,\big\rangle_{\rm eq.} + \int_{-\infty}^{\infty}\chi_{BA}(t-t')\,f(t')\,{\rm d}t' + \mathcal{O}(f^2).} \tag{VI.8}$$

In this definition, the upper boundary of the integral extends to $+\infty$, i.e. formally involves the generalized force in the future of the time t at which the effect $\langle\hat{B}_{\rm I}(t)\rangle_{\rm n.eq.}$ is considered, which violates causality. To restore the latter, one may either restrict the integral to the interval $-\infty < t' \leq t$—which is indeed what cones out of the explicit calculation of the linear response, as we shall see in Sec. VI.2.1 below—, or define the response function such that it vanishes for $t' > t$, i.e. $\tau \equiv t-t' < 0$, as we shall do:

$$\chi_{BA}(\tau) = 0 \quad \text{for} \quad \tau < 0. \tag{VI.9}$$

[(bp)] N. Bohr, 1885–1962

To emphasize the causality, the linear response function χ_{BA} is often called *after-effect function*.

If the perturbation $f(t)$ is an impulse, i.e. $f(t) \propto \delta(t)$, relation (VI.8) shows that the linear response $\langle\hat{B}_{\rm I}(t)\rangle_{\rm n.eq.} - \langle\hat{B}\rangle_{\rm eq.}$ is directly proportional to $\chi_{BA}(t)$. Accordingly, χ_{BA} is also referred to as *impulse response function*—in particular in signal theory—or, to account simultaneously for its causality property, *retarded Green*(bq) *function* or *retarded propagator*.

Remarks:

∗ Equation (VI.8) represents a linear (in the regime of small excitations), causal, and time-translation invariant relation between an “input” $f(t')$ and an “output” $\langle\hat{B}_{\rm I}(t)\rangle_{\rm n.eq.} - \langle\hat{B}\rangle_{\rm eq.}$. The response function χ_{BA} thus plays the role of a *linear filter*.

∗ The response described by the formula (VI.8) is obviously no longer “Markovian” as were the relations between fluxes and affinities considered in Sec.I.2. The memoryless case could a priori be recovered by taking the response function $\chi_{BA}(\tau)$ proportional to $\delta(\tau)$, yet we shall later see that such a behavior cannot be accommodated within linear response theory, for it leads to the violation of important relations (see the “Comparison” paragraph at the end of § VI.4.3 c).

∗ The dependence of the linear response function on the observables $\hat{A}$ and $\hat{B}$ is readily obtained, yet we postpone its derivation for later (Sec. VI.2.1). Without any calculation, it should be clear to the reader that if the departure from equilibrium $\langle\hat{B}_{\rm I}(t)\rangle_{\rm n.eq.} - \langle\hat{B}\rangle_{\rm eq.}$ is to be linear in the perturbation, i.e. of order $\mathcal{O}(f)$, then Eq. (VI.8) implies that χ_{BA} should be of order $\mathcal{O}(f^0)$, that is, χ_{BA} depends only on *equilibrium* quantities.

∗ The causality property (VI.9) encoded in the linear response functions will strongly constraint its Fourier transform, as we shall see in Sec. VI.3.1.

∗ Relation (VI.8) is sometimes called *Kubo*(br) *formula*, although the denomination is also often attached to another, equivalent form of the equation [see Eq. (VI.51) below].

∗ Since $\hat{B}$ is a Hermitian operator, its expectation values in or out of equilibrium are real numbers. Assuming $f(t) \propto \delta(t)$, one then finds that the retarded propagator $\chi_{BA}(\tau)$ is real-valued.

VI.1.2 b Generalized susceptibility

The integral in the defining relation (VI.8) is a time convolution, which suggests Fourier transforming to frequency space.

Accordingly, one introduces the Fourier transform of the response function as

$$\tilde{\chi}_{BA}(\omega) = \lim_{\varepsilon\to 0^+} \int_{-\infty}^{\infty} \chi_{BA}(\tau)\, {\rm e}^{{\rm i}\omega\tau}\, {\rm e}^{-\varepsilon\tau}\, {\rm d}\tau, \tag{VI.10a}$$

where an exponential factor ${\rm e}^{-\varepsilon\tau}$ with $\varepsilon > 0$ was inserted for the sake of ensuring the convergence—as we shall see later, one can easily check that $\chi_{BA}(\tau)$ does not diverge as $\tau \to \infty$, so that this factor is sufficient. $\tilde{\chi}_{BA}(\omega)$ is referred to as *(generalized) susceptibility*, *generalized admittance* or *frequency response function*.

The inverse Fourier transform reads(76)

$$\chi_{BA}(\tau) = \int_{-\infty}^{\infty} \tilde{\chi}_{BA}(\omega)\, {\rm e}^{-{\rm i}\omega\tau}\, \frac{{\rm d}\omega}{2\pi}. \tag{VI.10b}$$

---

(76) Here we implicitly assume that $\tilde{\chi}_{BA}(\omega)$ fulfills some integrability condition.

(bq) G. Green, 1793–1841 (br) R. Kubo, 1920–1995

Remark: If the equilibrated system is subject to a sinusoidal force $f(t) = f_\omega \cos\omega t = \mathrm{Re}\big(f_\omega\, \mathrm{e}^{-\mathrm{i}\omega t}\big)$ with $f_\omega \in \mathbb{R}$, then its linear response (VI.8) reads[(77)]

$$\big\langle \hat{B}_{\mathrm{I}}(t)\big\rangle_{\text{n.eq.}} = \big\langle \hat{B}\,\big\rangle_{\text{eq.}} + \mathrm{Re}\big(f_\omega \tilde{\chi}_{BA}(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\big),$$

i.e. the response is also sinusoidal, although it will in general be out of phase with the exciting force, since $\tilde{\chi}_{BA}(\omega)$ is complex.

VI.1.3 Non-symmetrized and symmetrized correlation functions

In this section and the next two ones, we consider generic operators $\hat{A}$ and $\hat{B}$, which might not necessarily be Hermitian unless we specify it.

In the equilibrium state $\hat{\rho}_{\text{eq.}}$, the *non-symmetrized correlation function* between the operators $\hat{A}$ and $\hat{B}$ at different times t, t' is defined as

$$C_{BA}(t,t') \equiv \big\langle \hat{B}_{\mathrm{I}}(t)\hat{A}_{\mathrm{I}}(t')\big\rangle_{\text{eq.}}. \tag{VI.11}$$

Due the stationarity of the equilibrium state, the expectation value on the right-hand side is invariant under time translations, so that $C_{BA}(t,t') = C_{BA}(t-t',0)$.

Mathematically, the latter identity is easily checked by inserting explicitly the terms $\mathrm{e}^{\pm\mathrm{i}\hat{H}_0 t/\hbar}$, $\mathrm{e}^{\pm\mathrm{i}\hat{H}_0 t'/\hbar}$ in definition (VI.11) and by using the invariance of the trace under cyclic permutations and the commutativity of $\hat{\rho}_{\text{eq.}}$ and $\hat{H}_0$.

One may thus replace the two variables by their difference $\tau \equiv t - t'$ and define equivalently

$$C_{BA}(\tau) \equiv \big\langle \hat{B}_{\mathrm{I}}(\tau)\hat{A}\,\big\rangle_{\text{eq.}}, \tag{VI.12}$$

where we used $\hat{A}_{\mathrm{I}}(0) = \hat{A}$.[(75)]

Introducing the representation of the operators in the basis of the energy eigenstates, a straightforward calculation gives the alternative form

$$C_{BA}(\tau) = \sum_{n,n'} \pi_n B_{nn'} A_{n'n}\, \mathrm{e}^{-\mathrm{i}\omega_{n'n}\tau}, \tag{VI.13}$$

from where follows at once the Fourier transform

$$\tilde{C}_{BA}(\omega) \equiv \int_{-\infty}^{\infty} C_{BA}(\tau)\,\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau = 2\pi \sum_{n,n'} \pi_n\, B_{nn'} A_{n'n}\, \delta(\omega - \omega_{n'n}). \tag{VI.14}$$

This expression shows the generic property of the dependence of correlation functions on the excitations frequencies of the system. More precisely, $\tilde{C}_{BA}(\omega)$ clearly diverges when the angular frequency ω coincides with one of the Bohr frequencies of the system, unless the associated matrix element of $\hat{A}$ or $\hat{B}$ vanishes—for instance, because the corresponding transition between energy eigenstates is forbidden by some selection rule.

Even if $\hat{A}$ and $\hat{B}$ are Hermitian, the correlation function (VI.12) is generally not real-valued: with the cyclicity of the trace, one finds

$$C_{BA}(\tau)^* = \big\langle \hat{B}_{\mathrm{I}}(\tau)\hat{A}\,\big\rangle^*_{\text{eq.}} = \big\langle \hat{A}^\dagger \hat{B}^\dagger_{\mathrm{I}}(\tau)\big\rangle_{\text{eq.}} = C_{A^\dagger B^\dagger}(-\tau), \tag{VI.15}$$

which has in general no obvious relation to $C_{BA}(\tau)$.

[(77)] Strictly speaking, the condition $f(t) \to 0$ as $t \to -\infty$ does not hold here, yet one can easily branch the sinusoidal perturbation adiabatically—mathematically with a factor $\mathrm{e}^{\varepsilon t}$ for $-\infty < t \leq 0$ with $\epsilon \to 0^+$—and recover the same result for $t > 0$.

To characterize the correlations between observables $\hat{A}$ and $\hat{B}$ by a real-valued quantity, one introduces the *symmetric correlation function*

$$\boxed{S_{BA}(\tau) \equiv \frac{1}{2}\Big\langle \big\{\hat{B}_{\rm I}(\tau), \hat{A}\big\}_+ \Big\rangle_{\rm eq.}\,,} \tag{VI.16}$$

where $\{\,\cdot\,,\,\cdot\,\}_+$ denotes the anticommutator of two operators. Using definition (VI.12), one finds at once

$$S_{BA}(\tau) = \frac{1}{2}\big[C_{BA}(\tau) + C_{AB}(-\tau)\big]. \tag{VI.17}$$

In the case of observables, the associated operators are Hermitian, $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, so that relation (VI.15) reads $C_{AB}(-\tau) = C_{BA}(\tau)^*$, which implies that $S_{BA}(\tau)$ is a real number.

The representation (VI.13) of the non-symmetrized correlation function gives

$$C_{AB}(-\tau) = \sum_{n,n'} \pi_n A_{nn'} B_{n'n}\, {\rm e}^{{\rm i}\omega_{n'n}\tau} = \sum_{n,n'} \pi_n A_{nn'} B_{n'n}\, {\rm e}^{-{\rm i}\omega_{nn'}\tau} = \sum_{n,n'} \pi_{n'} A_{n'n} B_{nn'}\, {\rm e}^{-{\rm i}\omega_{n'n}\tau},$$

where the second identity comes from the obvious identity $\omega_{nn'} = -\omega_{n'n}$, while the last ones follows from exchanging the dummy indices n, n'. Taking the half sum of this decomposition and Eq. (VI.13) and Fourier transforming, one finds

$$\tilde{S}_{BA}(\omega) = \pi \sum_{n,n'} (\pi_n + \pi_{n'}) B_{nn'} A_{n'n}\, \delta(\omega - \omega_{n'n}). \tag{VI.18}$$

VI.1.4 Spectral density

Instead of considering (half) the expectation value of the anticommutator of $\hat{B}_{\rm I}(\tau)$ and $\hat{A}$, one can also think of introducing that of the commutator. Inserting a factor $1/\hbar$ to ensure a proper behavior in the classical limit, we thus define

$$\boxed{\xi_{BA}(\tau) \equiv \frac{1}{2\hbar}\Big\langle \big[\hat{B}_{\rm I}(\tau), \hat{A}\big] \Big\rangle_{\rm eq.}\,.} \tag{VI.19}$$

Repeating identically the steps leading from the definition (VI.16) of the symmetrized correlation function to its Fourier transform (VI.18), one finds the so-called *spectral function* (or *spectral density*)

$$\boxed{\tilde{\xi}_{BA}(\omega) \equiv \frac{\pi}{\hbar} \sum_{n,n'} (\pi_n - \pi_{n'}) B_{nn'} A_{n'n}\, \delta(\omega - \omega_{n'n}),} \tag{VI.20}$$

with as usual $\omega_{n'n}$ the Bohr frequencies of the system, and $A_{nn'}$, $B_{nn'}$ the matrix elements of the operators $\hat{A}$, $\hat{B}$ between energy eigenstates.

The spectral density is in general complex-valued, yet becomes real-valued when one considers the autocorrelation of an observable, i.e. for $\hat{B} = \hat{A}^\dagger = \hat{A}$, in which case $B_{nn'}A_{n'n} = |A_{n'n}|^2$.

VI.1.5 Canonical correlation function

Last we introduce Kubo's *canonical correlation function*, defined as [50]

$$\boxed{K_{BA}(\tau) \equiv \frac{1}{\beta}\int_0^\beta \Big\langle {\rm e}^{\lambda\hat{H}_0}\hat{A}\,{\rm e}^{-\lambda\hat{H}_0}\hat{B}_{\rm I}(\tau) \Big\rangle_{\rm eq.} {\rm d}\lambda = \frac{1}{\beta}\int_0^\beta \Big\langle \hat{A}_{\rm I}(-{\rm i}\hbar\lambda)\hat{B}_{\rm I}(\tau) \Big\rangle_{\rm eq.} {\rm d}\lambda,} \tag{VI.21}$$

for a system governed by the Hamilton operator $\hat{H}_0$, where β is the inverse temperature of the equilibrium state.

Using the explicit form of the equilibrium distribution $\hat{\rho}_{\rm eq.}$—or equivalently, of the populations π_n of the energy eigenstates at canonical equilibrium—, one finds the Fourier transform

$$\tilde{K}_{BA}(\omega) \equiv \int_{-\infty}^{\infty} K_{BA}(\tau)\, {\rm e}^{{\rm i}\omega\tau}\, {\rm d}\tau = 2\pi \sum_{n,n'} \frac{\pi_n - \pi_{n'}}{\beta\hbar\omega_{n'n}} B_{nn'} A_{n'n}\, \delta(\omega - \omega_{n'n}). \tag{VI.22}$$

Proof of the spectral decomposition (VI.22):
The equilibrium expectation value in the integrand of definition (VI.21) reads

$$\sum_{n,n'} \pi_{n'}\, {\rm e}^{\lambda E_{n'}} A_{n'n}\, {\rm e}^{-\lambda E_n} B_{nn'} {\rm e}^{{\rm i}\omega_{nn'}\tau} = \sum_{n,n'} \pi_{n'}\, {\rm e}^{\lambda\hbar\omega_{n'n}} B_{nn'} A_{n'n} {\rm e}^{-{\rm i}\omega_{n'n}\tau}.$$

The integration over λ is straightforward and gives

$$K_{BA}(\tau) = \frac{1}{\beta} \sum_{n,n'} \frac{{\rm e}^{\beta\hbar\omega_{n'n}} - 1}{\hbar\omega_{n'n}} \pi_{n'} B_{nn'} A_{n'n} {\rm e}^{-{\rm i}\omega_{n'n}\tau} = \sum_{n,n'} \frac{\pi_n - \pi_{n'}}{\beta\hbar\omega_{n'n}} B_{nn'} A_{n'n} {\rm e}^{-{\rm i}\omega_{n'n}\tau}, \tag{VI.23}$$

where the second identity comes from ${\rm e}^{\beta\hbar\omega_{n'n}} = \pi_n/\pi_{n'}$, which follows from Eq. (VI.2b). This alternative representation of the Kubo correlation function leads at once to the Fourier transform (VI.22).

VI.2 Physical meaning of the correlation functions

In the previous Section, we left a few issues open. First we defined the linear response function by introducing its role in a given physical situation, but did not attempt to compute it, which will now be done perturbatively in Sec. VI.2.1. Adopting then a somewhat opposite approach, we introduced several correlation functions mathematically, without discussing the physical phenomena they embody. Again, we shall remedy this now, in Secs. VI.2.2–VI.2.4.

VI.2.1 Calculation of the linear response function

We consider the quantum-mechanical system of Sec. VI.1.1, submitted to the small perturbation $\hat{W}(t) = -f(t)\hat{A}$ described in Sec. VI.1.2.

Let $\hat{\rho}_{\rm I}(t) \equiv {\rm e}^{{\rm i}\hat{H}_0 t/\hbar}\, \hat{\rho}\, {\rm e}^{-{\rm i}\hat{H}_0 t/\hbar}$ denote the interaction-picture representation of the density operator. Since the free evolution under the effect of $\hat{H}_0$ is accounted for by the transformation, the evolution in the presence of the perturbation (VI.7) is governed by

$$\frac{{\rm d}\hat{\rho}_{\rm I}(t)}{{\rm d}t} = \frac{1}{{\rm i}\hbar}\big[\hat{W}_{\rm I}(t), \hat{\rho}_{\rm I}(t)\big], \quad \text{with} \quad \hat{W}_{\rm I}(t) = -f(t)\hat{A}_{\rm I}(t), \tag{VI.24}$$

where $\hat{A}_{\rm I}(t)$ is defined according to Eq. (VI.3). To first order in $f(t)$, the solution to this equation with the initial condition $\hat{\rho}_{\rm I}(-\infty) = \hat{\rho}_{\rm eq.}$ is

$$\hat{\rho}_{\rm I}(t) = \hat{\rho}_{\rm eq.} + \frac{\rm i}{\hbar}\int_{-\infty}^{t} f(t')\big[\hat{A}_{\rm I}(t'), \hat{\rho}_{\rm I}(t')\big]\, {\rm d}t' = \hat{\rho}_{\rm eq.} + \frac{\rm i}{\hbar}\int_{-\infty}^{t} f(t')\big[\hat{A}_{\rm I}(t'), \hat{\rho}_{\rm eq.}\big]\, {\rm d}t' + \mathcal{O}(f^2).$$

Multiplying with $\hat{B}_{\rm I}(t)$ and taking the trace, this leads to

$$\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} = {\rm Tr}\big[\hat{B}_{\rm I}(t)\hat{\rho}_{\rm I}(t)\big] = \big\langle \hat{B}\,\big\rangle_{\rm eq.} + \frac{\rm i}{\hbar}\int_{-\infty}^{t} f(t')\, {\rm Tr}\big\{\hat{B}_{\rm I}(t)\big[\hat{A}_{\rm I}(t'), \hat{\rho}_{\rm eq.}\big]\big\}\, {\rm d}t' + \mathcal{O}(f^2). \tag{VI.25}$$

Developing explicitly the commutator and using the invariance of the trace under cyclic permutations, so as to isolate $\hat{\rho}_{\rm eq.}$, one finds that the trace in the integrand can be rewritten as

$${\rm Tr}\big\{\hat{\rho}_{\rm eq.}\big[\hat{B}_{\rm I}(t), \hat{A}_{\rm I}(t')\big]\big\} = \Big\langle \big[\hat{B}_{\rm I}(t), \hat{A}_{\rm I}(t')\big]\Big\rangle_{\rm eq.} = \Big\langle \big[\hat{B}_{\rm I}(t-t'), \hat{A}_{\rm I}(0)\big]\Big\rangle_{\rm eq.},$$

where the last identity comes from inserting the terms ${\rm e}^{{\rm i}\hat{H}_0 t/\hbar}$, ${\rm e}^{{\rm i}\hat{H}_0 t'/\hbar}$ and their complex conjugates and invoking the invariance of the trace under cyclic permutations and the commutativity of $\hat{\rho}_{\rm eq.}$ and $\hat{H}_0$. In addition, one can insert a Heaviside function $\Theta(t-t')$ in the integrand, so as to extend the upper bound of the integral to ∞.

All in all, this yields the *Kubo formula*

$$\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} = \big\langle \hat{B}\,\big\rangle_{\rm eq.} + \int_{-\infty}^{\infty} \chi_{BA}(t-t')\, f(t')\,{\rm d}t' + \mathcal{O}(f^2), \qquad \text{(VI.8)}$$

with χ_{BA} explicitly given as

$$\boxed{\chi_{BA}(\tau) = \frac{\rm i}{\hbar}\Big\langle\big[\hat{B}_{\rm I}(\tau),\hat{A}\big]\Big\rangle_{\rm eq.}\Theta(\tau).} \qquad \text{(VI.26)}$$

As had already been anticipated, the retarded propagator (VI.26), which characterizes the behavior of the system when it is driven out of equilibrium by an external perturbation, can actually be expressed in terms of a two-time average in the equilibrium state.

Remarks:

$*$ When it can be performed, the computation of the interaction-picture representation of the operator conjugate to the perturbing force $f(t)$ allows one to derive the response function at once.

$*$ The first identity in Eq. (VI.25) embodies *Duhamel's*[(bs)] *principle* for the solution of the linear differential equation (VI.24), expressing it in terms of the initial condition (here at $t_0 = -\infty$) and the history between t_0 and t.

$*$ Searching for solutions to the evolution equation (VI.24) of the type

$$\hat{\rho}_I(t) = \hat{\rho}_I^{(0)}(t) + \hat{\rho}_I^{(1)}(t) + \cdots + \hat{\rho}_I^{(k)}(t) + \cdots \quad \text{with} \quad \hat{\rho}_I^{(k)}(t) = \mathcal{O}\big(f(t)^k\big),$$

where necessarily $\hat{\rho}_I^{(0)}(t) = \hat{\rho}_{\rm eq.}$, one easily finds the recurrence relation

$$\frac{{\rm d}\hat{\rho}_{\rm I}^{(k)}(t)}{{\rm d}t} = \frac{1}{{\rm i}\hbar}\big[\hat{W}_{\rm I}(t),\hat{\rho}_{\rm I}^{(k-1)}(t)\big],$$

which under consideration of the initial conditions $\hat{\rho}_I^{(k)}(-\infty) = \delta_{k0}\hat{\rho}_{\rm eq.}$ leads to

$$\hat{\rho}_{\rm I}^{(k)}(t) = \delta_{k0}\hat{\rho}_{\rm eq.} + \frac{\rm i}{\hbar}\int_{-\infty}^{t} f(t')\big[\hat{A}_{\rm I}(t'),\hat{\rho}_{\rm I}^{(k-1)}(t')\big]\,{\rm d}t'.$$

This relation allows one to derive the response to arbitrary order, i.e. to go beyond linear response.

Expression (VI.26) leads at once to the alternative representation

$$\chi_{BA}(\tau) = \frac{\rm i}{\hbar}\Theta(\tau)\sum_{n,n'}(\pi_n - \pi_{n'})B_{nn'}A_{n'n}\,{\rm e}^{-{\rm i}\omega_{n'n}\tau}. \qquad \text{(VI.27)}$$

This identity shows that $\chi_{BA}(\tau)$ does not diverge at $\tau\to\infty$.

Inserting this form in the definition (VI.10a) of the generalized susceptibility, one finds

$$\boxed{\tilde{\chi}_{BA}(\omega) = \frac{1}{\hbar}\sum_{n,n'}(\pi_n - \pi_{n'})B_{nn'}A_{n'n}\lim_{\varepsilon\to 0^+}\frac{1}{\omega_{n'n}-\omega-{\rm i}\varepsilon}.} \qquad \text{(VI.28)}$$

This decomposition is sometimes referred to as the *Lehmann*[(bt)] (*spectral*) *representation*.

Comparing this expression of the generalized susceptibility $\tilde{\chi}_{BA}(\omega)$ with the definition (VI.20) of the spectral density $\tilde{\xi}_{BA}(\omega)$ relative to the same operators $\hat{A}$ and $\hat{B}$, one finds a first relation between both functions, namely

$$\tilde{\chi}_{BA}(\omega) = \frac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{\tilde{\xi}_{BA}(\omega')}{\omega'-\omega-{\rm i}\varepsilon}\,{\rm d}\omega'. \qquad \text{(VI.29)}$$

[(bs)] J.-M. Duhamel, 1797–1872 [(bt)] H. Lehmann, 1924–1998

VI.2.2 Dissipation

Consider a system subject to a harmonic perturbation $\hat{W}\cos\omega t$ with constant $\hat{W}$. A straightforward application of time-dependent perturbation theory in quantum mechanics yields for the transition rate from an arbitrary initial state $|\phi_{\rm i}\rangle$ to the set of corresponding final states $|\phi_{\rm f}\rangle$

$$\Gamma_{\rm i\to f} = \frac{\pi}{2\hbar^2}\sum_{\rm f}\big|\langle\phi_{\rm f}|\,\hat{W}\,|\phi_{\rm i}\rangle\big|^2\big[\delta(\omega_{\rm fi}-\omega)+\delta(\omega_{\rm if}-\omega)\big]. \tag{VI.30}$$

If the system absorbs energy from the perturbation, then $E_{\rm f} > E_{\rm i}$, so that only the first δ-term contributes. On the other hand, energy release from the system, corresponding to "induced emission", requires $E_{\rm f} < E_{\rm i}$, i.e. involves the second δ-term.

We apply this result to the system of Sec. VI.1.1, initially at thermodynamic equilibrium. Let $f(t) = f_\omega\cos\omega t$, with constant real f_ω, be a "force" coupling to an operator $\hat{A}$ acting on the system, leading to perturbation (VI.7) in the Hamiltonian.

The probability per unit time that the system absorb a quantum $\hbar\omega$ of energy starting from an initial state $|\phi_n\rangle$ is given by Eq. (VI.30) with $|\phi_{\rm i}\rangle = |\phi_n\rangle$ and keeping only the first δ-term. Multiplying this probability by the initial state population π_n and by the absorbed energy $\hbar\omega$, and summing over all possible initial states, one finds the total energy received per unit time by the system

$$\mathscr{P}_{\rm gain} = \frac{\pi f_\omega^2}{2\hbar^2}\sum_{n,n'}\pi_n\hbar\omega\big|A_{nn'}\big|^2\delta(\omega_{n'n}-\omega).$$

An analogous reasoning gives for the total energy emitted by the system per unit time

$$-\mathscr{P}_{\rm loss} = \frac{\pi f_\omega^2}{2\hbar^2}\sum_{n,n'}\pi_n\hbar\omega\big|A_{nn'}\big|^2\delta(\omega_{nn'}-\omega) = \frac{\pi f_\omega^2}{2\hbar^2}\sum_{n,n'}\pi_{n'}\hbar\omega\big|A_{nn'}\big|^2\delta(\omega_{n'n}-\omega),$$

where in the second identity we have exchanged the roles of the two dummy indices. Note that since the system is in thermodynamic equilibrium, the lower energy levels are more populated, so that the absorption term is larger than the emission term.

Adding $\mathscr{P}_{\rm gain}$ and $\mathscr{P}_{\rm loss}$ yields the net total energy exchanged—and actually absorbed by the system—per unit time, $\mathrm{d}E_{\rm tot}/\mathrm{d}t$, when it is submitted to a perturbation $-f_\omega\hat{A}\cos\omega t$, namely

$$\frac{\mathrm{d}E_{\rm tot}}{\mathrm{d}t} = \frac{\pi f_\omega^2\omega}{2\hbar}\sum_{n,n'}(\pi_n-\pi_{n'})\big|A_{nn'}\big|^2\delta(\omega_{n'n}-\omega).$$

Under consideration of the definition (VI.20) of the spectral density, this also reads

$$\frac{\mathrm{d}E_{\rm tot}}{\mathrm{d}t} = \frac{f_\omega^2}{2}\,\omega\,\tilde{\xi}_{A^\dagger A}(\omega). \tag{VI.31}$$

The spectral function $\tilde{\xi}_{A^\dagger A}(\omega)$ thus characterizes the dissipation of energy in the system when it is submitted to a small perturbation proportional to $\hat{A}\cos\omega t$.

VI.2.3 Relaxation

Let us now assume that the external "force" at equilibrium in the excitation (VI.7) acting on the system of Sec. VI.1.1 is given by

$$f(t) = f\,\mathrm{e}^{\varepsilon t}\,\Theta(-t), \tag{VI.32}$$

with f constant, where at the end of calculations we shall take the limit $\varepsilon\to 0^+$. This force represents a perturbation turned on from $t = -\infty$ over the typical scale ε^{-1}, slowly driving the system out of its initial equilibrium state. At $t = 0$, the perturbation is turned off, and the system then relaxes to the original equilibrium state. We shall now compute the departure from equilibrium of the expectation value of an operator $\hat{B}$ due to this excitation.

Inserting Eq. (VI.32) in the Kubo formula (VI.8), one finds

$$\begin{aligned}\frac{1}{f}\Big[\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.}\Big] &= \int_{-\infty}^{\infty} \chi_{BA}(t-t')\,{\rm e}^{\varepsilon t'}\,\Theta(-t')\,{\rm d}t' \\ &= \int_{-\infty}^{\infty} {\rm e}^{\varepsilon t'}\,\Theta(-t')\int_{-\infty}^{\infty} \tilde{\chi}_{BA}(\omega)\,{\rm e}^{-{\rm i}\omega(t-t')}\,\frac{{\rm d}\omega}{2\pi}\,{\rm d}t',\end{aligned}$$

where we have introduced the generalized susceptibility. Exchanging the order of the integrations and performing that over t' give

$$\frac{1}{f}\Big[\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.}\Big] = \int_{-\infty}^{\infty} \frac{\tilde{\chi}_{BA}(\omega)}{{\rm i}\omega+\varepsilon}\,{\rm e}^{-{\rm i}\omega t}\,\frac{{\rm d}\omega}{2\pi}.$$

That is, the linear response $\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.}$ is proportional to the inverse Fourier transform of the ratio of the generalized susceptibility over ${\rm i}\omega+\varepsilon = {\rm i}(\omega-{\rm i}\varepsilon)$.

Expressing $\tilde{\chi}_{BA}(\omega)$ in terms of the spectral function with Eq. (VI.29) and exchanging the order of the integrals, the above relation becomes

$$\frac{1}{f}\Big[\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.}\Big] = \frac{1}{\pi}\lim_{\varepsilon'\to 0^+}\int_{-\infty}^{\infty} \tilde{\xi}_{BA}(\omega')\int_{-\infty}^{\infty} \frac{{\rm e}^{-{\rm i}\omega t}}{(\omega-{\rm i}\varepsilon)(\omega'-\omega-{\rm i}\varepsilon')}\,\frac{{\rm d}\omega}{2\pi{\rm i}}\,{\rm d}\omega'. \qquad \text{(VI.33)}$$

The integration over ω is then straightforward with the theorem of residues, where the term ${\rm e}^{-{\rm i}\omega t}$ dictates whether the integration contour consisting of the real axis and a half-circle at infinity should be closed in the upper (for $t<0$) or in the lower (for $t>0$) complex half-plane of the variable ω.

- For $t<0$, one has to consider the only pole of the integrand in the upper half-plane, which lies at $\omega = {\rm i}\varepsilon$. The corresponding residue is ${\rm e}^{\varepsilon t}/(\omega'-{\rm i}\varepsilon-{\rm i}\varepsilon')$, which yields

  $$\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.} = \frac{f\,{\rm e}^{\varepsilon t}}{\pi}\lim_{\varepsilon'\to 0^+}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega'-{\rm i}\varepsilon-{\rm i}\varepsilon'}\,{\rm d}\omega' = \frac{f(t)}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega'-{\rm i}\varepsilon}\,{\rm d}\omega'.$$

  Taking now the limit $\varepsilon\to 0^+$ and using relation (VI.29) for $\omega=0$, one obtains

  $$\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.} = f(t)\,\tilde{\chi}_{BA}(0) \quad \text{for } t\le 0. \qquad \text{(VI.34)}$$

  This result is easily interpreted: the system is driven out of equilibrium so slowly that the departure of the expectation value of $\hat{B}_{\rm I}(t)$ from the equilibrium value can be computed with the help of the *static* susceptibility, i.e. $\tilde{\chi}_{BA}(0)$ at zero frequency.

- For $t>0$, the only pole in the lower half-plane of the integrand in Eq. (VI.33) is at $\omega=\omega'-{\rm i}\varepsilon'$. This leads to

  $$\begin{aligned}\frac{1}{f}\Big[\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} - \big\langle \hat{B}\,\big\rangle_{\rm eq.}\Big] &= \frac{1}{\pi}\lim_{\varepsilon'\to 0^+}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega'-{\rm i}\varepsilon-{\rm i}\varepsilon'}\,{\rm e}^{-{\rm i}(\omega'-{\rm i}\varepsilon')t}\,{\rm d}\omega' \quad \text{for } t>0 \qquad \text{(VI.35)} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega'-{\rm i}\varepsilon}\,{\rm e}^{-{\rm i}\omega' t}\,{\rm d}\omega'.\end{aligned}$$

  Replacing the spectral density by its explicit expression (VI.20) and inverting the sum and the integral, one finds

  $$\begin{aligned}\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega'-{\rm i}\varepsilon}\,{\rm e}^{-{\rm i}\omega' t}\,{\rm d}\omega' &= \frac{1}{\hbar}\sum_{n,n'}(\pi_n-\pi_{n'})B_{nn'}A_{n'n}\int_{-\infty}^{\infty} \frac{{\rm e}^{-{\rm i}\omega' t}}{\omega'-{\rm i}\varepsilon}\,\delta(\omega-\omega_{n'n})\,{\rm d}\omega' \\ &= \frac{1}{\hbar}\sum_{n,n'}(\pi_n-\pi_{n'})B_{nn'}A_{n'n}\,\frac{{\rm e}^{-{\rm i}\omega_{n'n}t}}{\omega_{n'n}-{\rm i}\varepsilon}.\end{aligned}$$

In the limit $\varepsilon \to 0^+$, one recognizes on the right-hand side the spectral decomposition (VI.23) of the canonical correlation function

$$\frac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{\tilde{\xi}_{BA}(\omega)}{\omega-\mathrm{i}\varepsilon}\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega = \beta K_{BA}(t). \tag{VI.36}$$

Inserting this identity in the above expression of the departure from equilibrium of the average value of $\hat{B}_{\mathrm{I}}(t)$ finally gives

$$\frac{1}{f}\Big[\big\langle\hat{B}_{\mathrm{I}}(t)\big\rangle_{\text{n.eq.}} - \big\langle\hat{B}\,\big\rangle_{\text{eq.}}\Big] = \beta K_{BA}(t) \quad \text{for } t>0. \tag{VI.37}$$

That is, the Kubo correlation function describes the relaxation from an out-of-equilibrium state—which justifies why Kubo called it[(78)] "relaxation function" in his original paper [50].

Remark: Instead of letting ε' go to 0^+ in Eq. (VI.35), one may as well take first the (trivial) limit $\varepsilon \to 0^+$. The remaining denominator of the integrand is exactly the factor multiplying t in the complex exponential. Differentiating both sides of the identity with respect to t and considering afterwards $\varepsilon' \to 0^+$ then gives

$$\frac{1}{f}\frac{\mathrm{d}}{\mathrm{d}t}\Big[\big\langle\hat{B}_{\mathrm{I}}(t)\big\rangle_{\text{n.eq.}} - \big\langle\hat{B}\,\big\rangle_{\text{eq.}}\Big] = -2\mathrm{i}\xi_{BA}(t).$$

Taking the initial condition at time $t=0$ from Eq. (VI.34), one finally obtains

$$\frac{1}{f}\Big[\big\langle\hat{B}_{\mathrm{I}}(t)\big\rangle_{\text{n.eq.}} - \big\langle\hat{B}\,\big\rangle_{\text{eq.}}\Big] = \tilde{\chi}_{BA}(0) - 2\mathrm{i}\int_0^t \xi_{BA}(t')\,\mathrm{d}t' \quad \text{for } t>0. \tag{VI.38}$$

Thus, the relaxation at $t>0$ involves the integral of the Fourier transform of the spectral function.

VI.2.4 Fluctuations

Taking $\hat{B}=\hat{A}$ and setting $\tau=0$ in definitions (VI.12) or (VI.16), one finds

$$C_{AA}(\tau{=}0) = S_{AA}(\tau{=}0) = \big\langle\hat{A}^2\big\rangle_{\text{eq.}}.$$

If $\hat{A}$ is centered, $\langle\hat{A}\rangle_{\text{eq.}}=0$, then $\langle\hat{A}^2\rangle_{\text{eq.}}$ is the variance of repeated measurements of the expectation value of $\hat{A}$, i.e. it is a measure of the *fluctuations* of the value taken by this observable.

When $\tau\neq 0$ or when $\hat{B}$ and $\hat{A}$ differ, $C_{BA}(\tau)$ and $S_{BA}(\tau)$ are no longer variances, but rather *covariances*—to be accurate, this holds if both $\hat{A}$ and $\hat{B}$ are centered—, which measure the degree of correlation between the two observables (see Appendix B.4.1).

VI.3 Properties and interrelations of the correlation functions

The various time-correlation functions χ_{BA}, C_{BA}, S_{BA}, ξ_{BA}, K_{BA} introduced in Sec. VI.1, as well as their respective Fourier transforms, are obviously not independent from each other. Accordingly, one finds in the literature results expressed in terms of different functions, which however turn out to be totally equivalent.

In this section, we investigate some of the mathematical relations between different correlators, starting with those between the real and imaginary parts of the generalized susceptibility (Sec. VI.3.1), which arise due to the causality of the retarded propagator. We then list in Secs. VI.3.2 and VI.3.3 a few properties of the correlations functions either in t- or in ω-representation, and we gather the most important identities connecting them with each other.

Some of the relations of these two sections are explored at further length in Sec. VI.3.4, which discusses the general formulation of the fluctuation–dissipation theorem, and Sec. VI.3.5, in which

---

[(78)] or rather, to be precise, the product $\beta K_{BA}(t)$, which he denotes by $\Phi_{BA}(t)$.

we prove the Onsager reciprocal relations. Eventually, we derive in Sec. VI.3.6 identities between integrals involving either the spectral density or the Fourier transform of the canonical correlation function and equilibrium expectation values of appropriate commutators.

VI.3.1 Causality and dispersion relations

The causality property of the retarded propagator, encoded in the $\Theta(\tau)$-factor, leads to important integral relations between the real and imaginary of the generalized susceptibility. We now derive these identities,(79) which hold irrespective of the exact functional form of $\tilde{\chi}_{BA}(\omega)$.

VI.3.1 a Analytic continuation of the generalized susceptibility

It is convenient to extend the definition of the (complex-valued) susceptibility $\tilde{\chi}_{BA}$ to complex values of the variable ω. Consider thus the integral

$$\int_{-\infty}^{\infty} \chi_{BA}(\tau)\, \mathrm{e}^{\mathrm{i}z\tau}\, \mathrm{d}\tau$$

for a complex number $z = x + \mathrm{i}y$, which naturally extends definition (VI.10a). Given the $\Theta(\tau)$ function in $\chi_{BA}(\tau)$, this integral actually reads

$$\int_{0}^{\infty} \chi_{BA}(\tau)\, \mathrm{e}^{\mathrm{i}z\tau}\, \mathrm{d}\tau = \int_{0}^{\infty} \chi_{BA}(\tau)\, \mathrm{e}^{\mathrm{i}x\tau}\, \mathrm{e}^{-y\tau}\, \mathrm{d}\tau.$$

If the imaginary part y of z is non-negative, this integral is convergent and thus defines an analytic function in the upper complex half-plane

$$\widehat{\chi}_{BA}(z) \equiv \int_{-\infty}^{\infty} \chi_{BA}(\tau)\, \mathrm{e}^{\mathrm{i}z\tau}\, \mathrm{d}\tau \quad \text{for } \operatorname{Im} z > 0. \tag{VI.39}$$

The generalized susceptibility is the value of this function at the lower boundary:

$$\tilde{\chi}_{BA}(\omega) = \lim_{\varepsilon \to 0^+} \widehat{\chi}_{BA}(\omega + \mathrm{i}\varepsilon), \tag{VI.40}$$

cf. Eq. (VI.10a).

Spectral representations of $\widehat{\chi}_{BA}(z)$

Inserting the alternative expression (VI.27) of the response function in the definition of $\widehat{\chi}_{BA}(z)$, one finds

$$\widehat{\chi}_{BA}(z) = \frac{1}{\hbar} \sum_{n,n'} (\pi_n - \pi_{n'}) B_{nn'} A_{n'n} \frac{1}{\omega_{n'n} - z}. \tag{VI.41}$$

While definition (VI.39) only holds for $z > 0$, this spectral representation allows analytical continuations to the rest of the complex plane. Relation (VI.41) then shows that the singularities of $\widehat{\chi}_{BA}(z)$ are poles (of order one), situated on the real axis at each Bohr frequency $\omega_{n'n}$, with residues which can directly be read off the above expression.

Let $\tilde{\chi}'_{BA}(\omega)$ and $\tilde{\chi}''_{BA}(\omega)$ denote the real and imaginary parts of the generalized susceptibility $\tilde{\chi}_{BA}(\omega)$. Since the response function $\chi_{BA}(\tau)$ is real-valued (see last remark of § VI.1.2 a), $\tilde{\chi}'_{BA}(\omega)$ resp. $\tilde{\chi}''_{BA}(\omega)$ represents its cosine resp. sine Fourier transform. That is, $\tilde{\chi}'_{BA}(\omega)$ is actually the Fourier transform of the even part $\frac{1}{2}[\chi_{BA}(\tau) + \chi_{BA}(-\tau)]$, while $\tilde{\chi}''_{BA}(\omega)$ is the Fourier transform of $\frac{1}{2\mathrm{i}}[\chi_{BA}(\tau) - \chi_{BA}(-\tau)]$.

Since $\chi_{BA}(\tau)$ vanishes for $\tau < 0$, it can be rewritten as

$$\chi_{BA}(\tau) = [\chi_{BA}(\tau) - \chi_{BA}(-\tau)]\, \Theta(\tau) = 2\mathrm{i}\Theta(\tau) \frac{1}{2\mathrm{i}} [\chi_{BA}(\tau) - \chi_{BA}(-\tau)].$$

---

(79) The proof given below differs from the "traditional" one, as found e.g. in Ref. [51, Sec. 7.10.C & D], which relies on the integral along a properly chosen contour including parts of the real axis of $\widehat{\chi}_{BA}(z')/(z'-z)$, assuming (often unexpressedly) the analyticity of $\widehat{\chi}_{BA}$ in the whole upper complex half-plane, including the real axis.

Inserting this expression in Eq. (VI.39), one obtains for $\operatorname{Im} z > 0$

$$\widehat{\chi}_{BA}(z) = 2\mathrm{i}\int_0^\infty \mathrm{e}^{\mathrm{i}z\tau} \int_{-\infty}^{\infty} \tilde{\chi}''_{BA}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega\tau}\, \frac{\mathrm{d}\omega}{2\pi}\, \mathrm{d}\tau.$$

The integral over τ is readily performed and yields the spectral representation of the analytic continuation

$$\widehat{\chi}_{BA}(z) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\chi}''_{BA}(\omega)}{\omega - z}\, \mathrm{d}\omega \quad \text{for } \operatorname{Im} z > 0 \tag{VI.42a}$$

as function of the imaginary part of the susceptibility.

Similarly, starting from the identity $\chi_{BA}(\tau) = [\chi_{BA}(\tau)+\chi_{BA}(-\tau)]\,\Theta(\tau)$ one finds the alternative representation

$$\widehat{\chi}_{BA}(z) = \frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \frac{\tilde{\chi}'_{BA}(\omega)}{\omega - z}\, \mathrm{d}\omega \quad \text{for } \operatorname{Im} z > 0, \tag{VI.42b}$$

this time as function of the real part of the generalized susceptibility.

VI.3.1 b Causality and dispersion relations

The spectral decompositions (VI.42) can be exploited to derive *dispersion relations* which relate the real and imaginary parts of the generalized susceptibility to each other.

Renaming the dummy integration variable ω' and setting $z = \omega + \mathrm{i}\varepsilon$ with $\operatorname{Im} z = \varepsilon > 0$, these decompositions read

$$\widehat{\chi}_{BA}(\omega+\mathrm{i}\varepsilon) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\chi}''_{BA}(\omega')}{\omega' - \omega - \mathrm{i}\varepsilon}\, \mathrm{d}\omega', \qquad \widehat{\chi}_{BA}(\omega+\mathrm{i}\varepsilon) = \frac{1}{\mathrm{i}\pi}\int_{-\infty}^{\infty} \frac{\tilde{\chi}'_{BA}(\omega')}{\omega' - \omega - \mathrm{i}\varepsilon}\, \mathrm{d}\omega'.$$

In the limit $\varepsilon \to 0^+$, the left-hand side of either of these identities tends according to Eq. (VI.40) towards $\chi_{BA}(\omega) = \chi'_{BA}(\omega) + \mathrm{i}\chi''_{BA}(\omega)$. Transforming the right-hand sides with relation (A.2b), one obtains

$$\boxed{\begin{aligned} \operatorname{Re} \tilde{\chi}_{BA}(\omega) &= \frac{1}{\pi}\mathscr{P}\int_{-\infty}^{\infty} \frac{\operatorname{Im} \tilde{\chi}_{BA}(\omega')}{\omega' - \omega}\, \mathrm{d}\omega', \\ \operatorname{Im} \tilde{\chi}_{BA}(\omega) &= -\frac{1}{\pi}\mathscr{P}\int_{-\infty}^{\infty} \frac{\operatorname{Re} \tilde{\chi}_{BA}(\omega')}{\omega' - \omega}\, \mathrm{d}\omega', \end{aligned}} \tag{VI.43}$$

where $\mathscr{P}$ denotes the Cauchy principal value(80) of the integral.

These identities, known as *Kramers–Kronig*(bu) *relations* (or *Plemelj*(bv) *formulas*(81)), are a consequence of the causality of linear response. They allow the reconstruction of the whole susceptibility from either its real or its imaginary part.

Remark: Causality provides the analyticity of $\widehat{\chi}_{BA}(z)$ in the upper complex half-plane—not including the (whole) real axis, since the poles lie there, at the Bohr frequencies. Yet it tells nothing on the behavior at large $|z|$, which motivates a posteriori footnote 76 above. In particular, $\widehat{\chi}_{BA}(z)$ could converge towards a non-zero constant value χ_∞, in which case the linear response function $\chi_{BA}(\tau)$ contains an instantaneous contribution $\chi_\infty\delta(\tau)$—which shows that χ_∞ is necessarily real. The Kramers–Kronig relations (VI.43) are then invalid. One can however perform appropriate "subtractions" (cf. Ref. [52], Chapter 1.7) so as to recover new dispersion relations.(82)

---

(80) Cf. Appendix A for a quick reminder.

(81) To be accurate, the latter relate two analytic functions inside and outside of a closed contour in the complex plane—the Kramers–Kronig relations corresponding to the case where the contour consists of the real axis and a contribution at infinity.

(82) In the simplest case, one only needs to subtract χ_∞ from the real part of $\tilde{\chi}_{BA}(\omega)$ in the left member of the first resp. the right member of the second of relations (VI.43).

(bu) R. Kronig, 1904–1995 (bv) J. Plemelj, 1873–1967

VI.3.2 Properties and relations of the time-correlation functions

VI.3.2 a Properties of the time-correlation functions

We now list a few properties of the various time-correlation functions, without providing their respective proofs, starting with the symmetric and canonical correlation functions.(83)

- $S_{BA}(0) = S_{AB}(0) \quad , \quad K_{BA}(0) = K_{AB}(0);$ (VI.44a)
- $S_{BA}(\tau) = S_{AB}(-\tau) \quad , \quad K_{BA}(\tau) = K_{AB}(-\tau);$ (VI.44b)

  in particular $S_{AA}(\tau) = S_{AA}(-\tau) \quad , \quad K_{AA}(\tau) = K_{AA}(-\tau),$ (VI.44c)

that is, S_{AA} and K_{AA} are even functions.
Note that similar properties do not hold for C_{BA} when $\hat{B} \neq \hat{A}$. However, one still has

$$C_{AA}(\tau) = C_{AA}(-\tau). \tag{VI.44d}$$

Considering now complex conjugation, one finds [cf. Eq. (VI.15)]

- $S_{BA}(\tau)^* = S_{A^\dagger B^\dagger}(-\tau) = S_{B^\dagger A^\dagger}(\tau) \quad , \quad K_{BA}(\tau)^* = K_{A^\dagger B^\dagger}(-\tau) = K_{B^\dagger A^\dagger}(\tau);$ (VI.45a)
- if $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, $S_{BA}(\tau)$ and $K_{BA}(\tau)$ are real numbers; (VI.45b)

  in particular for $\hat{B} = \hat{A}^\dagger = \hat{A}$, $S_{AA}(0)$ and $K_{AA}(0)$ are positive real numbers. (VI.45c)

The latter property for Hermitian operators $\hat{A}$ also holds for $C_{AA}(0)$.
Given the antisymmetrization in the definition of $\xi_{BA}(\tau)$, the corresponding properties differ:

- $\xi_{BA}(0) = \xi_{AB}(0) = 0;$ (VI.46a)
- $\xi_{BA}(\tau) = -\xi_{AB}(-\tau);$ (VI.46b)

  in particular ξ_{AA} is odd: $\xi_{AA}(\tau) = -\xi_{AA}(-\tau),$ (VI.46c)

Turning to complex conjugation, one finds

- $\xi_{BA}(\tau)^* = \xi_{A^\dagger B^\dagger}(-\tau) = -\xi_{B^\dagger A^\dagger}(\tau);$ (VI.47a)
- if $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, $\xi_{BA}(\tau)$ is purely imaginary; . (VI.47b)

We shall come back to property (VI.46b) in Sec. VI.3.5, in which we shall take into account the specific behaviour of the operators $\hat{A}$, $\hat{B}$ under time reversal.

VI.3.2 b Interrelations between time-correlation functions

The explicit expression (VI.26) of the generalized susceptibility shows that it is simply related to the inverse Fourier transform (VI.19) of the spectral density according to

$$\boxed{\chi_{BA}(\tau) = 2\mathrm{i}\Theta(\tau)\,\xi_{BA}(\tau).} \tag{VI.48}$$

Since $\chi_{BA}(\tau)$, which was defined for Hermitian operators only, is real-valued (see last remark of Sec. VI.1.2 a), one recovers property VI.47b.

Let us define an operator $\hat{\dot{A}}$ by the relation

$$\hat{\dot{A}} \equiv \frac{1}{\mathrm{i}\hbar}\big[\hat{A}, \hat{H}_0\big], \tag{VI.49}$$

i.e. such that its matrix elements are given by $(\dot{A})_{nn'} = (E_{n'} - E_n)A_{nn'}/\mathrm{i}\hbar = \mathrm{i}\omega_{nn'}A_{nn'}$. If $\hat{A}$ is an

---

(83) The properties involving C_{BA}, S_{BA} or ξ_{BA} can be read at once from their definitions (VI.12) resp. (VI.16), or invoking their respective spectral representations. Those pertaining to K_{BA} can be shown with the help of the decomposition (VI.23), in particular using the invariance of the ratio under the exchange $n \leftrightarrow n'$.

observable, then $\dot{\hat{A}}$ coincides with the value taken at $t = 0$ by the derivative $\mathrm{d}\hat{A}_{\mathrm{I}}(t)/\mathrm{d}t$ for a system evolving with $\hat{H}_0$ only, i.e. in the absence of external perturbation.

Replacing $\hat{A}$ by $\dot{\hat{A}}$ in the spectral form (VI.23) of Kubo's correlation function, one finds

$$K_{B\dot{A}}(\tau) = \mathrm{i}\sum_{n,n'} \frac{\pi_n - \pi_{n'}}{\beta\hbar} B_{nn'} A_{n'n} \mathrm{e}^{-\mathrm{i}\omega_{n'n}\tau},$$

i.e.

$$K_{B\dot{A}}(\tau) = \frac{2\mathrm{i}}{\beta}\xi_{BA}(\tau). \tag{VI.50}$$

In turn, relation (VI.48), becomes

$$\boxed{\chi_{BA}(\tau) = \beta\Theta(t) K_{B\dot{A}}(\tau).} \tag{VI.51}$$

This relation is sometimes referred to as *Kubo formula*, since in his original article [50] Kubo expressed the linear response to an perturbation with the help of $\beta K_{B\dot{A}}(\tau)$ instead of the retarded propagator $\chi_{BA}(\tau)$ used in Sec. VI.1.2.

Identifying the right-hand sides of Eqs. (VI.38) and (VI.37) and differentiating the resulting relation with respect to time, one finds

$$\frac{\mathrm{d}K_{BA}(t)}{\mathrm{d}t} = -\frac{2\mathrm{i}}{\beta}\xi_{BA}(t).$$

Equation (VI.50) then yields

$$\frac{\mathrm{d}K_{BA}(t)}{\mathrm{d}t} = -K_{B\dot{A}}(t). \tag{VI.52}$$

VI.3.3 Properties and relations in frequency space

VI.3.3 a Detailed balance relation and properties of the spectral density

Recalling the spectral decomposition (VI.14) of the Fourier transform of the non-symmetrized correlation function

$$\tilde{C}_{BA}(\omega) = \sum_{n,n'} \pi_n\, B_{nn'} A_{n'n}\, \delta(\omega_{n'n} - \omega).$$

one sees that the exchange of the dummy indices n and n' and the relation $\pi_n/\pi_{n'} = \mathrm{e}^{-\beta\hbar\omega_{nn'}}$ yield, under consideration of the constraint imposed by the δ-term, the *detailed balance relation*

$$\boxed{\tilde{C}_{BA}(-\omega) = \tilde{C}_{AB}(\omega)\, \mathrm{e}^{-\beta\hbar\omega}.} \tag{VI.53}$$

This relation is a generic property of systems in canonical equilibrium.

The two obvious limits of this relation can be readily discussed. For $\hbar\omega \ll k_BT$, i.e. in the "classical regime", one finds the symmetric (in particular when $\hat{B} = \hat{A}$) relation $\tilde{C}_{BA}(-\omega) \simeq \tilde{C}_{AB}(\omega)$. On the other hand, the asymmetry—which reflects the difference between the probabilities for the absorption or emission of energy by the system—becomes large in the "quantum limit" $\hbar\omega \gg k_BT$, and in particular for T vanishingly small, in which case $\tilde{C}_{BA}(-\omega) \simeq 0$ for negative frequencies.

Either by Fourier transforming the identities (VI.46b) and (VI.47a) or by invoking directly the definition (VI.20), one finds that the spectral density obeys the properties

- $\tilde{\xi}_{BA}(\omega) = -\tilde{\xi}_{AB}(-\omega)$; (VI.54a)
- $\tilde{\xi}_{BA}(\omega)^* = \tilde{\xi}_{A^\dagger B^\dagger}(\omega) = -\tilde{\xi}_{B^\dagger A^\dagger}(-\omega)$; (VI.54b)
- if $\hat{A} = \hat{A}^\dagger$ and $\hat{B} = \hat{B}^\dagger$, $\tilde{\xi}_{BA}(\omega)^* = -\tilde{\xi}_{BA}(-\omega) = \tilde{\xi}_{AB}(\omega)$. (VI.54c)

As we shall now see, the functions $\tilde{\chi}_{BA}(\omega)$, $\tilde{C}_{BA}(\omega)$, $\tilde{S}_{BA}(\omega)$ and $\tilde{K}_{BA}(\omega)$ can all be expressed in terms of the spectral density $\tilde{\xi}_{BA}(\omega)$. There follows relations similar to Eqs. (VI.54) for the other spectral representations, which we shall not list.

VI.3.3 b Relations between different correlation functions in frequency space

Relation between $\tilde{\chi}_{BA}(\omega)$ and $\tilde{\xi}_{BA}(\omega)$

Using the decomposition (VI.20) of the spectral density, relation (VI.41) can be rewritten as

$$\widehat{\chi}_{BA}(z) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega)}{\omega - z}\,\mathrm{d}\omega. \tag{VI.55}$$

This identity constitutes yet another spectral representation of $\widehat{\chi}_{BA}(z)$, valid in the whole complex plane.

Renaming the dummy integration variable ω', setting $z = \omega + \mathrm{i}\varepsilon$ and taking the limit $\varepsilon \to 0^+$ under consideration of Eq. (VI.40), one naturally recovers Eq. (VI.29)

$$\tilde{\chi}_{BA}(\omega) = \frac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty} \frac{\tilde{\xi}_{BA}(\omega')}{\omega' - \omega - \mathrm{i}\varepsilon}\,\mathrm{d}\omega'. \tag{VI.29}$$

Equation (VI.55) can be further exploited to yield another relation between $\tilde{\chi}_{BA}(\omega)$ and $\tilde{\xi}_{BA}(\omega)$. Writing the principal value of $1/(\omega' - \omega)$ [cf. Eq. (A.2b)] in two different ways and subtracting them, one finds

$$\lim_{\varepsilon\to 0^+}\left[\widehat{\chi}_{BA}(\omega + \mathrm{i}\varepsilon) - \widehat{\chi}_{BA}(\omega - \mathrm{i}\varepsilon)\right] = 2\mathrm{i}\tilde{\xi}_{BA}(\omega), \tag{VI.56}$$

i.e. the difference between the values of $\widehat{\chi}_{BA}(z)$ in the upper and lower complex half-planes on each side of the point $\omega \in \mathbb{R}$ is proportional to $\tilde{\xi}_{BA}(\omega)$. Along a portion of the real axis where $\tilde{\xi}_{BA}(\omega)$ is continuous—which might happen in a system in the thermodynamic limit, when the Bohr frequencies span a continuous spectrum—and non-vanishing, $\widehat{\chi}_{BA}(z)$ thus has a cut.

The first term in Eq. (VI.56) is given by Eq. (VI.40). For the value in the lower half-plane, using Eq. (VI.41) gives

$$\widehat{\chi}_{BA}(\omega - \mathrm{i}\varepsilon) = \frac{1}{\hbar}\sum_{n,n'}(\pi_n - \pi_{n'})B_{nn'}A_{n'n}\,\frac{1}{\omega_{n'n} - \omega + \mathrm{i}\varepsilon} = \frac{1}{\hbar}\sum_{n,n'}\left[(\pi_n - \pi_{n'})B^*_{nn'}A^*_{n'n}\,\frac{1}{\omega_{n'n} - \omega - \mathrm{i}\varepsilon}\right]^*.$$

Recognizing in $A^*_{n'n}$, $B^*_{n'n}$ the matrix elements of $\hat{A}^\dagger$, $\hat{B}^\dagger$ in the basis $\{\,|\phi_n\rangle\}$, the rightmost term can be rewritten as $[\widehat{\chi}_{A^\dagger B^\dagger}(\omega + \mathrm{i}\varepsilon)]^*$. Invoking Eq. (VI.40) again, one finds

$$\lim_{\varepsilon\to 0^+}\widehat{\chi}_{BA}(\omega - \mathrm{i}\varepsilon) = \lim_{\varepsilon\to 0^+}[\widehat{\chi}_{A^\dagger B^\dagger}(\omega + \mathrm{i}\varepsilon)]^* = [\tilde{\chi}_{A^\dagger B^\dagger}(\omega)]^*.$$

All in all, Eq. (VI.56) thus becomes

$$\tilde{\xi}_{BA}(\omega) = \frac{1}{2\mathrm{i}}\left[\tilde{\chi}_{BA}(\omega) - \tilde{\chi}_{A^\dagger B^\dagger}(\omega)^*\right]. \tag{VI.57}$$

Since the susceptibilities $\tilde{\chi}_{BA}(\omega)$ and $\tilde{\chi}_{A^\dagger B^\dagger}(\omega)$ are in general not equal, even if $\hat{A}$ and $\hat{B}$ are Hermitian, $\tilde{\xi}_{BA}(\omega)$ will differ from the imaginary part of $\tilde{\chi}_{BA}(\omega)$.

In the specific case $\hat{B} = \hat{A}^\dagger$, Eq. (VI.57) shows that the spectral function is the imaginary part of the generalized susceptibility

$$\boxed{\tilde{\xi}_{A^\dagger A}(\omega) = \operatorname{Im}\tilde{\chi}_{A^\dagger A}(\omega).} \tag{VI.58}$$

As we have seen in Sec. VI.2.2, the spectral density characterizes energy dissipation in the system. As a consequence, the imaginary part of the susceptibility $\tilde{\chi}_{BA}(\omega)$ is often referred to as “dissipative part”, even if $\hat{B} \neq \hat{A}^\dagger$.(84)

Remark: The relation between $\operatorname{Im}\tilde{\chi}_{A^\dagger A}(\omega)$ and dissipation can be recovered by the following heuristic argument. Viewing $\hat{A}(t)$ as the “generalized displacement” conjugate to the force $f(t)$ in the Hamiltonian (in the Heisenberg picture with respect to $\hat{H}_0$), then the power dissipated by the

---

(84) This denomination can actually be dangerous if $\hat{A}$ and $\hat{B}$ behave differently under time reversal, see the second remark at the end of Sec. VI.3.5.

system is the product of the force with the “velocity”, namely

$$\frac{\mathrm{d}E_{\text{tot.}}}{\mathrm{d}t} = f(t)\left\langle \frac{\mathrm{d}\hat{A}(t)}{\mathrm{d}t} \right\rangle_{\text{n.eq.}}.$$

Assuming a harmonic force $f(t) = f_\omega \hat{A}\cos\omega t = f_\omega \operatorname{Re}\big(\hat{A}\mathrm{e}^{-\mathrm{i}\omega t}\big)$, the linear response of $\hat{A} = \hat{A}^\dagger$ is given by $\langle \hat{A}^\dagger\rangle_{\text{n.eq.}} = f_\omega\big[\operatorname{Re}\tilde{\chi}_{A^\dagger A}(\omega)\cos\omega t + \operatorname{Im}\tilde{\chi}_{A^\dagger A}(\omega)\sin\omega t\big]$. Differentiating with respect to t yields at once the instantaneous power, which after averaging over one period of the force yields for the mean rate of energy dissipation

$$\overline{\frac{\mathrm{d}E_{\text{tot.}}}{\mathrm{d}t}} = \frac{f_\omega^2}{2}\,\omega \operatorname{Im}\tilde{\chi}_{A^\dagger A}(\omega),$$

which is of course equivalent to Eq. (VI.31).

Relation between $\tilde{C}_{BA}(\omega)$ and $\tilde{\xi}_{BA}(\omega)$

Comparing the spectral decomposition (VI.14) of the Fourier transform of the non-symmetrized correlation function with the spectral density (VI.20), one sees that the only change is the replacement of $2\pi_n$ by $(\pi_n - \pi_{n'})/\hbar$.

The specific form (VI.2b) of the canonical equilibrium populations leads to the identity

$$\pi_n = (\pi_n - \pi_{n'})\frac{\pi_n}{\pi_n - \pi_{n'}} = (\pi_n - \pi_{n'})\frac{1}{1-\mathrm{e}^{-\beta(E_{n'}-E_n)}} = (\pi_n - \pi_{n'})\frac{1}{1-\mathrm{e}^{-\beta\hbar\omega_{n'n}}}.$$

As the term $\delta(\omega_{n'n} - \omega)$ in Eq. (VI.14) or (VI.20) imposes $\omega_{n'n} = \omega$ in the exponent, one finds

$$\boxed{\tilde{C}_{BA}(\omega) = \frac{2\hbar}{1-\mathrm{e}^{-\beta\hbar\omega}}\,\tilde{\xi}_{BA}(\omega).} \tag{VI.59}$$

Relation between $\tilde{K}_{BA}(\omega)$ and $\tilde{\xi}_{BA}(\omega)$

Consider the spectral representation (VI.22) of the Fourier transform of Kubo’s canonical correlation function. The $\delta(\omega - \omega_{n'n})$-term allows us to replace the Bohr frequencies in the denominator by ω. Comparison with the decomposition (VI.20) of the spectral density then yields at once the identity

$$\boxed{\tilde{K}_{BA}(\omega) = \frac{2}{\beta}\frac{\tilde{\xi}_{BA}(\omega)}{\omega}.} \tag{VI.60}$$

Relation between $\tilde{S}_{BA}(\omega)$ and $\tilde{\xi}_{BA}(\omega)$

As was done above for $\tilde{C}_{BA}(\omega)$, one sees that the spectral decomposition of the Fourier transform of the symmetric correlation function and the spectral function (VI.20) only differ in that the latter involves the difference $\pi_n - \pi_{n'}$ of the populations of different energy eigenstates, while the former involves their sum. Invoking again the form (VI.2b) of the equilibrium populations, one obtains the identity

$$\pi_n + \pi_{n'} = (\pi_n - \pi_{n'})\frac{\pi_n + \pi_{n'}}{\pi_n - \pi_{n'}} = (\pi_n - \pi_{n'})\frac{1+\mathrm{e}^{-\beta\hbar\omega_{n'n}}}{1-\mathrm{e}^{-\beta\hbar\omega_{n'n}}} = (\pi_n - \pi_{n'})\coth\frac{\beta\hbar\omega_{n'n}}{2}.$$

As before, $\omega_{n'n}$ is set to ω by the term $\delta(\omega_{n'n} - \omega)$, so that the argument of the hyperbolic cotangent in the rightmost member is actually independent of n and n'. Equation (VI.18) then yields

$$\boxed{\tilde{S}_{BA}(\omega) = \hbar\coth\frac{\beta\hbar\omega}{2}\,\tilde{\xi}_{BA}(\omega).} \tag{VI.61}$$

For $|\beta\hbar\omega| \ll 1$, one has $\coth(\frac{1}{2}\beta\hbar\omega) \sim 2/\beta\hbar\omega$. One thus finds with the help of relation (VI.60)

$$\tilde{S}_{BA}(\omega) \sim \frac{2}{\beta\omega}\,\tilde{\xi}_{BA}(\omega) = \tilde{K}_{BA}(\omega).$$

That is, $\tilde{S}_{BA}(\omega)$ and $\tilde{K}_{BA}(\omega)$ tend towards each other in the classical limit $\hbar \to 0$

VI.3.3 c Recapitulation of the various correlation functions

Let us summarize the main results we have found above for the various correlation functions we have introduced, indicating the physical phenomenon in which they naturally appear, as well as various relations between them.

	t-space	ω-space
Spectral function (dissipation)	$\xi_{BA}(t) = \dfrac{1}{2\hbar}\Big\langle \big[\hat{B}_{\rm I}(t),\hat{A}\big]\Big\rangle_{\rm eq.}$ $= \dfrac{\beta}{2{\rm i}}\,K_{B\dot{A}}(t)$	$\tilde{\xi}_{BA}(\omega) = \dfrac{\pi}{\hbar}\sum_{n,n'}(\pi_n - \pi_{n'}) \times$ $\times\, B_{nn'}A_{n'n}\,\delta(\omega_{n'n}-\omega)$
Response function / *susceptibility*	$\chi_{BA}(t) = 2{\rm i}\Theta(t)\xi_{BA}(t)$ $= \beta\Theta(t)K_{B\dot{A}}(t)$	$\tilde{\chi}_{BA}(\omega) = \dfrac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\dfrac{\tilde{\xi}_{BA}(\omega')}{\omega'-\omega-{\rm i}\varepsilon}\,{\rm d}\omega'$
Symmetric correlation *function* (fluctuations)	$S_{BA}(t) = \dfrac{1}{2}\Big\langle \big\{\hat{B}_{\rm I}(t),\hat{A}\big\}_+\Big\rangle_{\rm eq.}$	$\tilde{S}_{BA}(\omega) = \hbar\coth\dfrac{\beta\hbar\omega}{2}\,\tilde{\xi}_{BA}(\omega)$
Canonical correlation *function* (relaxation for $t \geq 0$)	$K_{BA}(t) =$ $\dfrac{1}{\beta}\int_0^{\beta}\Big\langle {\rm e}^{\lambda\hat{H}_0}\hat{A}\,{\rm e}^{-\lambda\hat{H}_0}\hat{B}_{\rm I}(t)\,{\rm d}\lambda\Big\rangle_{\rm eq.}$	$\tilde{K}_{BA}(\omega) = \dfrac{2}{\beta\omega}\,\tilde{\xi}_{BA}(\omega)$

Table VI.1 – Summary of the various correlation functions in linear response theory.

VI.3.4 Fluctuation–dissipation theorem

Equations (VI.59), (VI.60), and (VI.61) relate the Fourier transforms of the non-symmetrized, canonical and symmetric correlation functions to the spectral function $\tilde{\xi}_{BA}(\omega)$. We now discuss the physical content of these relations and present an example of application.

VI.3.4 a First fluctuation–dissipation theorem

Consider first the special case $\hat{B} = \hat{A}$, with $\hat{A}$ an observable, thus Hermitian. Together with Eq. (VI.58), one has the series of identities

$$\operatorname{Im}\tilde{\chi}_{AA}(\omega) = \tilde{\xi}_{AA}(\omega) = \frac{\beta\omega}{2}\tilde{K}_{AA}(\omega) = \frac{1-{\rm e}^{\beta\hbar\omega}}{2\hbar}\tilde{C}_{AA}(\omega) = \frac{\tanh\frac{\beta\hbar\omega}{2}}{\hbar}\tilde{S}_{AA}(\omega). \tag{VI.62}$$

The two leftmost terms are related to dissipation in the system when it is excited by a perturbation coupling to $\hat{A}$ (Sec. VI.2.2). That is, they represent (part of) the dynamical response of the system when it is driven out of equilibrium by an external constraint. Meanwhile, the two rightmost terms encode the temporal (auto)correlation and spontaneous fluctuations of $\hat{A}$ in the system at thermodynamic equilibrium. These two pairs of correlation functions thus model a priori different physical phenomena: their interrelation expressed by Eq. (VI.62) is thus non-trivial, and constitutes the so-called *fluctuation–dissipation theorem*.

Traditionally, the denomination *fluctuation–dissipation theorem* is rather attached to relations in which the Fourier transform of the correlation function which stands for fluctuations is explicitly written as a time integral; for instance,

$$\operatorname{Im}\tilde{\chi}_{AA}(\omega) = \frac{1}{2k_BT}\int_{-\infty}^{\infty}\omega K_{AA}(t)\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t, \tag{VI.63}$$

where β has been replaced by its explicit expression in terms of the temperature, or

$$\operatorname{Im}\tilde{\chi}_{AA}(\omega) = \frac{\tanh\frac{\beta\hbar\omega}{2}}{\hbar}\int_{-\infty}^{\infty} S_{AA}(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t. \tag{VI.64}$$

More generally, for an arbitrary pair of observables $\hat{A}$, $\hat{B}$ one can simply Fourier transform the Kubo formula [Eq. (VI.51)]

$$\chi_{BA}(t) = \frac{1}{k_BT}\,\Theta(t)K_{B\dot{A}}(t),$$

which gives

$$\boxed{\tilde{\chi}_{BA}(\omega) = \frac{1}{k_BT}\int_0^{\infty} K_{B\dot{A}}(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t.} \tag{VI.65}$$

Following Kubo [53], this identity is referred to as first fluctuation–dissipation theorem.(85)

Remarks:

∗ Kubo's canonical function can be associated with a mechanical reaction of the system when it is perturbed (Sec. VI.2.3). Yet the third term of Eq. (VI.62) also becomes identical to the two rightmost ones in the classical limit, in which case it is rather related to the equilibrium dynamics of the fluctuations of $\hat{A}$. Accordingly, in relation (VI.65) the "fluctuation" part of the theorem is played by the canonical correlation function.

∗ As explained in the second remark at the end of Sec. VI.3.5, which part, real of imaginary, of the susceptibility is dissipative depends on the time-reversal signatures of the two observables $\hat{A}$ and $\hat{B}$. In practice, one often considers Eq. (VI.65) with $\hat{B} = \dot{\hat{A}}$, so that $\hat{A}$ and $\hat{B}$ have opposite time-reversal parities, in which case the dissipative part of $\tilde{\chi}_{BA}(\omega)$ is the real part, as e.g. in the example of next paragraph.

VI.3.4 b Application: Johnson–Nyquist noise

Consider an arbitrary passive electrical circuit,(86) which can either be closed on itself or submitted to a voltage $V_{\text{ext.}}(t)$, in thermodynamic equilibrium at temperature T. Let $I(t)$ denote the electric current through the circuit. In the absence of external voltage, $I(t)$ vanishes at equilibrium.

Assume first that the circuit is submitted to $V_{\text{ext.}}(t)$. The average electric current $\langle I(t)\rangle_{\text{n.eq.}}$ in the circuit can be computed within the (classical) theory of linear response, where the angular brackets denote the result of a "typical" measurement, as obtained by averaging over many repeated measurements so as to minimize the uncertainty of a single observation.

The external voltage $V_{\text{ext.}}(t)$ couples to the electric charge Q traversing the circuit, which thus plays the role of the excited (classical) observable A. In turn, the responding observable B is here the electric current $I(t)$. Going to frequency space, the response of $\langle I(t)\rangle_{\text{n.eq.}}$ to the excitation $V_{\text{ext.}}(t)$ is governed by the generalized admittance $\tilde{\chi}_{IQ}(\omega)$, which is simply the inverse of the *electrical impedance* $Z(\omega)$ of the circuit:

$$\big\langle I(t)\big\rangle_{\text{n.eq.}} = \tilde{\chi}_{IQ}(\omega)\tilde{V}_{\text{ext.}}(\omega) \quad\text{with}\quad \tilde{\chi}_{IQ}(\omega) = \frac{1}{Z(\omega)}.$$

Consider now the situation when the circuit is closed on itself, i.e. $V_{\text{ext.}} = 0$. The circuit is in an equilibrium state, and the electrical current $I(t)$ fluctuates around its average value $\langle I(t)\rangle_{\text{eq.}} = 0$.

The fluctuation–dissipation theorem (VI.65) relates $\tilde{\chi}_{IQ}(\omega)$ to the fluctuations of the electrical current. Since $\dot{Q} = I$, one thus has under consideration of the classical limit (VI.130) of the canonical correlation function

$$\tilde{\chi}_{IQ}(\omega) = \frac{1}{k_BT}\int_0^{\infty}\big\langle I(t)I(0)\big\rangle_{\text{eq.}}\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t. \tag{VI.66}$$

---

(85) … or fluctuation–dissipation theorem of the first kind.

(86) … consisting of *linear* elements only: resistors, inductors, capacitors and memristors.

Since $\langle I(t)I(0)\rangle_{\rm eq.}$ is a real and even function of t, the complex conjugate of the right member can be expressed as the integral of the same integrand between $-\infty$ and 0. Taking the half sum of Eq. (VI.66) and its complex conjugate thus yields

$$\operatorname{Re}\frac{1}{Z(\omega)} = \frac{R(\omega)}{|Z(\omega)|^2} = \frac{1}{2k_BT}\int_{-\infty}^{\infty} \big\langle I(t)I(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t, \tag{VI.67}$$

where $R(\omega)$ denotes the real part—the *resistive part*—of the impedance $Z(\omega)$. This is a relation between the resistance and impedance on the one hand and the fluctuations of the current in the electrical circuit on the other side. Performing the inverse Fourier transform and setting $t = 0$, one finds

$$\big\langle I^2\big\rangle_{\rm eq.} = \frac{2k_BT}{\pi}\int_0^{\infty} \frac{R(\omega)}{|Z(\omega)|^2}\,{\rm d}\omega, \tag{VI.68}$$

where the evenness of $R(\omega)/|Z(\omega)|^2$, which can be read directly off Eq. (VI.67), has been used. These thermal fluctuations of the electrical current were first measured by Johnson(bw) [54] and computed by Nyquist(bx) [55], which is why they are referred to as *Johnson–Nyquist noise*.

Let $V(t)$ denote the fictitious fluctuating voltage which, if applied to the circuit, would give rise to the same fluctuating current $I(t)$. One can show that the Fourier transforms of their autocorrelation functions are related to each other through

$$\int_{-\infty}^{\infty} \big\langle I(t)I(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t = \frac{1}{|Z(\omega)|^2}\int_{-\infty}^{\infty} \big\langle V(t)V(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t. \tag{VI.69}$$

Comparing this relation with Eq. (VI.67), one finds

$$\frac{k_BT}{\pi}R(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \big\langle V(t)V(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t, \tag{VI.70}$$

which constitutes the *Nyquist theorem* relating the real part of the circuit impedance to the Fourier transform of the time-autocorrelation function of the voltage fluctuations at thermodynamic equilibrium.

> Proof of Eq. (VI.69):
> The operation leading from $V(t)$ to $I(t)$ is an instance of *linear filter*, i.e. an operation relating an "input" $y_{\rm in.}(t)$ and an "output" $y_{\rm out.}(t)$ such that a) the output depends linearly on the input; b) the filter properties are independent of time; and c) the output cannot predate the input (causality). $y_{\rm out.}(t)$ is then expressed in function of $y_{\rm in.}(t-t')$ by a convolution over time, as in relation (VI.8), which in frequency space becomes a simple multiplication
>
> $$\tilde{y}_{\rm out.}(\omega) = G(\omega)\tilde{y}_{\rm in.}(\omega),$$
>
> with $G(\omega)$ the *transfer function* of the filter. Here, $G(\omega)$ is the admittance $1/Z(\omega)$.
>
> If $y_{\rm in.}(t)$ and $y_{\rm out.}(t)$ are now fluctuating quantities that can be viewed as stationary stochastic processes, their spectral functions are respectively proportional to $|\tilde{y}_{\rm in.}(\omega)|^2$ and $|\tilde{y}_{\rm out.}(\omega)|^2$
>
> $$S_{\rm out.}(\omega) = |G(\omega)|^2 S_{\rm in.}(\omega).$$
>
> According to the Wiener–Khinchin theorem (C.46), these spectral functions are the Fourier transforms of the respective autocorrelation functions, that is
>
> $$\int_{-\infty}^{\infty} \big\langle y_{\rm out.}(t)y_{\rm out.}(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t = |G(\omega)|^2\int_{-\infty}^{\infty} \big\langle y_{\rm in.}(t)y_{\rm in.}(0)\big\rangle_{\rm eq.}\,{\rm e}^{{\rm i}\omega t}\,{\rm d}t,$$
>
> which with $y_{\rm in.}(t) = V(t)$, $y_{\rm out.}(t) = I(t)$, $G(\omega) = 1/Z(\omega)$ proves Eq. (VI.69). □

Remark: One may "guess" that k_BT in the numerator on the right-hand side of the Nyquist relation (VI.70) is actually the classical limit $k_BT \gg \hbar\omega$ of $\frac{1}{2}\hbar\omega\coth(\hbar\omega/2k_BT)$, which appears for

---

(bw) J. B. Johnson, 1887–1970 (bx) H. Nyquist, 1889–1976 instance on the right-hand side of Eq. (VI.61).(87) That is, relation (VI.70) would be the high-temperature limit of

$$\frac{\hbar\omega}{2\pi}\coth\frac{\hbar\omega}{2k_BT}\,R(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\big\langle V(t)V(0)\big\rangle_{\text{eq.}}\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t.$$

The inverse Fourier transform of this identity reads

$$\big\langle V(t)V(0)\big\rangle_{\text{eq.}}=\int_{-\infty}^{\infty}\hbar\omega\coth\frac{\hbar\omega}{2k_BT}\,R(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\frac{\mathrm{d}\omega}{2\pi}=\frac{1}{\pi}\int_{-\infty}^{\infty}\bigg(\frac{1}{\mathrm{e}^{\hbar\omega/k_BT}-1}+\frac{1}{2}\bigg)\hbar\omega R(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega.$$

Setting $t=0$, one recovers the "generalized Nyquist relation"

$$\big\langle V^2\big\rangle_{\text{eq.}}=\frac{2}{\pi}\int_0^{\infty}\bigg(\frac{1}{\mathrm{e}^{\hbar\omega/k_BT}-1}+\frac{1}{2}\bigg)\hbar\omega R(\omega)\,\mathrm{d}\omega \tag{VI.71}$$

which was the first quantum-mechanical instance of fluctuation–dissipation relation, as derived by Callen and Welton(by) [56].

VI.3.5 Onsager relations

Using the symmetries of a problem often allows one to deduce interesting relations as well as simplifications. We discuss here a first example, in the case of symmetry under time reversal. A further example will be given illustrated on an explicit example in Sec. VI.4.3, when discussing quantum Brownian motion.

Equation (VI.46b) relates ξ_{BA}, i.e. the response of $\hat{B}$ to a excitation coupled to $\hat{A}$, to ξ_{AB}, which describes the "reciprocal" situation of the shift of the expectation value of $\hat{A}$ caused by a perturbation coupling to $\hat{B}$. More precisely, it is a relation between $\xi_{BA}(t)$ and $\xi_{AB}(-t)$, that is with reversed time direction, which is slightly unsatisfactory.

To obtain an equation relating $\xi_{BA}(t)$ and $\xi_{AB}(t)$, with the same time direction in both correlation functions, one needs to introduce the time reversal operator $\hat{\mathscr{T}}$ and to discuss the behavior of the various observables under its operation.

VI.3.5 a Time reversal in quantum mechanics

Accordingly, let us briefly recall some properties of the operator $\hat{\mathscr{T}}$ which represents the action of the time-reversal operation on spinless particles.(88) These follow from the fact that $\hat{\mathscr{T}}$ is an *antiunitary* operator, i.e. an antilinear operator whose adjoint equals its inverse.

Let $\hat{\mathcal{A}}$ denote an antilinear operator. If $|1\rangle$, $|2\rangle$ are two kets of the Hilbert space $\mathscr{H}$ on which $\hat{\mathcal{A}}$ is acting, and λ_1, λ_2 two complex constants, one has

$$\hat{\mathcal{A}}\big(\lambda_1|1\rangle+\lambda_2|2\rangle\big)=\lambda_1^*\hat{\mathcal{A}}\,|1\rangle+\lambda_2^*\hat{\mathcal{A}}\,|2\rangle. \tag{VI.72a}$$

That is, if $\lambda\in\mathbb{C}$

$$\hat{\mathcal{A}}\lambda=\lambda^*\hat{\mathcal{A}}. \tag{VI.72b}$$

If $\langle\phi|$ is a bra (of the dual space to $\mathscr{H}$), the action of $\hat{\mathcal{A}}$ on $\langle\phi|$ defines a new bra $\langle\phi|\,\hat{\mathcal{A}}$ such that for any ket $|\psi\rangle$, one has the identity

$$\big(\langle\phi|\,\hat{\mathcal{A}}\big)|\psi\rangle=\Big[\langle\phi|\big(\hat{\mathcal{A}}\,|\psi\rangle\big)\Big]^*. \tag{VI.72c}$$

Note that the brackets cannot be dropped, contrary to the case of linear operators: one must specify whether $\hat{\mathcal{A}}$ acts on the ket or on the bra.

---

(87) This educated guess is motivated by the fact that $\frac{1}{2}\hbar\omega\coth(\hbar\omega/2k_BT)$ is actually the average energy at temperature T of the harmonic oscillator with frequency ω.

(88) For further details, see e.g. Messiah [29] Chapter 15, in particular Secs. 3–5 & 15–22.

(by) T. A. Welton, 1918–2010

The adjoint operator $\hat{\mathcal{A}}^\dagger$ of the antilinear operator $\hat{\mathcal{A}}$ is such that for all $|\phi\rangle$, $\hat{\mathcal{A}}^\dagger|\phi\rangle$ is the ket conjugate to the bra $\langle\phi|\hat{\mathcal{A}}$. For all $|\phi\rangle$, $|\psi\rangle$, the usual property of the scalar product reads

$$\langle\psi|\big(\hat{\mathcal{A}}^\dagger|\phi\rangle\big) = \Big[\big(\langle\phi|\hat{\mathcal{A}}\big)|\psi\rangle\Big]^*.$$

Comparing with Eq. (VI.72c), this gives the relation

$$\langle\psi|\big(\hat{\mathcal{A}}^\dagger|\phi\rangle\big) = \langle\phi|\big(\hat{\mathcal{A}}\,|\psi\rangle\big). \tag{VI.72d}$$

Eventually, the antiunitarity of $\hat{\mathscr{T}}$ means that it obeys the additional identity

$$\hat{\mathscr{T}}\hat{\mathscr{T}}^\dagger = \hat{\mathscr{T}}^\dagger\hat{\mathscr{T}} = \hat{\mathbb{1}}. \tag{VI.72e}$$

It follows from this relation that if vectors $\{\,|\phi_n\rangle\}$ form an orthonormal basis, then so do the transformed vectors $\{\hat{\mathscr{T}}|\phi_n\rangle\}$, which shall be denoted as $\{\,|\hat{\mathscr{T}}\phi_n\rangle\}$.

Signature of an operator under time reversal

Many operators $\hat{O}$ acting on a system transform in a simple way under time reversal, namely according to

$$\hat{\mathscr{T}}\hat{O}\hat{\mathscr{T}}^\dagger = \epsilon_O\hat{O} \quad \text{with } \epsilon_O = +1 \text{ or } -1. \tag{VI.73}$$

ϵ_O is the *signature* of the operator $\hat{O}$ under time reversal.
For instance, time reversal does not modify positions, but changes velocities

$$\hat{\mathscr{T}}\hat{\vec{X}}\hat{\mathscr{T}}^\dagger = \hat{\vec{X}}, \qquad \hat{\mathscr{T}}\hat{\vec{V}}\hat{\mathscr{T}}^\dagger = -\hat{\vec{V}}, \tag{VI.74}$$

that is $\epsilon_X = 1$, $\epsilon_V = -1$.

Consider now a system with Hamilton operator $\hat{H}$, assumed to be invariant under time reversal, i.e. $\hat{\mathscr{T}}\hat{H}\hat{\mathscr{T}}^\dagger = \hat{H}$. Let $\hat{O}_{\rm H}(t)$ denote the Heisenberg representation (II.36) of an observable $\hat{O}$; one may then write

$$\hat{\mathscr{T}}\,\hat{O}_{\rm H}(t)\hat{\mathscr{T}}^\dagger = \hat{\mathscr{T}}\,{\rm e}^{{\rm i}\hat{H}t/\hbar}\,\hat{O}\,{\rm e}^{-{\rm i}\hat{H}t/\hbar}\,\hat{\mathscr{T}}^\dagger = \hat{\mathscr{T}}\,{\rm e}^{{\rm i}\hat{H}t/\hbar}\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}\,\hat{O}\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}\,{\rm e}^{-{\rm i}\hat{H}t/\hbar}\hat{\mathscr{T}}^\dagger,$$

where the (anti)unitarity (VI.72e) was used. Invoking now Eqs. (VI.72b) and (VI.73), and again the antiunitarity property yields the relation

$$\hat{\mathscr{T}}\,\hat{O}_{\rm H}(t)\hat{\mathscr{T}}^\dagger = {\rm e}^{-{\rm i}\hat{H}t/\hbar}\,\epsilon_O\hat{O}\,{\rm e}^{{\rm i}\hat{H}t/\hbar} = \epsilon_O\hat{O}_{\rm H}(-t), \tag{VI.75}$$

which shows that the time reversal operator acts on operators as it is supposed to, inversing the direction of time evolution.

Remark: In the presence of an external magnetic field $\vec{\mathscr{B}}_{\rm ext.}$, the Hamiltonian $\hat{H}$ is not invariant under time reversal. As a matter of fact, the magnetic field couples to operators of the system with signature -1, as for instance the velocity of particles or their spins, so that the transformed of $\hat{H}$ under time reversal corresponds to the Hamiltonian of the same system in the opposite magnetic field $-\vec{\mathscr{B}}_{\rm ext.}$

$$\hat{\mathscr{T}}\,\hat{H}\big[\vec{\mathscr{B}}_{\rm ext.}\big]\hat{\mathscr{T}}^\dagger = \hat{H}\big[-\vec{\mathscr{B}}_{\rm ext.}\big]. \tag{VI.76}$$

VI.3.5 b Behavior of correlation functions under time reversal

Let us come back to the generic setup of this Chapter, namely to a system with unperturbed Hamiltonian $\hat{H}_0$. We assume that the latter is invariant under time reversal, so that $\hat{H}_0$ and $\hat{\mathscr{T}}$ commute. As a consequence, the canonical equilibrium density operator $\hat{\rho}_{\rm eq.}$ also commutes with the time reversal operator.

Considering operators $\hat{A}$ and $\hat{B}$ with definite signatures under time reversal and their respective Heisenberg representations (VI.3) with respect to $\hat{H}_0$, we can compute the equilibrium expectation value $\big\langle\hat{B}_{\rm I}(t)\hat{A}\big\rangle_{\rm eq.}$:

$${\rm Tr}\big[\hat{\rho}_{\rm eq.}\hat{B}_{\rm I}(t)\hat{A}\big] = \sum_n\langle\phi_n|\hat{\rho}_{\rm eq.}\hat{B}_{\rm I}(t)\hat{A}|\phi_n\rangle = \sum_n\langle\phi_n|\Big(\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}\hat{\rho}_{\rm eq.}\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}\hat{B}_{\rm I}(t)\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}\hat{A}\hat{\mathscr{T}}^\dagger\hat{\mathscr{T}}|\phi_n\rangle\Big).$$

Using the identities $\hat{\mathscr{T}}\hat{\rho}_{\text{eq.}}\hat{\mathscr{T}}^\dagger = \hat{\rho}_{\text{eq.}}$, $\hat{\mathscr{T}}\hat{B}_{\text{I}}(t)\hat{\mathscr{T}}^\dagger = \epsilon_B\hat{B}_{\text{I}}(-t)$ and $\hat{\mathscr{T}}\hat{A}\hat{\mathscr{T}}^\dagger = \epsilon_A\hat{A}$ in the right-most member of the equation, this becomes

$$\begin{aligned}\big\langle\hat{B}_{\text{I}}(t)\hat{A}\big\rangle_{\text{eq.}} &= \epsilon_A\epsilon_B\sum_n\langle\phi_n|\Big(\hat{\mathscr{T}}^\dagger\hat{\rho}_{\text{eq.}}\hat{B}_{\text{I}}(-t)\hat{A}|\hat{\mathscr{T}}\phi_n\rangle\Big)\\ &= \epsilon_A\epsilon_B\sum_n\langle\hat{\mathscr{T}}\phi_n|\hat{A}^\dagger\hat{B}_{\text{I}}(-t)^\dagger\hat{\rho}_{\text{eq.}}\Big(\hat{\mathscr{T}}|\phi_n\rangle\Big),\end{aligned}$$

where the second identity follows from applying property (VI.72d) with $|\phi\rangle = \hat{\rho}_{\text{eq.}}\hat{B}_{\text{I}}(-t)\hat{A}|\hat{\mathscr{T}}\phi_n\rangle$, $\hat{\mathcal{A}} = \hat{\mathscr{T}}$, and $\langle\psi| = \langle\phi_n|$, while using the hermiticity of $\hat{\rho}_{\text{eq.}}$. The term $\big(\hat{\mathscr{T}}|\phi_n\rangle\big)$ can then be rewritten as $|\hat{\mathscr{T}}\phi_n\rangle$. Since the vectors $\{\hat{\mathscr{T}}|\phi_n\rangle\}$ form an orthogonal basis, the sum on the right-hand side actually represents a trace

$$\big\langle\hat{B}_{\text{I}}(t)\hat{A}\big\rangle_{\text{eq.}} = \epsilon_A\epsilon_B\,\text{Tr}\big[\hat{A}^\dagger\hat{B}_{\text{I}}(-t)^\dagger\hat{\rho}_{\text{eq.}}\big] = \epsilon_A\epsilon_B\big\langle\hat{A}^\dagger\hat{B}_{\text{I}}(-t)^\dagger\big\rangle_{\text{eq.}}.$$

If both $\hat{A}$ and $\hat{B}$ are Hermitian and using stationarity, this gives

$$\big\langle\hat{B}_{\text{I}}(t)\hat{A}\big\rangle_{\text{eq.}} = \epsilon_A\epsilon_B\big\langle\hat{A}_{\text{I}}(t)\hat{B}\,\big\rangle_{\text{eq.}}.$$

One similarly shows $\langle\hat{A}\hat{B}_{\text{I}}(t)\rangle_{\text{eq.}} = \epsilon_A\epsilon_B\langle\hat{B}\hat{A}_{\text{I}}(t)\rangle_{\text{eq.}}$, which leads to the *reciprocity relations*

$$\xi_{BA}(t) = \epsilon_A\epsilon_B\,\xi_{AB}(t), \tag{VI.77a}$$

$$\boxed{\tilde{\xi}_{BA}(\omega) = \epsilon_A\epsilon_B\,\tilde{\xi}_{AB}(\omega).} \tag{VI.77b}$$

This result constitutes the generalization of the Onsager reciprocal relations introduced in § I.2.2 b, which are the zero-frequency limit of the second relation here (see also § VI.4.1 a).

When the system is in an external magnetic field $\vec{\mathscr{B}}_{\text{ext.}}$, one shows with the help of Eq. (VI.76) that relation (VI.77b) generalizes to

$$\tilde{\xi}_{BA}(\omega,\vec{\mathscr{B}}_{\text{ext.}}) = \epsilon_A\epsilon_B\,\tilde{\xi}_{AB}(\omega,-\vec{\mathscr{B}}_{\text{ext.}}). \tag{VI.78}$$

Starting from this relation or, in the absence of magnetic field, from Eq. (VI.77b), one easily derives similar relations for the other spectral representations $\tilde{\chi}_{BA}(\omega)$, $\tilde{S}_{BA}(\omega)$, $\tilde{K}_{BA}(\omega)$.

Remarks:

∗ Combined with relation (VI.54b), Eq. (VI.77) shows that the spectral function $\tilde{\xi}_{BA}(\omega)$ is real and odd when $\epsilon_A\epsilon_B = 1$, or purely imaginary and even if $\epsilon_A\epsilon_B = -1$.

∗ Using relation (VI.48) between the linear response function and the Fourier transform of the spectral function, one deduces from Eq. (VI.77) the identities

$$\chi_{BA}(t) = \epsilon_A\epsilon_B\,\chi_{AB}(t) \quad , \quad \tilde{\chi}_{BA}(\omega) = \epsilon_A\epsilon_B\,\tilde{\chi}_{AB}(\omega). \tag{VI.79}$$

Accordingly, relation (VI.57) becomes

$$\tilde{\xi}_{BA}(\omega) = \frac{1}{2\text{i}}\big[\tilde{\chi}_{BA}(\omega) - \epsilon_A\epsilon_B\,\tilde{\chi}_{BA}(\omega)^*\big].$$

If observables $\hat{A}$ and $\hat{B}$ have the same parity under time reversal, then $\tilde{\xi}_{BA}(\omega)$ is the imaginary part of $\tilde{\chi}_{BA}(\omega)$, as in the case $\hat{B} = \hat{A}^\dagger$ [Eq. (VI.58)]. On the other hand, if they have opposite parities, then the above identity reads $\tilde{\xi}_{BA}(\omega) = -\text{i}\,\text{Re}\,\tilde{\chi}_{BA}(\omega)$: the dissipative part of the susceptibility is now its real part.

Accordingly, the rather standard notation $\chi''_{BA}(\tau)$ for the function called in these notes $\xi_{BA}(\tau)$ can be misleading in a twofold way: firstly, despite the double-primed notation is *not* the imaginary part of the retarded propagator $\chi_{BA}(\tau)$ even though $\tilde{\chi}''_{BA}(\omega)$ is that of $\tilde{\chi}_{BA}(\omega)$. Secondly, $\chi''_{BA}(\tau)$ is the inverse Fourier transform of $\tilde{\chi}''_{BA}(\omega)$ only if $\hat{A}$ and $\hat{B}$ behave similarly under time reversal.

VI.3.6 Sum rules

Consider definition (VI.19) with $\tau = t - t'$. Rewriting the right-hand side with the help of the stationarity property and expressing $\xi_{BA}(\tau)$ as inverse Fourier transform of the spectral density, one obtains

$$\int_{-\infty}^{\infty} \tilde{\xi}_{BA}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega(t-t')}\, \frac{\mathrm{d}\omega}{2\pi} = \frac{1}{2\hbar} \Big\langle \big[\hat{B}_{\mathrm{I}}(t), \hat{A}_{\mathrm{I}}(t')\big] \Big\rangle_{\mathrm{eq.}}. \tag{VI.80}$$

Let us differentiate this identity k times with respect to t and l times with respect to t':

$$(-\mathrm{i})^{k-l} \int_{-\infty}^{\infty} \omega^{k+l}\, \tilde{\xi}_{BA}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega(t-t')}\, \frac{\mathrm{d}\omega}{2\pi} = \frac{1}{2\hbar} \left\langle \left[\frac{\mathrm{d}^k \hat{B}_{\mathrm{I}}(t)}{\mathrm{d}t^k}, \frac{\mathrm{d}^l \hat{A}_{\mathrm{I}}(t')}{\mathrm{d}t'^l} \right] \right\rangle_{\mathrm{eq.}}.$$

Given Eq. (VI.3), each successive differentiation on the right-hand side gives rise to a commutator (with $\hat{H}_0$) divided by $\mathrm{i}\hbar$. This leads to k nested commutators in the left member of the commutator, and l nested commutators in its right member. Setting then $t' = t = 0$, one finds

$$\frac{(-1)^l}{\pi} \int_{-\infty}^{\infty} \omega^{k+l}\, \tilde{\xi}_{BA}(\omega)\, \mathrm{d}\omega = \frac{1}{\hbar^{l+k+1}} \left\langle \Big[\underbrace{\Big[\cdots \Big[[\hat{B}, \hat{H}_0], \hat{H}_0 \Big] \cdots \Big]}_{k \text{ commutators}}, \underbrace{\Big[\cdots \Big[[\hat{A}, \hat{H}_0], \hat{H}_0 \Big] \cdots \Big]}_{l \text{ commutators}} \Big] \right\rangle_{\mathrm{eq.}}. \tag{VI.81}$$

The term on the left-hand side of this identity is, up to the prefactor, the moment of order $k + l$ of the spectral function $\tilde{\xi}_{BA}(\omega)$. The larger $k + l$ is, the more sensitive the moment becomes to large values of ω, that is, to the short-time behavior of the inverse Fourier transform $\xi_{BA}(t)$.

The *sum rules* (VI.81) for the various values of k, l express the moments of the spectral function in terms of equilibrium expectation values of commutators. If the latter can be computed, using commutation relations, then the sum rules represent conditions that theoretical models for the spectral function $\tilde{\xi}_{BA}(\omega)$ should satisfy.

According to Eq. (VI.50), the right-hand side of Eq. (VI.80) also equals $\beta K_{B\dot{A}}(t - t')/2\mathrm{i}$:

$$\frac{1}{2\hbar} \Big\langle \big[\hat{B}_{\mathrm{I}}(t), \hat{A}(t')\big] \Big\rangle_{\mathrm{eq.}} = \frac{1}{2\mathrm{i}} \int_0^{\beta} \Big\langle \dot{\hat{A}}(t' - \mathrm{i}\hbar\lambda)\, \hat{B}_{\mathrm{I}}(t) \Big\rangle_{\mathrm{eq.}}\, \mathrm{d}\lambda.$$

Differentiating as above k times with respect to t and l times with respect to t', and setting $t' = t$, one obtains the alternative sum rules

$$\frac{-1}{(\mathrm{i}\hbar)^{l+k+1}} \left\langle \Big[\underbrace{\Big[\cdots \Big[[\hat{B}, \hat{H}_0], \hat{H}_0 \Big] \cdots \Big]}_{k \text{ commutators}}, \underbrace{\Big[\cdots \Big[[\hat{A}, \hat{H}_0], \hat{H}_0 \Big] \cdots \Big]}_{l \text{ commutators}} \Big] \right\rangle_{\mathrm{eq.}} = \int_0^{\beta} \left\langle \frac{\mathrm{d}^{l+1} \hat{A}_{\mathrm{I}}(-\mathrm{i}\hbar\lambda)}{\mathrm{d}t^{l+1}}\, \frac{\mathrm{d}^k \hat{B}_{\mathrm{I}}(0)}{\mathrm{d}t^k} \right\rangle_{\mathrm{eq.}} \mathrm{d}\lambda. \tag{VI.82}$$

Up to a factor β^{-1}, the right-hand side of this identity is the canonical correlation function of the $(l + 1)$-th time derivative of $\hat{A}$ and the k-th derivative of $\hat{B}$, taken at $t = 0$.

Examples of applications of these sum rules will be given in § VI.4.3 c.

VI.4 Examples and applications

The presentation of this Chapter is missing.

VI.4.1 Green–Kubo relation

An important application of the formalism of linear response is the calculation of the transport coefficients which appear in the phenomenological constitutive relations of non-equilibrium thermodynamics.

VI.4.1 a Linear response revisited

Under consideration of relation (VI.48) or (VI.51) and dropping the nonlinear terms, the Kubo formula (VI.8) can be recast as either

$$\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} = \big\langle \hat{B}\,\big\rangle_{\rm eq.} + 2{\rm i}\int_0^\infty \xi_{BA}(\tau)\,f(t-\tau)\,{\rm d}\tau \tag{VI.83}$$

or equivalently

$$\boxed{\big\langle \hat{B}_{\rm I}(t)\big\rangle_{\rm n.eq.} = \big\langle \hat{B}\,\big\rangle_{\rm eq.} + \int_0^\infty \beta K_{B\dot{A}}(\tau)\,f(t-\tau)\,{\rm d}\tau,} \tag{VI.84}$$

where the latter is in fact the form originally given by Kubo.

Let us assume that the expectation value of $\hat{B}$ at equilibrium vanishes: for instance, $\hat{B}$ is a flux as introduced in Chapter I. To emphasize this interpretation, we denote the responding observable by $\hat{\mathscr{J}}_b$, instead of $\hat{B}$. Accordingly, to increase the similarity with Sec. I.2, we call the generalized force $\mathscr{F}_a(t)$ instead of $f(t)$, and we rename $\hat{A}_a$ the observable coupling to $\mathscr{F}_a(t)$.

Fourier transforming relations (VI.83) or (VI.84) leads to

$$\big\langle \tilde{\hat{\mathscr{J}}}_b(\omega)\big\rangle_{\rm n.eq.} = L_{ba}(\omega)\tilde{\mathscr{F}}_a(\omega), \tag{VI.85a}$$

where we also adopted a new notation for the susceptibility:

$$L_{ba}(\omega) \equiv \beta\int_0^\infty K_{\mathscr{J}_b\dot{A}_a}(\tau)\,{\rm e}^{{\rm i}\omega\tau}\,{\rm d}\tau = 2{\rm i}\int_0^\infty \xi_{\mathscr{J}_bA_a}(\tau)\,{\rm e}^{{\rm i}\omega\tau}\,{\rm d}\tau. \tag{VI.85b}$$

Summing over different generalized forces, whose effects simply add up at the linear approximation, the Fourier-transformed Kubo formula (VI.85a) is a straightforward generalization of Eq. (I.31) accounting for frequency dependent fluxes and affinities. The real novelty is that the “generalized kinetic coefficient” $L_{ba}(\omega)$ is no longer a phenomenological factor as in Sec. I.2: instead, there is now an explicit formula to compute it using time-correlation functions at equilibrium of the system, Eq. (VI.85b).

In Chapter I, the considered affinities and fluxes were implicitly quasi-static—the gradients of intensive thermodynamic quantities were time-independent—, which is the regime for which transport coefficients are defined. For instance, the electrical conductivity $\sigma_{\rm el.}$ is defined as the proportionality factor between a constant electric field and the ensuing direct current. Accordingly, the kinetic coefficients L_{ba} in the constitutive relation I.31 are actually the zero-frequency limits of the corresponding susceptibilities:

$$\boxed{L_{ba} = \lim_{\omega\to 0}\frac{1}{k_BT}\int_0^\infty K_{\mathscr{J}_b\dot{A}_a}(\tau)\,{\rm e}^{{\rm i}\omega\tau}\,{\rm d}\tau = \frac{1}{k_BT}\int_0^\infty K_{\mathscr{J}_b\dot{A}_a}(\tau)\,{\rm d}\tau.} \tag{VI.86}$$

This constitutes the general form of the *Green–Kubo-relation* [57, 58, 50].

VI.4.1 b Example: electrical conductivity

As example of application of the Green–Kubo-relation, consider a system made of electrically charged particles, with charges q_i, submitted to a classical external electric field $\vec{\mathscr{E}}(t)$, which plays the role of the generalized force. In the dipolar approximation, this field perturbs the Hamiltonian by coupling to the electric dipole moment of the system

$$\hat{W}(t) = -\vec{\mathscr{E}}(t)\cdot\hat{\vec{D}} \quad\text{where}\quad \hat{\vec{D}} \equiv \sum_i q_i \hat{\vec{r}}_i$$

with $\hat{\vec{r}}_i$ the position operator of the i-th particle.

Quite obviously, we are interested in the electric current due to this perturbation, namely

$$\hat{\vec{J}}_{\text{el.}}(t) \equiv \sum_i q_i \frac{\mathrm{d}\hat{\vec{r}}_i(t)}{\mathrm{d}t}.$$

In the notations of the present Chapter, $\hat{A}$ is the component of $\hat{\vec{D}}$ along the direction of $\vec{\mathscr{E}}$—let us for simplicity denote this component by $\hat{D}_z$—, while the role of $\hat{B}$ is played by any component $\hat{J}_{\text{el.},j}$ of $\hat{\vec{J}}_{\text{el.}}$. Additionally, one sees that $\hat{B} = \dot{\hat{A}}$ in the case $\hat{B} = \hat{J}_{\text{el.},z}$.

Focussing on the latter case, Kubo's linear response formula reads in Fourier space

$$\hat{J}_{\text{el.},z}(\omega) = \tilde{\chi}_{J_z D_z}(\omega)\mathscr{E}_z(\omega) \equiv \sigma_{zz}(\omega)\mathscr{E}_z(\omega),$$

where for obvious reasons we have introduced for the relevant generalized susceptibility the alternative notation $\sigma_{zz}(\omega)$. The zero-frequency limit of σ_{zz} is the electrical conductivity, which according to the Green–Kubo-relation (VI.86) is given by

$$\sigma_{\text{el.}} = \frac{1}{k_B T}\int_0^\infty K_{J_z J_z}(\tau)\,\mathrm{d}\tau,$$

i.e. by the integral of the time-autocorrelation function of the component of the electric current along the direction of the electric field.

Let us again emphasize that given a microscopic model of the system, the canonical correlation function can be calculated, which then allows one to *compute* the electrical conductivity, which in paragraph I.2.3 c was a mere phenomenological coefficient.

VI.4.2 Harmonic oscillator

To illustrate the various correlation functions introduced in Sec. VI.1 and their properties and relations (Sec. VI.3), consider a free one-dimensional harmonic oscillator in thermal equilibrium. The Hamilton operator is

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_0^2\hat{x}^2 = \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \tag{VI.87a}$$

with the usual commutation relations $[\hat{a},\hat{a}] = [\hat{a}^\dagger,\hat{a}^\dagger] = 0$, $[\hat{a},\hat{a}^\dagger] = 1$, while the relation between the position and momentum operators and the annihilation and creation operators reads

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega_0}}\big(\hat{a}+\hat{a}^\dagger\big), \quad \hat{p} = \frac{1}{\mathrm{i}}\sqrt{\frac{m\hbar\omega_0}{2}}\big(\hat{a}-\hat{a}^\dagger\big). \tag{VI.87b}$$

We consider an external perturbation coupling to the position, corresponding to an extra term $\hat{W}(t) = -f(t)\hat{x}$ in the Hamiltonian.

The computations in that case are quite straightforward, thanks to the fact that the interaction-picture representation of the operators are easily computed. Thus, the Heisenberg equation (II.37) with the Hamiltonian (VI.87a) yields

$$\frac{\mathrm{d}\hat{a}_{\text{I}}(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{a}_{\text{I}}(t),\hat{H}_0\big] = -\mathrm{i}\omega_0\,\hat{a}_{\text{I}}(t),$$

so that for instance

$$\hat{x}(t) = \sqrt{\frac{\hbar}{2m\omega_0}}\big(\hat{a}\,\mathrm{e}^{-\mathrm{i}\omega_0 t} + \hat{a}^\dagger\mathrm{e}^{\mathrm{i}\omega_0 t}\big). \tag{VI.88}$$

One can then compute all correlation functions and their Fourier transforms explicitly.

Non-symmetrized correlation function

Let us start with the non-symmetrized correlation function $C_{xx}(t)$. From definition (VI.12), one has

$$C_{xx}(t) = \frac{\hbar}{2m\omega_0}\Big\langle \big(\hat{a}\,\mathrm{e}^{-\mathrm{i}\omega_0 t} + \hat{a}^\dagger \mathrm{e}^{\mathrm{i}\omega_0 t}\big)\big(\hat{a} + \hat{a}^\dagger\big)\Big\rangle_{\text{eq.}}.$$

Now, the annihilation or creation operators couple states with different energies, so that their representations in a basis of energy eigenstate ordered by increasing energy are strictly upper or lower triangular matrices.(89) As a consequence, their squares $\hat{a}^2$ or $\hat{a}^{\dagger\,2}$ are also strictly triangular in the basis where $\hat{\rho}_{\text{eq.}}$ is diagonal, so that neither $\hat{\rho}_{\text{eq.}}\hat{a}^2$ nor $\hat{\rho}_{\text{eq.}}\hat{a}^{\dagger\,2}$ has nonvanishing diagonal elements: the corresponding terms do not contribute to $C_{xx}(t)$. There remains

$$C_{xx}(t) = \frac{\hbar}{2m\omega_0}\Big\langle \hat{a}\hat{a}^\dagger \mathrm{e}^{-\mathrm{i}\omega_0 t} + \hat{a}^\dagger\hat{a}\,\mathrm{e}^{\mathrm{i}\omega_0 t}\Big\rangle_{\text{eq.}} = \frac{\hbar}{2m\omega_0}\Big\langle \mathrm{e}^{-\mathrm{i}\omega_0 t} + \hat{a}^\dagger\hat{a}\big(\mathrm{e}^{-\mathrm{i}\omega_0 t} + \mathrm{e}^{\mathrm{i}\omega_0 t}\big)\Big\rangle_{\text{eq.}}.$$

One recognizes in the rightmost member the particle number operator $\hat{N} \equiv \hat{a}^\dagger\hat{a}$, whose expectation value at equilibrium is readily computed with the matrix elements (VI.2b) and the known energy levels $E_n = (n+\frac{1}{2})\hbar\omega_0$ of the free harmonic oscillator

$$\big\langle \hat{N}\big\rangle_{\text{eq.}} = \sum_{n=0}^{\infty} n\frac{\mathrm{e}^{-\beta E_n}}{Z(\beta)} = \frac{1}{\mathrm{e}^{\beta\hbar\omega_0} - 1} \equiv f^{(\mathrm{B})}(\hbar\omega_0),$$

which is the Bose–Einstein occupation factor. One then finds

$$C_{xx}(t) = \frac{\hbar}{m\omega_0}\bigg[f^{(\mathrm{B})}(\hbar\omega_0)\cos(\omega_0 t) + \frac{1}{2}\mathrm{e}^{-\mathrm{i}\omega_0 t}\bigg], \tag{VI.89a}$$

where the second, temperature-independent term comes from the ground state.

Remark: In the classical limit $\hbar \to 0$, one has $f^{(\mathrm{B})}(\hbar\omega_0) \sim k_B T/\hbar\omega_0$ and thus

$$C_{xx}(t) \underset{\hbar\to 0}{\sim} \frac{k_B T}{m\omega_0^2}\cos(\omega_0 t).$$

Setting $t = 0$, one recovers the equipartition theorem $\frac{1}{2}m\omega_0^2\langle x^2\rangle = \frac{1}{2}k_B T$.

Fourier transforming Eq. (VI.89a), one finds

$$\tilde{C}_{xx}(\omega) = \frac{\pi\hbar}{m\omega_0}\big[\big(f^{(\mathrm{B})}(\hbar\omega_0) + 1\big)\delta(\omega_0 - \omega) + f^{(\mathrm{B})}(\hbar\omega_0)\delta(\omega_0 + \omega)\big]. \tag{VI.89b}$$

The first term term describes the absorption of an excitation—and contains the contribution from the ground state—, while the second term stands for emission. Since $f^{(\mathrm{B})}(\hbar\omega_0)+1 = f^{(\mathrm{B})}(\hbar\omega_0)\,\mathrm{e}^{\beta\hbar\omega_0}$, one easily checks that this Fourier transform obeys the detailed balance relation (VI.53).

Spectral function

From the correlation function $C_{xx}(t)$, one quickly deduces $\xi_{xx}(t)$:

$$\xi_{xx}(t) = \frac{1}{2\hbar}\big[C_{xx}(t) - C_{xx}(-t)\big] = \frac{\sin(\omega_0 t)}{2\mathrm{i}m\omega_0}, \tag{VI.90a}$$

whose Fourier transform is trivial and yields the spectral function

$$\tilde{\xi}_{xx}(\omega) = \frac{\pi}{2m\omega_0}\big[\delta(\omega_0 - \omega) - \delta(\omega_0 + \omega)\big]. \tag{VI.90b}$$

Together with Eq. (VI.89b), this obeys relation (VI.59).

Retarded propagator

The linear response function then follows from relation (VI.48)

$$\chi_{xx}(t) = 2\mathrm{i}\Theta(t)\xi_{xx}(t) = \frac{\sin(\omega_0 t)}{m\omega_0}\,\Theta(t). \tag{VI.91a}$$

---

(89) That is, triangular with vanishing diagonal terms.

Equation (VI.10a) yields the corresponding generalized susceptibility

$$\tilde{\chi}_{xx}(\omega)=\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\chi_{xx}(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{e}^{-\varepsilon t}\,\mathrm{d}t=\frac{1}{2m\omega_0}\lim_{\varepsilon\to 0^+}\left(\frac{1}{\omega_0-\omega-\mathrm{i}\varepsilon}+\frac{1}{\omega_0+\omega+\mathrm{i}\varepsilon}\right). \tag{VI.91b}$$

Using relation (A.2b), one easily checks that the imaginary part of $\tilde{\chi}_{xx}(\omega)$ is exactly the spectral function (VI.90b), in agreement with Eq. (VI.58), while the real part is

$$\operatorname{Re}\tilde{\chi}_{xx}(\omega)=\frac{1}{2m\omega_0}\left(\mathcal{P}\frac{1}{\omega_0-\omega}+\mathcal{P}\frac{1}{\omega_0+\omega}\right), \tag{VI.91c}$$

as also given by the first Kramers–Kronig relation (VI.43).

Symmetric correlation function

The symmetric correlation function follows at once from $C_{xx}(t)$, Eq. (VI.89a)

$$S_{xx}(t)=\frac{1}{2}\big[C_{xx}(t)+C_{xx}(-t)\big]=\frac{\hbar}{m\omega_0}\left[f^{(\mathrm{B})}(\hbar\omega_0)+\frac{1}{2}\right]\cos(\omega_0 t).$$

Using the relation $f^{(\mathrm{B})}(\hbar\omega_0)+\frac{1}{2}=\frac{1}{2}\coth\frac{\beta\hbar\omega_0}{2}$, this gives

$$S_{xx}(t)=\frac{\hbar}{2m\omega_0}\coth\frac{\beta\hbar\omega_0}{2}\cos(\omega_0 t). \tag{VI.92a}$$

The Fourier transform is straightforward and reads

$$\tilde{S}_{xx}(\omega)=\frac{\pi\hbar}{2m\omega_0}\coth\frac{\beta\hbar\omega_0}{2}\big[\delta(\omega_0-\omega)+\delta(\omega_0+\omega)\big]. \tag{VI.92b}$$

Rewriting this under consideration of the δ-distributions as

$$\tilde{S}_{xx}(\omega)=\frac{\pi\hbar}{2m\omega_0}\coth\frac{\beta\hbar\omega}{2}\big[\delta(\omega_0-\omega)-\delta(\omega_0+\omega)\big].$$

one finds relation (VI.61) with the spectral function (VI.90b).

Canonical correlation function

Eventually, one can compute Kubo's canonical correlation function. A straightforward integration yields

$$K_{xx}(t)=\frac{1}{\beta}\int_0^{\beta}C_{xx}(t+\mathrm{i}\hbar\lambda)\,\mathrm{d}\lambda=\frac{\cos(\omega_0 t)}{\beta m\omega_0^2}. \tag{VI.93a}$$

In turn, the Fourier transform reads

$$\tilde{K}_{xx}(\omega)=\frac{\pi}{\beta m\omega_0^2}\big[\delta(\omega_0-\omega)+\delta(\omega_0+\omega)\big]. \tag{VI.93b}$$

Using as above the constraints imposed by the Dirac distributions to replace ω_0 in the denominator of the prefactor by ω, and accordingly the plus sign between the two δ-terms by a minus sign, one recovers relation (VI.60) with the spectral function (VI.90b).

Remark: The expressions of the linear response function $\chi_{xx}(t)$ and the canonical correlation function $K_{xx}(t)$ determined above, Eqs. (VI.91a) and (VI.93a), are both independent of $\hbar$. They thus coincide with the classical correlation functions, which one finds for the linear response of a classical harmonic oscillator. This can be traced back to the fact that the position and momentum operators obey (in their Heisenberg-picture representations with respect to $\hat{H}_0$) the same equations of motion as the position and momentum of a classical oscillator. Since these equations of motion are linear, they also govern the evolution of the expectation values, which thus have the same solutions in the classical and quantum-mechanical cases.

VI.4.3 Quantum Brownian motion

In this section, we investigate a system consisting of a "heavy" particle interacting with a bath of many lighter particles at thermodynamic equilibrium—which provides a quantum-mechanical analogue to the problem of Brownian motion. We first introduce in § VI.4.3 a a rather general Hamilton operator for such a system, and investigate in § VI.4.3 b some of the properties of the spectral functions relating the velocity and position of the heavy particle that follow from the symmetries of the Hamiltonian. We then exemplify how the general sum rules derived in Sec. VI.3.6 constrain the spectral function—and in particular show that Langevin dynamics based on Eq. (V.1) cannot constitute the classical limit of such a model (§ VI.4.3 c). Eventually we focus on the case of the Caldeira–Leggett Hamiltonian already introduced in Sec. V.3.3, considered now at the quantum-mechanical level, and show that it describes a "quantum dissipative system" governed by a generalized Langevin equation (§ VI.4.3 d).

VI.4.3 a Description of the system

Consider a heavy particle of mass M interacting with a "bath" of (many) identical light particles of mass m. For the sake of simplicity, we assume that the various particles only interact by pairs—be it for the heavy–light or light–light interactions—with potentials that depend only on the distance between the particles.

The Hamiltonian of the system thus reads

$$\hat{H}_0 = \frac{\hat{\vec{p}}^{\,2}}{2M} + \sum_j \frac{\hat{\vec{p}}_j^{\,2}}{2m} + \sum_{j\neq k} w\big(|\hat{\vec{r}}_j - \hat{\vec{r}}_k|\big) + \sum_j W\big(|\hat{\vec{r}} - \hat{\vec{r}}_j|\big), \tag{VI.94}$$

with $\hat{\vec{p}}$, $\hat{\vec{r}}$ the momentum and position of the heavy particle and $\hat{\vec{p}}_j$, $\hat{\vec{r}}_j$ those of the j-th particle of the bath, while w, W denote the interaction potentials for light–light and heavy–light pairs, respectively. In § VI.4.3 d, we shall consider the special case in which w and W are harmonic potentials.

Let $\hat{x}$, $\hat{y}$, $\hat{z}$ denote the components of $\hat{\vec{r}}$ in a given, fixed Cartesian coordinate system, and $\hat{p}_x$, $\hat{p}_y$, $\hat{p}_z$ the corresponding coordinates of momentum. Using definition (VI.49), one can define the operators $\hat{\dot{x}}$, $\hat{\dot{y}}\ldots$, which using the Hamiltonian (VI.94) and the usual commutation relations are given by

$$\hat{\dot{x}} \equiv \frac{1}{\mathrm{i}\hbar}\big[\hat{x}, \hat{H}_0\big] = \frac{\hat{p}_x}{M} = \hat{v}_x, \tag{VI.95}$$

and similarly in the other two space directions, where $\hat{v}_x$, $\hat{v}_y$, $\hat{v}_z$ are the components of the velocity of the heavy particle.

We investigate the effect of applying an external force, for instance along the x-direction, amounting to a perturbation of the Hamiltonian $\hat{W} = -f(t)\hat{x}$. The "output" consists of any component $\hat{v}_i$ of the heavy-particle velocity. Throughout this section, we shall omit the index I denoting the interaction-picture representation of observables, yet any time-dependent operator is to be understood as being in that representation.

VI.4.3 b Constraints from the symmetries of the system

The Hamilton operator (VI.94) is invariant under several transformations. From the invariance under spatial and temporal translations follow as usual the conservation of the total momentum and of energy. Yet in the spirit of Sec. VI.3.5 we also want to discuss the consequences of the invariance under several discrete symmetries for the time correlation functions of the system.

Consequence of spatial symmetries

We first want to show that correlation functions of the type $\xi_{v_y x}(t)$—describing the effect on $\hat{v}_y$ of a force acting along the x-direction—all identically vanish.

For that purpose, consider the reflection with respect to the xz-plane, which is described by a unitary operator $\hat{\mathscr{S}}_y$. This symmetry operation leaves the Hamiltonian $\hat{H}_0$, and thus the equilibrium density operator $\hat{\rho}_{\rm eq.}$, invariant:

$$\hat{\mathscr{S}}_y\hat{H}_0\hat{\mathscr{S}}_y^{\dagger} = \hat{H}_0 \quad , \quad \hat{\mathscr{S}}_y\hat{\rho}_{\rm eq.}\hat{\mathscr{S}}_y^{\dagger} = \hat{\rho}_{\rm eq.}. \tag{VI.96a}$$

The reflection also leaves $\hat{x}$ invariant, yet it transforms $\hat{v}_y$ into $-\hat{v}_y$:

$$\hat{\mathscr{S}}_y\,\hat{x}\,\hat{\mathscr{S}}_y^{\dagger} = \hat{x} \quad , \quad \hat{\mathscr{S}}_y\,\hat{v}_y\,\hat{\mathscr{S}}_y^{\dagger} = -\hat{v}_y. \tag{VI.96b}$$

Eventually, since the operator $\hat{\mathscr{S}}_y$ is unitary, it transforms the eigenstates $\{\,|\phi_n\rangle\}$ of $\hat{H}_0$ into a set of states $\{\,|\hat{\mathscr{S}}_y\phi_n\rangle \equiv \hat{\mathscr{S}}_y\,|\phi_n\rangle\}$ which form an orthonormal basis.

We can now compute the equilibrium expectation value $\langle\hat{v}_y(t)\hat{x}\rangle_{\rm eq.}$:

$$\begin{aligned}\mathrm{Tr}\big[\hat{\rho}_{\rm eq.}\hat{v}_y(t)\hat{x}\big] &= \sum_n \big\langle\phi_n\big|\hat{\rho}_{\rm eq.}\hat{v}_y(t)\hat{x}\big|\phi_n\big\rangle = \sum_n \big\langle\phi_n\big|\hat{\mathscr{S}}_y^{\dagger}\hat{\mathscr{S}}_y\hat{\rho}_{\rm eq.}\hat{\mathscr{S}}_y^{\dagger}\hat{\mathscr{S}}_y\hat{v}_y(t)\,\hat{\mathscr{S}}_y^{\dagger}\hat{\mathscr{S}}_y\hat{x}\hat{\mathscr{S}}_y^{\dagger}\hat{\mathscr{S}}_y\big|\phi_n\big\rangle \\ &= \sum_n \big\langle\hat{\mathscr{S}}_y\phi_n\big|\big(\hat{\rho}_{\rm eq.}\big)\big(-\hat{v}_y(t)\big)\big(\hat{x}\big)\big|\hat{\mathscr{S}}_y\phi_n\big\rangle = -\sum_n \big\langle\hat{\mathscr{S}}_y\phi_n\big|\hat{\rho}_{\rm eq.}\hat{v}_y(t)\hat{x}\big|\hat{\mathscr{S}}_y\phi_n\big\rangle.\end{aligned}$$

Since the $\{\,|\hat{\mathscr{S}}_y\phi_n\rangle\}$ constitute an orthonormal basis, the last term is exactly the opposite of the trace of $\hat{\rho}_{\rm eq.}\hat{v}_y(t)\hat{x}$, i.e. equals $-\langle\hat{v}_y(t)\hat{x}\rangle_{\rm eq.}$. That is,

$$\langle\hat{v}_y(t)\hat{x}\rangle_{\rm eq.} = -\,\langle\hat{v}_y(t)\hat{x}\rangle_{\rm eq.} = 0,$$

from which one at once deduces

$$\xi_{v_yx}(t) = -\xi_{v_yx}(t) = 0. \tag{VI.97}$$

One shows in a similar way that only ξ_{v_xx}, ξ_{v_yy} and ξ_{v_zz} are non-zero.

Using the invariance of the Hamiltonian (VI.94) under arbitrary rotations, one can show that these three correlation functions are equal. This allows one to replace the problem of Brownian motion in three dimensions by three identical one-dimensional problems, as was done from the start in Chapter V.

Invariance under time reversal

The Hamiltonian (VI.94) is clearly invariant under time reversal. Using Eq. (VI.74), the reciprocity relation (VI.77) then reads

$$\tilde{\xi}_{v_xx}(\omega) = -\tilde{\xi}_{xv_x}(\omega). \tag{VI.98}$$

Together with Eq. (VI.54c), this shows that $\tilde{\xi}_{v_xx}(\omega)$ is purely imaginary and even in ω.

Considering now relation (VI.79), one similarly finds

$$\tilde{\chi}_{v_xx}(\omega) = -\tilde{\chi}_{xv_x}(\omega).$$

Inserting this identity in Eq. (VI.57), one finds

$$\tilde{\xi}_{v_xx}(\omega) = \frac{1}{2\mathrm{i}}\big[\tilde{\chi}_{v_xx}(\omega) - \tilde{\chi}_{xv_x}(\omega)^*\big] = \frac{1}{2\mathrm{i}}\big[\tilde{\chi}_{v_xx}(\omega) + \tilde{\chi}_{v_xx}(\omega)^*\big] = -\mathrm{i}\,\mathrm{Re}\,\tilde{\chi}_{v_xx}(\omega). \tag{VI.99}$$

Eventually, one can relate the spectral function $\tilde{\xi}_{v_xx}(\omega)$ for velocity–position correlations to the spectral function $\tilde{\xi}_{xx}(\omega)$ associated with position–position correlations, by Fourier transforming the identity

$$\xi_{v_xx}(t) = \frac{1}{2\hbar}\bigg\langle\bigg[\frac{\mathrm{d}\hat{x}(t)}{\mathrm{d}t},\hat{x}\bigg]\bigg\rangle_{\rm eq.} = \frac{1}{2\hbar}\frac{\mathrm{d}}{\mathrm{d}t}\big\langle\big[\hat{x}(t),\hat{x}\big]\big\rangle_{\rm eq.} = \frac{\mathrm{d}\xi_{xx}(t)}{\mathrm{d}t},$$

which gives

$$\tilde{\xi}_{v_xx}(\omega) = -\mathrm{i}\omega\tilde{\xi}_{xx}(\omega), \tag{VI.100}$$

which shows that $\tilde{\xi}_{xx}(\omega)$ is real and odd.

VI.4.3 c Fluctuation–dissipation theorem and sum rules

Thanks to relation (VI.95), the first fluctuation–dissipation theorem (VI.65) with $\hat{A} = \hat{x}$ and $\hat{B} = \hat{v}_x$ reads

$$\tilde{\chi}_{v_x x}(\omega) = \frac{1}{k_B T} \int_0^\infty K_{v_x v_x}(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t. \tag{VI.101}$$

This identity represents the quantum-mechanical generalization of relation (V.40), which had been derived in a classical context, in which the Langevin equation is postulated phenomenologically, and in the equilibrium regime $t \gg \tau_\mathrm{c}$—which implicitly requires the separation of scales $\tau_\mathrm{c} \ll \tau_\mathrm{r}$.

Sum rules

To obtain a first sum rule, consider Eqs. (VI.81)–(VI.82) with $\hat{A} = \hat{x}$, $\hat{B} = \hat{v}_x$ and $k = l = 0$. This gives

$$\frac{1}{\hbar}\big\langle \big[\hat{v}_x, \hat{x}\big]\big\rangle_{\mathrm{eq.}} = \frac{1}{\pi}\int_{-\infty}^{\infty} \tilde{\xi}_{v_x x}(\omega)\, \mathrm{d}\omega = -\mathrm{i}\beta K_{v_x v_x}(0). \tag{VI.102a}$$

Using $\hat{v}_x = \hat{p}_x / M$ and the usual commutator $\big[\hat{x}, \hat{p}_x\big] = \mathrm{i}\hbar$, the left member of this equation is easily computed. The first identity then reads

$$\int_{-\infty}^{\infty} \tilde{\xi}_{v_x x}(\omega)\, \mathrm{d}\omega = -\frac{\mathrm{i}\pi}{M}, \tag{VI.102b}$$

which shows that that the integral of the spectral function is fixed. In turn, the identity between the first and third terms in the sum rule (VI.102a) gives

$$\frac{1}{2} M K_{v_x v_x}(0) = \frac{1}{2} k_B T, \tag{VI.102c}$$

which in the classical limit $K_{v_x v_x}(0) \underset{\hbar\to 0}{\to} \langle v_x^2 \rangle$ is the equipartition theorem.

Taking now $k = 0$, $l = 1$, the sum rules (VI.81)–(VI.82) read

$$\frac{1}{\hbar^2}\Big\langle \Big[\hat{v}_x, \big[\hat{x}, \hat{H}_0\big]\Big]\Big\rangle_{\mathrm{eq.}} = -\frac{1}{\pi}\int_{-\infty}^{\infty} \omega\, \tilde{\xi}_{v_x x}(\omega)\, \mathrm{d}\omega = \beta K_{\dot{v}_x v_x}(0), \tag{VI.103}$$

where in the rightmost term the canonical correlation function correlates the acceleration $\dot{\hat{v}}_x$ to the velocity $\hat{v}_x$.

The Hamiltonian (VI.94) yields the commutation relation $\big[\hat{x}, \hat{H}_0\big] = \mathrm{i}\hbar \hat{p}_x / M = \mathrm{i}\hbar \hat{v}_x$, so that the left member of Eq. (VI.103) equals zero. This means that the first moment of the spectral function vanishes—which was clear since $\tilde{\xi}_{v_x x}(\omega)$ is an even function. It also implies that the acceleration and the velocity at the same instant are not correlated, which is also a normal property for a stationary physical quantity and its time derivative.

Consider eventually the sum rule obtained when $k = l = 1$ and again $\hat{A} = \hat{x}$, $\hat{B} = \hat{v}_x$. In that case, Eqs. (VI.81) and (VI.82) yield

$$\frac{1}{\hbar^3}\Big\langle \Big[\big[\hat{v}_x, \hat{H}_0\big], \big[\hat{x}, \hat{H}_0\big]\Big]\Big\rangle_{\mathrm{eq.}} = -\frac{1}{\pi}\int_{-\infty}^{\infty} \omega^2\, \tilde{\xi}_{v_x x}(\omega)\, \mathrm{d}\omega = \mathrm{i}\beta K_{\dot{v}_x \dot{v}_x}(0). \tag{VI.104a}$$

That is, the sum rule involves the second moment of the spectral function $\tilde{\xi}_{v_x x}(\omega)$ and the canonical autocorrelation function of the acceleration $\dot{\hat{v}}_x$ computed at equal times.

To compute the commutator on the left-hand side, one can first replace $\hat{v}_x$ by $\hat{p}_x / M$ and use as above $\big[\hat{x}, \hat{H}_0\big] = \mathrm{i}\hbar \hat{p}_x / M$, which gives

$$\frac{1}{\hbar^3}\Big\langle \Big[\big[\hat{v}_x, \hat{H}_0\big], \big[\hat{x}, \hat{H}_0\big]\Big]\Big\rangle_{\mathrm{eq.}} = \frac{\mathrm{i}}{\hbar^2}\frac{1}{M^2}\Big\langle \Big[\big[\hat{p}_x, \hat{H}_0\big], \hat{P}_x\Big]\Big\rangle_{\mathrm{eq.}}.$$

Using then

$$\big[\hat{p}_x, \hat{H}_0\big] = -\mathrm{i}\hbar \frac{\partial \hat{H}_0}{\partial \hat{x}} = -\mathrm{i}\hbar \frac{\partial \mathscr{V}(\hat{\vec{r}})}{\partial \hat{x}} \quad \text{with } \mathscr{V}(\hat{\vec{r}}) \equiv \sum_j W\big(|\hat{\vec{r}} - \hat{\vec{r}}_j|\big),$$

this becomes

$$\frac{1}{\hbar^3}\Big\langle\Big[\big[\hat{v}_x,\hat{H}_0\big],\big[\hat{x},\hat{H}_0\big]\Big]\Big\rangle_{\text{eq.}} = \frac{1}{\hbar}\frac{1}{M^2}\left\langle\left[\frac{\partial\mathscr{V}(\hat{\vec{r}})}{\partial\hat{x}},\hat{p}_x\right]\right\rangle_{\text{eq.}}.$$

Replacing again the commutator $\big[\hat{p}_x,\ \cdot\ \big]$ by $-\mathrm{i}\hbar\dfrac{\partial}{\partial\hat{x}}$, one eventually obtains

$$\frac{1}{\hbar^3}\Big\langle\Big[\big[\hat{v}_x,\hat{H}_0\big],\big[\hat{x},\hat{H}_0\big]\Big]\Big\rangle_{\text{eq.}} = \frac{\mathrm{i}}{M^2}\left\langle\frac{\partial^2\mathscr{V}(\hat{\vec{r}})}{\partial\hat{x}^2}\right\rangle_{\text{eq.}}.$$

Invoking the invariance of the Hamiltonian (VI.94) under rotations, the second derivative with respect to x equals one third of the Laplacian, which all in all gives

$$\frac{1}{3M^2}\Big\langle\triangle\mathscr{V}(\hat{\vec{r}})\Big\rangle_{\text{eq.}} = \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty}\omega^2\,\tilde{\xi}_{v_xx}(\omega)\,\mathrm{d}\omega = \beta K_{\dot{v}_x\dot{v}_x}(0). \tag{VI.104b}$$

This sum rule relates the potential in which the heavy particle is evolving, the second moment of the spectral function $\tilde{\xi}_{v_xx}(\omega)$ and the autocorrelation of the heavy-particle acceleration.

Comparison with the correlation functions of the Langevin model

The Langevin model investigated in Sec. V.1 provides a classical, phenomenological description of the motion of a heavy particle, whose mass will hereafter be denoted as M, interacting with light particles. In particular, we studied in Sec. V.1.6 the time evolution of the heavy-particle velocity, averaged over many realizations of the motion, when the particle is submitted to an external force $F_{\text{ext.}}(t)$.

Now, the latter actually couples to the position x of the Brownian particle: the change in energy caused by the external force for a displacement Δx is simply the mechanical work $F_{\text{ext.}}\,\Delta x$.(90) Thus, the complex admittance characterizing the (linear!) response of the average velocity $\langle v_x\rangle$ to $F_{\text{ext.}}$ is, in the language of the present chapter, the generalized susceptibility χ_{v_xx}. Equation (V.39b) can therefore be rewritten as

$$\tilde{\chi}_{v_xx}(\omega) = \frac{1}{M}\frac{1}{\gamma-\mathrm{i}\omega}. \tag{VI.105}$$

According to relation (VI.99), this amounts to a spectral function

$$\tilde{\xi}_{v_xx}(\omega) = -\frac{\mathrm{i}}{M}\frac{\gamma}{\omega^2+\gamma^2}. \tag{VI.106}$$

One then easily checks that this phenomenological spectral function is purely imaginary and even [reciprocal relation (VI.98)] and that it fulfills the lowest-order sum rule (VI.102b). On the other hand, the second moment, as well as all higher moments, of $\tilde{\xi}_{v_xx}(\omega)$ diverges, so that the sum rule (VI.104b) cannot be satisfied.

This shows that the phenomenological Langevin equation (V.1) does not describe the correct behavior at large frequencies, i.e. at short times. As already suggested in Sec. V.3, this is due to the instantaneity of the friction force, which is simply proportional to the velocity at the same time.

---

(90) This seemingly contrived formulation is used to avoid invoking some Hamilton function for the Langevin model, and its perturbation by the external force.

VI.4.3 d Caldeira–Leggett model

In the previous paragraph, we have seen that the phenomenological Langevin model for the motion of a Brownian particle submitted to an external force yields correlation functions which do not fulfill the sum rules of linear-response theory. This means that the model can actually not be the macroscopic manifestation of an underlying microscopic dynamical model.(91)

From Sec. V.3.3, we already know that a classical heavy particle interacting with a bath of classical independent harmonic oscillators—which constitutes a special case of the model introduced in § VI.4.3 a—actually obeys a generalized Langevin equation when the bath degrees of freedom are integrated out. Here we want to consider this model again, now in the quantum-mechanical case.

Since the dynamics along different directions decouple, we restrict the study to a one-dimensional system, whose Hamilton operator is given by [cf. Eq. (V.68b)]

$$\boxed{\hat{H}_0 = \frac{\hat{p}^2}{2M} + \sum_{j=1}^{N}\left[\frac{\hat{p}_j^2}{2m_j} + \frac{1}{2}m_j\omega_j^2\left(\hat{x}_j - \frac{C_j}{m_j\omega_j^2}\hat{x}\right)^2\right].} \tag{VI.107}$$

$\hat{x}$, $\hat{p}$, M are the position, momentum and mass of the heavy particle, while $\hat{x}_j$, $\hat{p}_j$ and m_j denote those of the harmonic oscillators, with resonant frequencies ω_j, with which the particle interacts.

Equations of motion

The Heisenberg equation (II.37) for the position and momentum of the heavy particle read

$$\frac{\mathrm{d}\hat{x}(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{x}(t),\hat{H}_0\big] = \frac{1}{M}\hat{p}(t), \tag{VI.108a}$$

$$\frac{\mathrm{d}\hat{p}(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{p}(t),\hat{H}_0\big] = \sum_{j=1}^{N}C_j\hat{x}_j(t) - \left(\sum_{j=1}^{N}\frac{C_j^2}{m_j\omega_j^2}\right)\hat{x}(t). \tag{VI.108b}$$

These equations are sometimes referred to as *Heisenberg–Langevin equations*.

The first term on the right hand side of the evolution equation for the momentum only depends on the bath degrees of freedom. Introducing the ladder operators $\hat{a}_j(t)$, $\hat{a}_j^\dagger(t)$ of the bath oscillators, it can be rewritten as

$$\hat{R}(t) \equiv \sum_{j=1}^{N}C_j\hat{x}_j(t) = \sum_{j=1}^{N}C_j\sqrt{\frac{\hbar}{2m_j\omega_j}}\left[\hat{a}_j(t)+\hat{a}_j^\dagger(t)\right]. \tag{VI.109}$$

The Heisenberg equation

$$\frac{\mathrm{d}\hat{a}_j(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\big[\hat{a}_j(t),\hat{H}_0\big] = -\mathrm{i}\omega_j\hat{a}_j(t) + \mathrm{i}\frac{C_j}{\sqrt{2\hbar m_j\omega_j}}\hat{x}(t)$$

obeyed by the annihilation operator for the j-th oscillator admits the solution

$$\hat{a}_j(t) = \hat{a}_j(t_0)\,\mathrm{e}^{-\mathrm{i}\omega_j(t-t_0)} + \mathrm{i}\frac{C_j}{\sqrt{2\hbar m_j\omega_j}}\int_{t_0}^{t}\hat{x}(t')\,\mathrm{e}^{-\mathrm{i}\omega_j(t-t')}\,\mathrm{d}t',$$

with t_0 an arbitrary initial time. Inserting this expression and its adjoint in Eq. (VI.109), the evolution equation (VI.108b) becomes

$$\frac{\mathrm{d}\hat{p}(t)}{\mathrm{d}t} = \sum_{j=1}^{N}\frac{C_j^2}{2m_j\omega_j}\left[\mathrm{i}\int_{t_0}^{t}\hat{x}(t')\,\mathrm{e}^{-\mathrm{i}\omega_j(t-t')}\,\mathrm{d}t' + \mathrm{h.c.}\right] + \hat{F}_{\mathrm{L}}(t) - \left(\sum_{j=1}^{N}\frac{C_j^2}{m_j\omega_j^2}\right)\hat{x}(t), \tag{VI.110a}$$

where the operator $\hat{F}_{\mathrm{L}}(t)$ is defined as

---

(91) More precisely, the Langevin equation cannot emerge as macroscopic limit valid on arbitrary time scales—or equivalently for all frequencies—of an underlying microscopic theory, although it might constitute an excellent approximation in a limited time / frequency range.

$$\hat{F}_{\rm L}(t) \equiv \sum_{j=1}^{N} C_j \sqrt{\frac{\hbar}{2m_j\omega_j}} \Big[\hat{a}_j(t_0)\, {\rm e}^{-{\rm i}\omega_j(t-t_0)} + \hat{a}_j^\dagger(t_0)\, {\rm e}^{{\rm i}\omega_j(t-t_0)}\Big]. \tag{VI.110b}$$

This operator corresponds to a Langevin force, which only depends on freely evolving operators of the bath.[92] In turn, the first term on the right-hand side of Eq. (VI.110a) describes a retarded friction force exerted on the heavy particle by the bath, and due to the perturbation of the latter by the former at earlier times.

Limiting case of a continuous bath

Introducing as in Sec. V.3.3 c the *spectral density of the coupling to the bath*

$$J(\omega) \equiv \frac{\pi}{2} \sum_j \frac{C_j^2}{m_j\omega_j}\, \delta(\omega-\omega_j), \tag{VI.111}$$

and its continuous approximation $J_c(\omega)$ [cf. Eq. (V.75)] the retarded force in Eq. (VI.110a) becomes

$$\begin{aligned}\sum_{j=1}^{N} \frac{C_j^2}{2m_j\omega_j} \bigg[{\rm i}\int_{t_0}^{t} \hat{x}(t')\, {\rm e}^{-{\rm i}\omega_j(t-t')}\, {\rm d}t' + {\rm h.c.}\bigg] &= \frac{1}{\pi}\int J(\omega) \bigg[{\rm i}\int_{t_0}^{t} \hat{x}(t')\, {\rm e}^{-{\rm i}\omega(t-t')}\, {\rm d}t' + {\rm h.c.}\bigg] {\rm d}\omega \\ &\simeq \frac{1}{\pi}\int J_c(\omega) \bigg[{\rm i}\int_{t_0}^{t} \hat{x}(t')\, {\rm e}^{-{\rm i}\omega(t-t')}\, {\rm d}t' + {\rm h.c.}\bigg] {\rm d}\omega, \end{aligned} \tag{VI.112}$$

while the third term in that same equation can be rewritten as

$$-\Bigg(\sum_{j=1}^{N} \frac{C_j^2}{m_j\omega_j^2}\Bigg) \hat{x}(t) = -\bigg(\frac{2}{\pi}\int \frac{J(\omega)}{\omega}\, {\rm d}\omega\bigg) \hat{x}(t) \simeq -\bigg(\frac{2}{\pi}\int \frac{J_c(\omega)}{\omega}\, {\rm d}\omega\bigg) \hat{x}(t). \tag{VI.113}$$

With a trivial change of integration variable from t' to $\tau = t-t'$ and some rewriting, the retarded force (VI.112) becomes after exchanging the order of integrations

$$\frac{1}{\pi}\int \frac{J_c(\omega)}{\omega} \bigg[{\rm i}\omega\int_0^{t-t_0} \hat{x}(t-\tau)\, {\rm e}^{-{\rm i}\omega\tau}\, {\rm d}t' + {\rm h.c.}\bigg] {\rm d}\omega = -\frac{1}{\pi}\int_0^{t-t_0} \hat{x}(t-\tau) \frac{\rm d}{{\rm d}\tau} \bigg[\int \frac{J_c(\omega)}{\omega} \big({\rm e}^{-{\rm i}\omega\tau} + {\rm e}^{{\rm i}\omega\tau}\big)\, {\rm d}\omega\bigg] {\rm d}\tau.$$

Introducing the "memory kernel" [cf. Eq. (V.74)]

$$\gamma(\tau) \equiv \frac{2}{\pi}\int \frac{J_c(\omega)}{M\omega} \cos\omega\tau\, {\rm d}\omega \tag{VI.114}$$

and performing an integration by parts, in which the equation of motion (VI.108a) allows us to replace the time derivative of $\hat{x}(t)$ by $\hat{p}(t)/M$, the friction force becomes

$$-M\int_0^{t-t_0} \hat{x}(t-\tau)\, \gamma'(\tau)\, {\rm d}\tau = M\Big[\gamma(0)\hat{x}(t) - \gamma(t-t_0)\hat{x}(t_0)\Big] - \int_0^{t-t_0} \hat{p}(t-\tau)\, \gamma(\tau)\, {\rm d}\tau.$$

In many simple cases, corresponding to oscillator baths with a "short memory", the kernel $\gamma(\tau)$ only takes significant values in a limit range of size ω_c^{-1} around $\tau = 0$. As soon as $|t-t_0| \gg \omega_c^{-1}$, the term $\gamma(t-t_0)$ in the above equation then becomes negligible, while the upper limit of the integral can be sent to $+\infty$ without affecting the result significantly. Deducing $\gamma(0)$ from Eq. (VI.114), the friction force (VI.112) reads

$$\frac{1}{\pi}\int J_c(\omega) \bigg[{\rm i}\int_{t_0}^{t} \hat{x}(t')\, {\rm e}^{-{\rm i}\omega(t-t')}\, {\rm d}t' + {\rm h.c.}\bigg] {\rm d}\omega = \bigg(\frac{2}{\pi}\int \frac{J_c(\omega)}{\omega}\, {\rm d}\omega\bigg) \hat{x}(t) - \int_0^{\infty} \hat{p}(t-\tau)\, \gamma(\tau)\, {\rm d}\tau.$$

The first term on the right hand side is exactly the negative of Eq. (VI.113): putting everything together, the evolution equation (VI.110a) takes the simple form of a generalized Langevin equation

$$\frac{{\rm d}\hat{p}(t)}{{\rm d}t} = -\int_0^{\infty} \hat{p}(t-\tau)\, \gamma(\tau)\, {\rm d}\tau + \hat{F}_{\rm L}(t). \tag{VI.115}$$

Dividing this equation by M, one obtains a similar evolution equation for the velocity operator $\hat{v}(t)$.

---

[92] One easily checks in a basis of energy eigenstates $\langle\hat{a}_j(t_0)\rangle_{\rm eq.} = \langle\hat{a}_j^\dagger(t_0)\rangle_{\rm eq.} = 0$ for all bath oscillators, which results in $\langle\hat{F}_{\rm L}(t)\rangle_{\rm eq.} = 0$.

Generalized susceptibility

Let us add to the Hamiltonian (VI.107) a perturbation $\hat{W} = -F_{\rm ext.}(t)\hat{x}(t)$ coupling to the position of the Brownian particle. One easily checks that this perturbation amounts to adding an extra term $F_{\rm ext.}(t)\hat{\mathbb{1}}$ on the right-hand side of Eq. (VI.115). Dividing the resulting equation by M, taking the average, and Fourier transforming, one obtains the generalized susceptibility [cf. Eq. (V.62)]

$$\tilde{\chi}_{vx}(\omega) = \frac{1}{M}\frac{1}{\tilde{\gamma}(\omega) - \mathrm{i}\omega}, \tag{VI.116a}$$

where $\tilde{\gamma}(\omega)$ is given by

$$\tilde{\gamma}(\omega) = \int \gamma(t)\,\Theta(t)\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t. \tag{VI.116b}$$

The Caldeira–Leggett Hamiltonian (VI.107) is invariant under time reversal. As already seen in § (VI.4.3 b), this leads to the proportionality between the spectral function $\tilde{\xi}_{vx}(\omega)$ and the real part of the generalized susceptibility $\tilde{\chi}_{vx}(\omega)$:

$$\tilde{\xi}_{vx}(\omega) = -\frac{\mathrm{i}}{M}\frac{\mathrm{Re}\,\tilde{\gamma}(\omega)}{|\tilde{\gamma}(\omega) - \mathrm{i}\omega|^2}.$$

If $\tilde{\gamma}(\omega)$ decreases quickly enough as $|\omega|$ goes to ∞—which depends on the specific behavior of $J_c(\omega)$ at infinity, see Eq. (VI.114)—, the spectral function $\tilde{\xi}_{vx}(\omega)$ can have moments to all orders, which can then obey the sum rules (VI.81).

Bibliography

Parts of this Chapter are strongly inspired by the (unpublished) lectures given by C. Cohen-Tannoudji in 1977–78 at the Collège de France [59], especially lectures V–VII.

Besides the chapters of textbooks listed below, the reader may also be interested in the reviews by Berne & Harp [60], Berne [61], Kubo [53] or Zwanzig [62].

- Kubo *et al.*, *Statistical physics II* [2], chapter 4.
- Landau & Lifshitz, *Course of theoretical physics. Vol. 5: Statistical physics part 1* [3], chapter XII § 118–126.
- Le Bellac, Mortessagne & Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* [9], chapter 9.1–9.2.
- Pottier, *Nonequilibrium statistical physics* [6], chapter 12–14.
- Reif, *Fundamentals of statistical and thermal physics* [35], chapter 15.16–15.18.
- Zwanzig, *Nonequilibrium statistical mechanics* [7], chapter 7.

Appendices to Chapter VI

VI.A Non-uniform phenomena

Until now in this chapter, we have implicitly considered only homogeneous systems, perturbed by uniform excitations. In this appendix, we generalize part of the formalism and results developed above to non-uniform systems. Much more can be found in the key article by L. Kadanoff[bz] and P. Martin[ca] [63].

VI.A.1 Space-time correlation functions

Consider a quantum-mechanical system at thermodynamic equilibrium. Under the influence of spontaneous fluctuations or of some external perturbation, its properties at a given instant and position, corresponding to some intensive variable, might depart from their equilibrium values. It then becomes interesting to quantify the correlation between such properties at different times and positions.

Associating space-dependent observables to intensive properties like the local density of particle number or of energy, we are thus led to consider averages of the type

$$\big\langle \hat{B}_{\rm I}(t,\vec{r})\,\hat{A}_{\rm I}(t',\vec{r}\,')\big\rangle_{\rm eq.}.$$

As usual, we can invoke the stationarity of the equilibrium state to show that such an average only depends on the time difference $\tau = t - t'$. If in addition the equilibrated system is invariant under spatial translations—as we assume from now on—, one also finds that the expectation value depends only the separation $\vec{r}-\vec{r}\,'$, not on $\vec{r}$ and $\vec{r}\,'$ separately.

Remark: The assumed invariance under arbitrary spatial translations is less warranted as that under time translations. Strictly speaking, the assumption can only hold in fluids, since a crystalline solid is only invariant under some discrete translations. In addition, it can only hold if the system—in particular its volume $\mathcal{V}$, which explicitly appears in some of the formulae below—is infinitely large. In practice, these mathematical caveats are in many cases irrelevant.

Quite obviously, the various time-correlation functions introduced in Sec. VI.1 can be generalized to functions of the spatial separation, which can all be expressed in terms of the non-symmetrized correlation function

$$C_{BA}(\tau,\vec{r}) \equiv \big\langle \hat{B}_{\rm I}(\tau,\vec{r})\,\hat{A}(0,\vec{0})\big\rangle_{\rm eq.}. \tag{VI.117}$$

It will be fruitful to investigate these functions not only in “direct” space, but also in reciprocal space, i.e. after performing a spatial Fourier transform. For a generic quantity $X(t,\vec{r})$, this transform is defined as (note the minus sign!)

$$\tilde{X}_{\rm I}(t,\vec{q}) \equiv \int_{\mathbb{R}^3} X_{\rm I}(t,\vec{r})\,{\rm e}^{-{\rm i}\vec{q}\cdot\vec{r}}\,{\rm d}^3\vec{r}. \tag{VI.118}$$

Thus, rewriting Eq. (VI.117) as $C_{BA}(\tau,\vec{r}) = \big\langle \hat{B}_{\rm I}(\tau,\vec{r}+\vec{r}\,')\,\hat{A}(0,\vec{r}\,')\big\rangle_{\rm eq.}$ and multiplying both sides of the identity by ${\rm e}^{-{\rm i}\vec{q}\cdot\vec{r}} = {\rm e}^{-{\rm i}\vec{q}\cdot(\vec{r}+\vec{r}\,')}\,{\rm e}^{{\rm i}\vec{q}\cdot\vec{r}\,'}$, one finds after integrating over $\vec{r}+\vec{r}\,'$ and $\vec{r}\,'$

$$\int C_{BA}(\tau,\vec{r})\,{\rm e}^{-{\rm i}\vec{q}\cdot\vec{r}}\,{\rm d}^3\vec{r} = \frac{1}{\mathcal{V}}\big\langle \tilde{\hat{B}}(\tau,\vec{q})\,\tilde{\hat{A}}(0,-\vec{q})\big\rangle_{\rm eq.},$$

where both sides were divided by the system volume $\mathcal{V} = \int {\rm d}^3\vec{r}\,'$.

[bz] L. P. Kadanoff, 1937–2015 [ca] P. C. Martin, 1931–2016

Fourier transforming with respect to τ, one defines

$$\tilde{C}_{BA}(\omega,\vec{q}) \equiv \int\left[\int_{-\infty}^{\infty} C_{BA}(\tau,\vec{r})\,\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau\right]\mathrm{e}^{-\mathrm{i}\vec{q}\cdot\vec{r}}\,\mathrm{d}^3\vec{r} = \frac{1}{\mathcal{V}}\int_{-\infty}^{\infty}\big\langle \tilde{\hat{B}}(\tau,\vec{q})\,\tilde{\hat{A}}(0,-\vec{q})\big\rangle_{\mathrm{eq.}}\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau, \qquad \text{(VI.119)}$$

where attention should be payed to the opposite arguments $\vec{q}$, $-\vec{q}$ of the observables.

VI.A.2 Non-uniform linear response

As example of these generalized correlation functions, the retarded propagator is the function which characterizes the local linear response to a non-uniform excitation by an inhomogeneous classical field $f(t,\vec{r})$

$$\hat{W}(t) = -\int f(t,\vec{r})\,\hat{A}(\vec{r})\,\mathrm{d}^3\vec{r}, \qquad \text{(VI.120)}$$

which generalizes Eq. (VI.7). Under this perturbation, the change of a local property reads

$$\big\langle \hat{B}_{\mathrm{I}}(t,\vec{r})\big\rangle_{\mathrm{n.eq.}} = \big\langle \hat{B}(\vec{r})\big\rangle_{\mathrm{eq.}} + \int_{\mathcal{V}}\left[\int_{-\infty}^{\infty}\chi_{BA}(t-t',\vec{r}-\vec{r}\,')\,f(t',\vec{r}\,')\,\mathrm{d}t'\right]\mathrm{d}^3\vec{r}\,' + \mathcal{O}(f^2), \qquad \text{(VI.121a)}$$

which defines $\chi_{BA}(\tau,\vec{r})$. Repeating the derivation of Sec. VI.2.1, one easily finds that the latter is given by the Kubo formula

$$\boxed{\chi_{BA}(\tau,\vec{r}) \equiv \frac{\mathrm{i}}{\hbar}\big\langle\big[\hat{B}_{\mathrm{I}}(\tau,\vec{r}),\hat{A}(0,\vec{0})\big]\big\rangle_{\mathrm{eq.}}\Theta(\tau).} \qquad \text{(VI.121b)}$$

In Fourier space, the Kubo formula (VI.121a) becomes (assuming that $\hat{B}$ is centered)

$$\big\langle \tilde{\hat{B}}(\omega,\vec{q})\big\rangle_{\mathrm{n.eq.}} = \tilde{\chi}_{BA}(\omega,\vec{q})\,\tilde{f}(\omega,\vec{q})\,, \qquad \text{(VI.122a)}$$

where the generalized susceptibility is given by

$$\tilde{\chi}_{BA}(\omega,\vec{q}) \equiv \int_{\mathbb{R}^3}\left[\lim_{\varepsilon\to0^+}\int_{-\infty}^{\infty}\chi_{BA}(\tau,\vec{r})\,\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{e}^{-\varepsilon\tau}\,\mathrm{d}\tau\right]\mathrm{e}^{-\mathrm{i}\vec{q}\cdot\vec{r}}\,\mathrm{d}^3\vec{r}. \qquad \text{(VI.122b)}$$

Dynamic structure factor

An important example of application is the coupling of a non-uniform external scalar potential $V_{\mathrm{ext.}}(t,\vec{r})$ to a system of particles in equilibrium, and more precisely to their number density, $\hat{n}(\vec{r})$:

$$\hat{W} = \int \hat{n}(\vec{r})\,V_{\mathrm{ext.}}(t,\vec{r})\,\mathrm{d}^3\vec{r}.$$

The response of the number density itself is given by the *compressibility* of the system

$$\big\langle \hat{n}(t,\vec{r})\big\rangle_{\mathrm{n.eq.}} = \big\langle \hat{n}(\vec{r})\big\rangle_{\mathrm{eq.}} + \int \chi_{nn}(t-t',\vec{r}-\vec{r}\,')\,\hat{n}(t',\vec{r}\,')\,\mathrm{d}t'\,\mathrm{d}\vec{r}\,'.$$

In that case, the autocorrelation function $\tilde{C}_{nn}(\omega,\vec{q})$ is usually denoted by $\tilde{S}(\omega,\vec{q})$ and called *dynamic structure factor*:(93)

$$\tilde{S}(\omega,\vec{q}) \equiv \frac{1}{\mathcal{V}}\int_{-\infty}^{\infty}\big\langle \hat{n}(\tau,\vec{q})\,\hat{n}(0,-\vec{q})\big\rangle_{\mathrm{eq.}}\,\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau = \frac{2\pi}{\mathcal{V}}\sum_{n,n'}\pi_n\big|(n_{\vec{q}})_{nn'}\big|^2\delta(\omega_{n'n}-\omega), \qquad \text{(VI.123)}$$

where we have introduced the Lehmann representation, involving the matrix elements $(n_{\vec{q}})_{nn'}$ of $\hat{n}(0,\vec{q})$, which obey the identity $(n_{\vec{q}})_{nn'} = (n_{-\vec{q}})^*_{nn'}$.

This dynamic structure function is directly measurable in a scattering experiment on the system. If a beam of particles with momentum $\vec{k}$ is sent onto the system, one can show that in the Born approximation, the intensity scattered with some momentum $\vec{k}\,'$, amounting to a momentum transfer $\vec{q} \equiv \vec{k}\,'-\vec{k}$, is proportional to the product of $\tilde{S}(-\omega,-\vec{q})$ and the *form factor*. The latter characterizes

---

(93) The factor $1/\mathcal{V}$ is sometimes omitted from the definition.

the scattering probability on a single center, in particular it quantifies how much the scattering center differs from a point-like particle. In turn, the dynamic structure function contains the information on the distribution and dynamics of the microscopic scattering centers in the system.(94)

Introducing the spectral function of the system

$$\tilde{\xi}(\omega,\vec{q}) \equiv \frac{1}{2\hbar\mathcal{V}}\int_{-\infty}^{\infty}\big\langle\big[\hat{n}(\tau,\vec{q}),\hat{n}(0,-\vec{q})\big]\big\rangle_{\text{eq.}}\,\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau$$

one easily checks that it is related to the dynamic structure factor according to

$$\tilde{\xi}(\omega,\vec{q}) = \frac{1}{2\hbar}\big[\tilde{S}(\omega,\vec{q})-\tilde{S}(-\omega,-\vec{q})\big] = \frac{1}{2\hbar}\big(1-\mathrm{e}^{-\beta\hbar\omega}\big)\tilde{S}(\omega,\vec{q}),$$

which expresses the fluctuation-dissipation theorem. Additionally, one finds the relation

$$\operatorname{Im}\tilde{\chi}_{nn}(\omega,\vec{q}) = -\tilde{\xi}(\omega,\vec{q})$$

with the generalized susceptibility of the system.

VI.A.3 Some properties of space-time autocorrelation functions

We now give a few relations obeyed by the non-symmetric correlation function, specializing to the case of identical observables $\hat{A}=\hat{B}$, that is of autocorrelation functions $C_{AA}(\tau,\vec{r})$. One can then check that the spatial Fourier transforms (we drop the factor $1/\mathcal{V}$) satisfy a few properties:

- $\big\langle\hat{A}(t,\vec{q})\,\hat{A}(0,-\vec{q})\big\rangle_{\text{eq.}} = \big\langle\hat{A}(0,\vec{q})\,\hat{A}(-t,-\vec{q})\big\rangle_{\text{eq.}}$; (VI.124a)
- $\big\langle\hat{A}(t,\vec{q})\,\hat{A}(0,-\vec{q})\big\rangle^{*}_{\text{eq.}} = \big\langle\hat{A}(0,\vec{q})\,\hat{A}(t,-\vec{q})\big\rangle_{\text{eq.}}$; (VI.124b)
- $\big\langle\hat{A}(t,\vec{q})\,\hat{A}(0,-\vec{q})\big\rangle^{*}_{\text{eq.}} = \big\langle\hat{A}(t-\mathrm{i}\hbar\beta,-\vec{q})\,\hat{A}(0,\vec{q})\big\rangle_{\text{eq.}}$. (VI.124c)

The latter identity is known as the *Kubo–Martin–Schwinger*(cb) *condition* for the observables of a system in canonical equilibrium.
From these identities follow a few properties of the double Fourier transform $\tilde{C}_{AA}(\omega,\vec{q})$:

- $\tilde{C}_{AA}(\omega,\vec{q})$ is a real number; (VI.125a)
- detailed balance condition: $\tilde{C}_{AA}(\omega,\vec{q}) = \tilde{C}_{AA}(-\omega,-\vec{q})\,\mathrm{e}^{\beta\hbar\omega}$, (VI.125b)

where the latter generalizes Eq. (VI.53).

VI.B Classical linear response

The linear response formalism can also be applied to systems which are described classically, as e.g. fluids obeying hydrodynamical laws. Two parallel strategies can then be adopted: either to consider the classical limit of the quantum-mechanical results, or to tackle the problem in the classical framework from the beginning. In this appendix we give examples of both approaches, which quite naturally lead to the same results.

VI.B.1 Classical correlation functions

Let A, B be classical observables associated with quantum-mechanical counterparts $\hat{A}$, $\hat{B}$. Using (without proof) the correspondence between quantum-mechanical and classical statistical-mechanical expectation values, the non-symmetrized correlation function $C_{BA}(\tau)$, Eq. (VI.12), becomes in the classical limit

$$\text{classical limit of } C_{BA}(\tau) = \big\langle B(\tau)A(0)\big\rangle_{\text{eq.}} \equiv C^{\text{cl.}}_{BA}(\tau), \qquad \text{(VI.126)}$$

---

(94) For more details, see Van Hove's original article on the topic [64], which is easily readable.

(cb) J. Schwinger, 1918–1994

where the expectation value is a Γ-space integral computed with the proper equilibrium phase-space distribution.

In the classical limit, operators become commuting numbers ("c-numbers"). Invoking the stationarity of the equilibrium state, one thus has

$$C^{\rm cl.}_{BA}(\tau) = C^{\rm cl.}_{AB}(-\tau), \tag{VI.127}$$

in contrast to the quantum-mechanical case where the identity is between $C_{BA}(\tau)$ and $C_{AB}(-\tau)^*$, see Eq. (VI.15). In the case of autocorrelations ($B = A$), the non-symmetrized correlation function (VI.126) is even.

Fourier transforming both sides of relation (VI.127), one finds at once

$$\tilde{C}^{\rm cl.}_{BA}(\omega) = \tilde{C}^{\rm cl.}_{AB}(-\omega), \tag{VI.128}$$

which is the classical limit $\hbar \to 0$ of the detailed balance relation, as was already discussed below Eq. (VI.53).

Thanks to the commutativity of the A and $B(\tau)$, the classical limit of $S_{BA}(\tau)$ is given by the same correlation function $\langle B(\tau)A(0)\rangle_{\rm eq.}$

$$\text{classical limit of } S_{BA}(\tau) = \big\langle B(\tau)A(0)\big\rangle_{\rm eq.} = C^{\rm cl.}_{BA}(\tau). \tag{VI.129}$$

Similarly, the various operators in the defining integral for Kubo's canonical correlation functions commute with each other in the classical limit, and one obtains

$$\text{classical limit of } K_{BA}(\tau) = \big\langle B(\tau)A(0)\big\rangle_{\rm eq.}\,\frac{1}{\beta}\int_0^\beta {\rm d}\lambda = \big\langle B(\tau)A(0)\big\rangle_{\rm eq.} = C^{\rm cl.}_{BA}(\tau). \tag{VI.130}$$

We thus recover the fact that S_{BA} and K_{BA} have the same classical limit, as mentioned for their Fourier transforms at the end of § VI.3.3 b.

Remark: More generally, even in the quantum-mechanical case if either $\hat{A}$ or $\hat{B}$ commutes with $\hat{H}_0$, then the non-symmetrized, symmetric and canonical correlation functions $C_{BA}(\tau)$, $S_{BA}(\tau)$, $K_{BA}(\tau)$ are equal.

In contrast, the linear response function $\chi_{BA}(\tau)$ and the Fourier transform $\xi_{BA}(\tau)$ of the spectral function are proportional to commutators divided by $\hbar$, see Eqs. (VI.26) and (VI.19). In the classical limit, these become proportional to some Poisson brackets, for instance [see also Eq. (VI.138b) hereafter]

$$\text{classical limit of } \xi_{BA}(\tau) = \frac{\rm i}{2}\big\langle\big\{B_N(\tau), A_N\big\}\big\rangle_{\rm eq.} \equiv \xi^{\rm cl.}_{BA}(\tau). \tag{VI.131}$$

VI.B.2 Classical Kubo formula

In this Subsection, we want to show how some results of linear response theory can be derived directly in classical mechanics, instead of taking the limit $\hbar \to 0$ in quantum-mechanical results.

For that purpose, we consider(95) a system of N pointlike particles with positions and conjugate momenta $\{q_i\}$, $\{p_i\}$ with $1 \leq i \leq 3N$. The Γ-space density and Hamilton function of this system are denoted by $\rho_N\big(t, \{q_i\}, \{p_i\}\big)$ and $H_N\big(t, \{q_i\}, \{p_i\}\big)$. The latter arises from slightly perturbing a time-independent Hamiltonian $H^{(0)}_N\big(\{q_i\}, \{p_i\}\big)$:

$$H_N\big(t, \{q_i\}, \{p_i\}\big) = H^{(0)}_N\big(\{q_i\}, \{p_i\}\big) - f(t)A_N\big(\{q_i\}, \{p_i\}\big), \tag{VI.132}$$

with $A_N\big(\{q_i\}, \{p_i\}\big)$ an observable of the system and $f(t)$ a time-dependent function which vanishes as $t \to -\infty$.

---

(95) This is the generic setup of Sec. II.2.1.

Let $\mathrm{i}\mathcal{L}_0$ be the Liouville operator (II.11) associated to $H_N^{(0)}$ and $\rho_{\rm eq.}$ the N-particle phase-space density corresponding to the canonical equilibrium of the unperturbed system

$$\rho_{\rm eq.}\big(\{q_i\},\{p_i\}\big) = \frac{1}{Z_N(\beta)}\,\mathrm{e}^{-\beta H_N^{(0)}(\{q_i\},\{p_i\})} \quad \text{with} \quad Z_N(\beta) = \int \mathrm{e}^{-\beta H_N^{(0)}(\{q_i\},\{p_i\})}\,\mathrm{d}^{6N}\mathcal{V}, \tag{VI.133}$$

where the Γ-space infinitesimal volume element is given by Eq. (II.4a). Averages computed with $\rho_{\rm eq.}$ will be denoted as $\langle\,\cdot\,\rangle_{\rm eq.}$, those computed with ρ_N as $\langle\,\cdot\,\rangle_{\rm n.eq.}$.

Let $B_N\big(\{q_i\},\{p_i\}\big)$ denote another observable of the system. We wish to compute its out-of-equilibrium expectation value at time t, $\langle B_N(t,\{q_i\},\{p_i\})\rangle_{\rm n.eq.}$, and in particular its departure from the equilibrium expectation value $\langle B_N(\{q_i\},\{p_i\})\rangle_{\rm eq.}$. The latter is time-independent, as follows from Eqs. (II.18)–(II.19) and the time-independence of $\rho_{\rm eq.}$.

For the sake of brevity we shall from now on drop the dependence of functions on the phase-space coordinates $\{q_i\},\{p_i\}$.

In analogy to the quantum-mechanical case (Sec. VI.2.1), we start by calculating the phase-space density $\rho_N(t)$, or equivalently its deviation from the equilibrium density

$$\delta\rho_N(t) \equiv \rho_N(t) - \rho_{\rm eq.}. \tag{VI.134}$$

Writing $\rho_N(t) = \rho_{\rm eq.} + \delta\rho_N(t)$ and using the stationarity of $\rho_{\rm eq.}$ the Liouville equation (II.10b) for the evolution of $\rho_N(t)$

$$\frac{\mathrm{d}\rho_N(t)}{\mathrm{d}t} + \big\{\rho_N(t), H_N(t)\big\} = 0$$

gives for $\delta\rho_N(t)$ to leading order in the perturbation

$$\begin{aligned}\frac{\mathrm{d}\delta\rho_N(t)}{\mathrm{d}t} &= \big\{H_N(t),\delta\rho_N(t)\big\} + \big\{-f(t)A_N,\rho_{\rm eq.}\big\} + \mathcal{O}(f^2)\\ &= -\mathrm{i}\mathcal{L}_0\delta\rho_N(t) - f(t)\big\{A_N,\rho_{\rm eq.}\big\} + \mathcal{O}(f^2).\end{aligned} \tag{VI.135}$$

In the second line, we took $f(t)$ outside of the Poisson brackets since it is independent of the phase-space coordinates.

This is an inhomogeneous first-order linear differential equation, whose solution reads

$$\delta\rho_N(t) = -\int_{-\infty}^{t} \mathrm{e}^{-\mathrm{i}(t-t')\mathcal{L}_0}\big\{A_N,\rho_{\rm eq.}\big\}f(t')\,\mathrm{d}t' + \mathcal{O}(f^2),$$

where we used $f(-\infty)=0$ which results in $\delta\rho_N(-\infty)=0$. Again, the independence of $f(t')$ from the Γ-space coordinates allows one to move it to the left of the time-translation operator $\mathrm{e}^{-\mathrm{i}(t-t')\mathcal{L}_0}$. Adding $\rho_{\rm eq.}$ to both sides then gives

$$\rho_N(t) = \rho_{\rm eq.} - \int_{-\infty}^{t} f(t')\,\mathrm{e}^{-\mathrm{i}(t-t')\mathcal{L}_0}\big\{A_N,\rho_{\rm eq.}\big\}\,\mathrm{d}t' + \mathcal{O}(f^2). \tag{VI.136}$$

Multiplying this identity left with B_N and integrating afterwards over phase space yields

$$\big\langle B_N(t)\big\rangle_{\rm n.eq.} = \big\langle B_N\big\rangle_{\rm eq.} - \int_{-\infty}^{t} f(t')\bigg[\int_\Gamma B_N\,\mathrm{e}^{-\mathrm{i}(t-t')\mathcal{L}_0}\big\{A_N,\rho_{\rm eq.}\big\}\,\mathrm{d}^{6N}\mathcal{V}\bigg]\mathrm{d}t' + \mathcal{O}(f^2). \tag{VI.137}$$

Using the unitarity of $\mathrm{e}^{-\mathrm{i}(t-t')\mathcal{L}_0}$, Eq. (II.20), the phase-space integral on the right-hand side can be recast as

$$\int_\Gamma \Big[\mathrm{e}^{\mathrm{i}(t-t')\mathcal{L}_0}B_N\Big]\big\{A_N,\rho_{\rm eq.}\big\}\,\mathrm{d}^{6N}\mathcal{V}.$$

Invoking Eq. (II.17), the term between square brackets is then $B_N(t-t')$ *as would follow from letting B_N evolve under the influence of $H_N^{(0)}$ only.*(96) Equation (VI.137) thus becomes

$$\big\langle B_N(t)\big\rangle_{\rm n.eq.} = \big\langle B_N\big\rangle_{\rm eq.} - \int_{-\infty}^{t} f(t')\bigg[\int_\Gamma B_N(t-t')\big\{A_N,\rho_{\rm eq.}\big\}\,\mathrm{d}^{6N}\mathcal{V}\bigg]\mathrm{d}t' + \mathcal{O}(f^2).$$

---

(96) That is, corresponding to the interaction picture in the quantum mechanical case.

By performing an integration by parts and using the fact that the phase-space distribution vanishes at infinity, one checks that the integral over phase space of $B_N(t-t')\{A_N, \rho_{\text{eq.}}\}$ equals that of $\rho_{\text{eq.}}\{B_N(t-t'), A_N\}$:

$$\langle B_N(t)\rangle_{\text{n.eq.}} = \langle B_N\rangle_{\text{eq.}} - \int_{-\infty}^{t} f(t')\left[\int_\Gamma \rho_{\text{eq.}}\{B_N(t-t'), A_N\}\,\mathrm{d}^{6N}\mathcal{V}\right]\mathrm{d}t' + \mathcal{O}(f^2).$$

The phase-space integral in this relation is now simply the equilibrium expectation value of the Poisson bracket $\{B_N(t-t'), A_N\}$. All in all, this gives

$$\begin{aligned}\langle B_N(t)\rangle_{\text{n.eq.}} &= \langle B_N\rangle_{\text{eq.}} - \int_{-\infty}^{t} f(t')\langle\{B_N(t-t'), A_N\}\rangle_{\text{eq.}}\mathrm{d}t' + \mathcal{O}(f^2)\\ &= \langle B_N\rangle_{\text{eq.}} + \int_{-\infty}^{\infty} f(t')\chi^{\text{cl.}}_{BA}(t-t')\,\mathrm{d}t' + \mathcal{O}(f^2),\end{aligned} \tag{VI.138a}$$

with

$$\chi^{\text{cl.}}_{BA}(\tau) \equiv -\langle\{B_N(\tau), A_N\}\rangle_{\text{eq.}}\Theta(\tau). \tag{VI.138b}$$

This result is—as it should be—what follows from the quantum-mechanical Kubo formula (VI.26) when making the usual substitution

$$\frac{1}{\mathrm{i}\hbar}[\,\cdot\,,\,\cdot\,] \to \{\,\cdot\,,\,\cdot\,\}.$$

Appendices

APPENDIX A

Some useful formulae

- Gaussian(cc) integrals: for $a \in \mathbb{C}$, $\operatorname{Re} a > 0$

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}, \tag{A.1a}$$

which after n differentiations with respect to a gives the even moments

$$\int_{-\infty}^{\infty} x^{2n}\, e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a^{2n+1}}}\,\frac{(2n)!}{2^{2n} n!}, \tag{A.1b}$$

while the odd moments trivially vanish.

Related integrals are

$$\int_{0}^{\infty} x^{2n}\, e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a^{2n+1}}}\,\frac{(2n)!}{2^{2n+1} n!} = \frac{\Gamma\big(n+\frac{1}{2}\big)}{2a^{n+1/2}}, \tag{A.1c}$$

which follows either of the parity of the integrand in Eq. (A.1b) or of the change of variable $y = ax^2$, which also yields

$$\int_{0}^{\infty} x^{2n+1}\, e^{-ax^2}\, dx = \frac{n!}{2a^{n+1}}. \tag{A.1d}$$

- Cauchy(cd) principal value:

Consider a function $f(x)$ over $\mathbb{R}$, with an isolated singularity at $x = 0$ and regular elsewhere. One then defines its *principal value* $\mathscr{P} f$ as a distribution (or generalized function) such that for every test function g

$$\mathscr{P} f(g) = \lim_{\varepsilon \to 0^+} \left[\int_{-\infty}^{-\varepsilon} f(x)\, g(x)\, dx + \int_{\varepsilon}^{\infty} f(x)\, g(x)\, dx \right]. \tag{A.2a}$$

In particular, the principal value of $f(x) = 1/x$ is such that

$$\mathscr{P}\frac{1}{x} = \lim_{\varepsilon \to 0^+} \frac{1}{x \pm i\varepsilon} \pm i\pi\delta(x). \tag{A.2b}$$

(cc) C. F. Gauss, 1777–1855 (cd) A. L. Cauchy, 1789–1857

APPENDIX B

Elements on random variables

This Appendix summarizes a few elements of probability theory, with a focus on random variables, adopting the point of view of a physicist interested in results more than in formal proofs.(97)

B.1 Definition

The notion of a *random variable*—or *stochastic variable*—X relies on two elements:

a) The set Ω—referred to as *sample space*, *universe* or *range*—of the possible values x (the *realisations*) describing the outcome of a "random experiment".

This set can be either discrete or continuous, or even partly discrete and partly continuous. Besides, the sample space can be multidimensional. Accordingly, one speaks of discrete, continuous or multidimensional random variables. The latter will in the following often be represented as vectors $\mathbf{X}$.

A physical instance of discrete resp. continuous one-dimensional random variable is the projection of the spin of a particle on a given axis, resp. the kinetic energy of a free particle. Examples of continuous 3-dimensional stochastic variables are the three components of the velocity $\vec{v}$ or those of the position $\vec{x}$ of a Brownian particle at a given instant.

b) The *probability distribution* on this set.

Consider first a continuous one-dimensional random variable defined on a real interval (or on a union of intervals) $\mathcal{I}$. The probability distribution is specified through a *probability density*, that is a nonnegative function $p_X(x)$

$$p_X(x) \geq 0 \quad \forall x \in \mathcal{I} \tag{B.1a}$$

normalized to 1 over its range of definition

$$\int_{\mathcal{I}} p_X(x)\,\mathrm{d}x = 1. \tag{B.1b}$$

$p_X(x)\,\mathrm{d}x$ represents the probability that X takes a value between x and $x + \mathrm{d}x$.

To account for the possible presence of discrete subsets in the sample space, the probability distribution may involve Dirac distributions:

$$p_X(x) = \sum_n p_n \delta(x - x_n) + \tilde{p}_X(x), \tag{B.2a}$$

with the normalization condition

$$\sum_n p_n + \int \tilde{p}_X(x)\,\mathrm{d}x = 1, \tag{B.2b}$$

---

(97) The presentation is strongly inspired from Chapter I of van Kampen's(ce) classic book [49].

(ce) N. G. van Kampen, 1921–2013

with $p_n > 0$ and $\tilde{p}_X$ a nonnegative function. If $\tilde{p}_X = 0$ identically over the range, then X is simply a discrete random variable. The corresponding probability density is then replaced by a *probability mass function*, which associates finite positive probabilities p_n to the respective realizations x_n.

The generalization to the case of a multidimensional stochastic variable is straightforward and involves D-dimensional integrals. The corresponding D-dimensional infinitesimal volume around a point $\mathbf{x}$ will hereafter be denoted by $\mathrm{d}^D\mathbf{x}$.

Remark: Physical quantities often possess a dimension, like length, mass, time... As a consequence, the probability density p_G for the distribution of the values g of such a quantity G must also have a dimension, namely the inverse of that of G, to ensure that the probability $p_G(g)\,\mathrm{d}g$ be dimensionless. This property can easily be checked on the various probability densities introduced in Sec. B.3.

In formal probability theory, one distinguishes between the sample space Ω—the set of all possible "outcomes" of a random experiment—and a set $\mathcal{F}$, which is a subset of the power set (set of all subsets) of Ω. The elements of $\mathcal{F}$, called "events", represent the... events, which one can observe. Eventually, one introduces a function, the "probability measure", $\mathcal{P}$ from $\mathcal{F}$ in the real interval $[0,1]$, which associates to each event $A \in \mathcal{F}$ a probability $\mathcal{P}(A)$ fulfilling the conditions

- $\mathcal{P}(\Omega) = 1$ [normalization, cf. Eq. (B.1b) or (B.2b)],
- $\forall A, B \in \mathcal{F}$, $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B)$ if $\mathcal{P}(A \cap B) = 0$—in particular when $A \cap B = \emptyset$—and otherwise $\mathcal{P}(A \cup B) < \mathcal{P}(A) + \mathcal{P}(B)$.

The triplet $(\Omega, \mathcal{F}, \mathcal{P})$ is called "probability space".

Consider then such a probability space. A one-dimensional random variable X is a function from Ω to $\mathbb{R}$ with the property

$$\forall x \in \mathbb{R}, \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F},$$

i.e. the set of all outcomes ω, whose realization $X(\omega)$ is smaller than x, is an event.

A *cumulative distribution function* F from $\mathbb{R}$ to $[0,1]$ is then associated to this random variable, which applies the real number x onto the probability $\mathcal{P}(X \leq x) \equiv \mathcal{P}(\{\omega \in \Omega \mid X(\omega) \leq x\})$. One then has

$$F(x) \equiv \mathcal{P}(X \leq x) = \int_{-\infty}^{x^+} p_X(x')\,\mathrm{d}x',$$

with $p_X(x)$ the probability density. (The notation x^+ means that when p_X contains a contribution $\delta(x)$, then the latter is also taken into account in the integral.)

B.2 Averages and moments

Besides the sample space Ω and the probability density p_X, other notions may be employed for the characterization of a random variable X. In this section we restrict the discussion to one-dimensional stochastic variables—multidimensional ones will be dealt with in Sec. B.4.

Consider a function f defined on the one-dimensional sample space Ω of a random variable X. The *expectation value* or *average value* of f is defined by

$$\boxed{\langle f(X)\rangle \equiv \int_\Omega f(x)\, p_X(x)\,\mathrm{d}x.} \tag{B.3}$$

The generalization of this definition to the case of multidimensional random variables is straightforward.

Remarks:

$*$ This expectation value is also denoted by $E\big(f(X)\big)$, in particular by mathematicians.

∗ Averaging a function is a linear operation.

The m-th *moment*—or “moment of order m”—of a one-dimensional random variable X (or equivalently of its probability distribution) is defined as the average value

$$\mu_m \equiv \langle X^m \rangle . \tag{B.4}$$

In particular, μ_1 is the expectation value of the random variable. In analogy with the arithmetic mean, μ_1 is often referred to as “mean value”.

In addition, the *variance* of the probability distribution is defined by

$$\sigma^2 \equiv \big\langle (X - \langle X \rangle)^2 \big\rangle = \mu_2 - \mu_1^2. \tag{B.5}$$

The positive square root σ is called *standard deviation*. The latter is often loosely referred to as “fluctuation”, because σ constitutes a typical measure for the dispersion of the realizations of a random variable about its expectation value, i.e. for the scale of the fluctuations of the quantity described by the random variable.

Remarks:

∗ The integral defining the m-th moment of a probability distribution might possibly diverge! See for instance the Cauchy–Lorentz distribution in Sec. B.3.7 below.

∗ In analogy to the variance (B.5), one also defines the m-th *central moment* (or *m-th moment about the mean*) $\big\langle (X - \langle X \rangle)^m \big\rangle$ for arbitrary m.

∗ If the random variable possesses a (physical) dimension, then its moments are also dimensioned quantities.

A further useful notion is that of the *characteristic function*, defined for $k \in \mathbb{R}$ by

$$G_X(k) \equiv \langle \mathrm{e}^{\mathrm{i}kX} \rangle = \int \mathrm{e}^{\mathrm{i}kx} p_X(x)\, \mathrm{d}x. \tag{B.6a}$$

One easily checks that the Taylor expansion of $G(x)$ about $k = 0$ reads

$$G_X(k) = \sum_{m=0}^{\infty} \frac{(\mathrm{i}k)^m}{m!} \mu_m, \tag{B.6b}$$

i.e. the m-th derivative of the characteristic function at the point $k = 0$ is related to the m-th moment of the probability distribution.

Remarks:

∗ More precisely, the *moment-generating function* is $\mathcal{G}_X(k) \equiv G_X(-\mathrm{i}k)$, whose successive derivatives at $k = 0$ are exactly equal to the moments μ_m. However, this moment-generating function may not always exist—e.g. in the case of the Cauchy–Lorentz distribution—, while the characteristic function always does.

∗ The logarithm of G_X (or $\mathcal{G}_X$) generates the successive *cumulants* κ_m of the probability distribution:

$$\ln G_X(k) = \sum_{m=1}^{\infty} \frac{(\mathrm{i}k)^m}{m!} \kappa_m, \tag{B.7}$$

which are sometimes more useful than the moments (see Sec. B.4). One easily checks for instance $\kappa_1 = \mu_1 = \langle X \rangle$ and $\kappa_2 = \sigma^2$.

B.3 Some usual probability distributions

In this section, we list some often encountered probability distributions together with a few of their properties, starting with discrete ones, before going on with continuous densities.

B.3.1 Discrete uniform distribution

Consider a random variable X with the finite discrete sample space $\Omega = \{x_1, \ldots, x_N\}$ where $N \in \mathbb{N}^*$. The *discrete uniform distribution*

$$p_n = \frac{1}{N} \quad \forall n \in \{1, 2, \ldots, N\} \tag{B.8}$$

corresponds to the case where all realizations of the random variable are equally probable.

The expectation value is then $\langle X \rangle = \dfrac{1}{N} \displaystyle\sum_{n=1}^{N} x_n$ and the variance $\sigma^2 = \dfrac{1}{N^2} \displaystyle\sum_{n=1}^{N} x_n^2 - \dfrac{1}{N} \left(\displaystyle\sum_{n=1}^{N} x_n \right)^2$.

B.3.2 Binomial distribution

Let p be a real number, $0 < p < 1$, and $N \in \mathbb{N}$ an integer.

The *binomial distribution* with parameters N and p is the probability distribution for a random variable with sample space $\Omega = \{1, 2, \ldots, n, \ldots N\}$ given by

$$p_n = \binom{N}{n} p^n (1-p)^{N-n}. \tag{B.9}$$

p_n is the probability that, when a random experiment with two possible outcomes ("success" and "failure", with respective probabilities p and $1 - p$) is repeated N times, one obtains exactly n "successes".

The expectation value is $\langle X \rangle = pN$ and the variance $\sigma^2 = Np(1-p)$.

B.3.3 Poisson distribution

Let λ be a positive real number. The *Poisson distribution* with parameter λ associates to the integer $n \in \mathbb{N} = \Omega$ (sample space) the probability

$$p_n = \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!}. \tag{B.10}$$

The corresponding average value and variance are $\langle X \rangle = \sigma^2 = \lambda$.

B.3.4 Continuous uniform distribution

Consider a continuous random variable X whose sample space Ω is the real interval $]a, b]$ with $a < b$. A constant probability density

$$p_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a < x \leq b \\ 0 & \text{otherwise} \end{cases} \tag{B.11}$$

on this range represents an instance of *continuous uniform distribution*. This is quite obviously the generalization to the continuous case of the discrete uniform distribution (B.8).

B.3.5 Gaussian distribution

Let $\Omega = \mathbb{R}$ represent the sample space for a continuous random variable X. The probability density

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{B.12}$$

is called *Gaussian distribution* (or *normal distribution*).

The average value is $\langle X\rangle = \mu$ and the variance σ^2. One can also easily check that the cumulants κ_m of all orders $m \geq 3$ vanish identically.

B.3.6 Exponential distribution

Let λ be a positive real number. A continuous random variable X with sample space $\Omega = \mathbb{R}_+$ is said to obey the *exponential distribution* with parameter λ if its probability density is given by

$$p_X(x) = \lambda\, \mathrm{e}^{-\lambda x}. \tag{B.13}$$

The expectation value is $\langle X\rangle = \dfrac{1}{\lambda}$ and the variance $\sigma^2 = \dfrac{1}{\lambda^2}$.

B.3.7 Cauchy–Lorentz distribution

Let x_0 and γ be two real numbers, with $\gamma > 0$. The *Cauchy–Lorentz distribution*, also called in physics (non-relativistic) *Breit*(cf)*–Wigner*(cg) *distribution* or shortly *Lorentzian*, is given by

$$p_X(x) = \frac{1}{\pi}\frac{\gamma}{(x-x_0)^2+\gamma^2} \tag{B.14}$$

for $x \in \Omega = \mathbb{R}$.

All moments of this distribution diverge! x_0 is the position of the maximum of the distribution—i.e. it corresponds to the most probable value of the random variable—, while 2γ represents the full width at half maximum (often abbreviated FWHM).

B.4 Multidimensional random variables

Let $\mathbf{X}$ be a D-dimensional random variable, whose components will be denoted $X_1, X_2, \ldots, X_D$. For the sake of brevity, we shall hereafter only consider the case of continuous variables.

B.4.1 Definitions

The probability density $p_{\mathbf{X}}(\mathbf{x})$ is also called *multivariate* or *joint probability density* of the D random variables. For commodity, we shall also denote this density by $p_D(x_1, \ldots, x_D) \equiv p_{\mathbf{X}}(\mathbf{x})$.

B.4.1 a Moments and averages

The *moments* of a multivariate probability density are defined as the expectation values

$$\left\langle X_1^{m_1} X_2^{m_2} \cdots X_D^{m_D}\right\rangle \equiv \int x_1^{m_1} x_2^{m_2} \cdots x_D^{m_D}\, p_D(x_1, \ldots, x_D)\, \mathrm{d}x_1\, \mathrm{d}x_2 \cdots \mathrm{d}x_D. \tag{B.15}$$

They are generated by the *characteristic function*

$$G_{\mathbf{X}}(k_1, \ldots, k_D) \equiv \left\langle \mathrm{e}^{\mathrm{i}(k_1X_1+\cdots+k_DX_D)}\right\rangle, \tag{B.16}$$

with $k_1, \ldots, k_D$ real auxiliary variables. The logarithm of this characteristic function generates the corresponding *cumulants*.

(cf) G. Breit, 1899–1981 (cg) E. P. Wigner, 1902–1995

Combinations of moments which plays an important role are the *covariances*

$$\big\langle\big(X_i - \langle X_i\rangle\big)\big(X_j - \langle X_j\rangle\big)\big\rangle = \langle X_i X_j\rangle - \langle X_i\rangle\langle X_j\rangle \tag{B.17}$$

for every i, $j \in \{1,\dots,D\}$. These are often combined into a symmetric *covariance matrix*, of which they constitute the elements. One easily checks that the covariances are actually the second-order cumulants of the joint probability distribution.

Useful, dimensionless measures are then the *correlation coefficients* obtained by dividing the covariance of random variables X_i, X_j by the product of their standard deviations (B.5)

$$c_{ij} \equiv \frac{\langle X_i X_j\rangle - \langle X_i\rangle\langle X_j\rangle}{\sqrt{\sigma_{X_i}^2 \sigma_{X_j}^2}}. \tag{B.18}$$

Obviously, the diagonal coefficients c_{ii} are identically equal to 1.

If the covariance—or equivalently the correlation coefficient—of two random variables X_i and X_j vanishes, then these variables are said to be *uncorrelated*.

B.4.1 b Marginal and conditional probability distributions

Consider an integer $r < D$ and choose r random variables among $X_1, X_2, \dots, X_D$—for the sake of simplicity, the first r ones X_1, …, X_r. The probability that the latter take values in the intervals $[x_1, x_1 + \mathrm{d}x_1]$, …, $[x_r, x_r + \mathrm{d}x_r]$, *irrespective* of the values taken by X_{r+1}, …, X_D, is

$$p_r(x_1,\dots,x_r)\,\mathrm{d}x_1\cdots\mathrm{d}x_r = \left[\int p_D(x_1,\dots,x_r,x_{r+1},\dots,x_D)\,\mathrm{d}x_{r+1}\cdots\mathrm{d}x_D\right]\mathrm{d}x_1\cdots\mathrm{d}x_r,$$

where the integral runs over the $(D-r)$-dimensional sample space of the variables $X_{r+1},\dots,X_D$, which have thus been "integrated out". The density

$$\boxed{p_r(x_1,\dots,x_r) \equiv \int p_D(x_1,\dots,x_r,x_{r+1},\dots,x_D)\,\mathrm{d}x_{r+1}\cdots\mathrm{d}x_D} \tag{B.19}$$

for the remaining random variables $X_1,\dots,X_r$ is then called *marginal distribution*.

If the random variables $X_{r+1},\dots,X_D$ take given realizations $x_{r+1},\dots,x_D$, then one can consider the probability distribution for the remaining random variables under this condition. Accordingly, one introduces the corresponding *conditional probability density*

$$p_{r|D-r}(x_1,\dots,x_r\,|\,x_{r+1},\dots,x_D). \tag{B.20}$$

One has then the identities

$$\begin{aligned} p_D(x_1,\dots,x_D) &= p_{r|D-r}(x_1,\dots,x_r\,|\,x_{r+1},\dots,x_D)\,p_{D-r}(x_{r+1},\dots,x_D)\\ &= p_{D-r|r}(x_{r+1},\dots,x_D\,|\,x_1,\dots,x_r)\,p_r(x_1,\dots,x_r). \end{aligned}$$

The first identity can be rewritten as

$$\boxed{p_{r|D-r}(x_1,\dots,x_r\,|\,x_{r+1},\dots,x_D) = \frac{p_D(x_1,\dots,x_D)}{p_{D-r}(x_{r+1},\dots,x_D)},} \tag{B.21}$$

which constitutes *Bayes'*[(ch)] *theorem*.

B.4.2 Statistical independence

When the identity

$$p_D(x_1,\dots,x_D) = p_r(x_1,\dots,x_r)\,p_{D-r}(x_{r+1},\dots,x_D) \tag{B.22}$$

holds for all realizations $x_1,\dots,x_r,x_{r+1},\dots,x_D$ of the random variables, then the sets of variables

---

[(ch)] T. Bayes, 1702–1761

$\{X_1, \ldots, X_r\}$ and $\{X_{r+1}, \ldots, X_D\}$ are said to be *statistically independent* (or shortly *independent*). In that case, the marginal probability distribution for X_1, …, X_r (resp. for X_{r+1}, …, X_D) equals the conditional distribution:

$$p_r(x_1, \ldots, x_r) = p_{r|D-r}(x_1, \ldots, x_r \,|\, x_{r+1}, \ldots, x_D).$$

Let X_1 and X_2 be two statistically independent random variables. For all functions f_1, f_2 defined on the respective sample spaces,(98) one then has the identity $\langle f_1(X_1)\, f_2(X_2)\rangle = \langle f_1(X_1)\rangle \langle f_2(X_2)\rangle$. In particular, all moments obey

$$\langle X_1^{m_1} X_2^{m_2}\rangle = \langle X_1^{m_1}\rangle \langle X_2^{m_2}\rangle \quad \forall m_1, m_2,$$

as one sees by considering the characteristic functions of the random variables.

The latter equation shows that the statistical independence of two random variables implies that they are uncorrelated. The converse is not however true, although both notions are often taken as identical.

B.4.3 Addition of random variables

Consider again two random variables X_1, X_2 defined on the same sample space, whose joint probability density is denoted by $p_{\mathbf{X}}(x_1, x_2)$.

Their sum $Y = X_1 + X_2$ constitutes a new random variable with the expectation value

$$\langle Y\rangle = \langle X_1\rangle + \langle X_2\rangle \tag{B.23}$$

and more generally the probability density

$$\begin{aligned} p_Y(y) &= \int p_{\mathbf{X}}(x_1, x_2)\, \delta(y - x_1 - x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 \\ &= \int p_{\mathbf{X}}(x_1, y - x_1)\, \mathrm{d}x_1 = \int p_{\mathbf{X}}(y - x_2, x_2)\, \mathrm{d}x_2. \end{aligned} \tag{B.24}$$

This corresponds to the characteristic function

$$G_Y(k) = G_{X_1,X_2}(k, k), \tag{B.25}$$

where $G_{X_1,X_2}(k_1, k_2)$ is the generating function for $p_{\mathbf{X}}(x_1, x_2)$.

If X_1 and X_2 are statistically independent, then Eq. (B.22) allows one to simplify Eq. (B.24) into

$$p_Y(y) = \int p_{X_1}(x_1)\, p_{X_2}(y - x_1)\, \mathrm{d}x_1 = \int p_{X_1}(y - x_2)\, p_{X_2}(x_2)\, \mathrm{d}x_2,$$

that is, into the convolution of p_{X_1} with p_{X_2}. In this case, the variance of Y is

$$\sigma_Y^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2. \tag{B.26}$$

This property generalizes to all cumulants of Y (however, not to its central moments!), as follows at once from the identity

$$G_Y(k) = G_{X_1}(k)\, G_{X_2}(k).$$

Remark: The latter equation shows at once that the sum of two Gaussian variables—and more generally, any linear combination of Gaussian variables—is itself a Gaussian variable.

---

(98) … and whose product can be defined in some way, in case the functions are neither real- nor complex-valued, as e.g. the scalar product of two vectors.

B.4.4 Multidimensional Gaussian distribution

Let A be a positive definite, symmetric $D \times D$ matrix and $\mathbf{B}$ a D-dimensional vector. The multivariate Gaussian distribution for random variables $\mathbf{X} = (X_1, \dots, X_D)$ is given by

$$p_{\mathbf{X}}(\mathbf{x}) = \sqrt{\frac{\det \mathsf{A}}{(2\pi)^D}}\, \mathrm{e}^{-\frac{1}{2}\mathbf{B}^T\cdot\mathsf{A}^{-1}\cdot\mathbf{B}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{D} A_{ij}x_ix_j - \sum_{i=1}^{D} B_ix_i\right], \tag{B.27}$$

with A_{ij} resp. B_i the elements of A resp. the components of $\mathbf{B}$, while $\mathbf{B}^T$ denotes the transposed vector of $\mathbf{B}$.

B.5 Central limit theorem

Consider a sequence $(X_1, X_2, \dots, X_n, \dots)$ of statistically independent one-dimensional random variables with the same sample space Ω and the same probability distribution. One assumes that both the expectation value μ and the variance σ^2 of the distribution exist. Let

$$Z_N \equiv \frac{1}{N}\sum_{n=1}^{N} X_n \tag{B.28a}$$

denote the N-th partial sum of these random variables, multiplied with an adequate normalization factor. Following the results of Sec. B.4.3, the expectation value of Z_N exactly equals μ while the variance is σ^2/N.

According to the *central limit theorem*,(99) the probability distribution for the random variable $\sqrt{N}(Z_N - \mu)$ converges for $N \to \infty$ towards the Gaussian distribution with expectation value $\mu_1 = 0$ and variance σ^2, i.e. for every real number z

$$p_{Z_N}(z) \underset{N\gg 1}{\sim} \frac{1}{\sqrt{2\pi\sigma^2/N}} \exp\left[-\frac{(z-\mu)^2}{2\sigma^2/N}\right]. \tag{B.28b}$$

This theorem underlies the important role of the Gaussian distribution and is related to the law of large numbers. Since the variance of the distribution of Z_N becomes smaller with increasing N, the possible realizations z become more and more concentrated about the expectation value μ: the distribution approaches a δ-distribution at the point μ.

Remarks:

∗ The convergence in Eq. (B.28b) is actually a weak convergence, or "convergence in distribution", analogous to the pointwise convergence of "usual" (i.e. non-stochastic) sequences.

∗ There exist further analogous theorems (the version above is the theorem of Lindeberg[ci]–Lévy[cj]) for statistically independent random variables with different probability distributions, for "nearly independent" random variables...

Elements of a proof:
The probability density for Z_N follows from the generalizations of Eqs. (B.24) and (B.22)

$$p_{Z_N}(z) = \int p_{X_1}(x_1)\cdots p_{X_N}(x_N)\,\delta\!\left(z - \frac{1}{N}\sum_{n=1}^{N} x_n\right) \mathrm{d}x_1\cdots\mathrm{d}x_N,$$

where $p_{X_1}, \dots, p_{X_N}$ actually all reduce to the same density p_X. Inserting the Fourier representation of the δ distribution, this becomes

---

(99)... in its simplest incarnation.

[ci] J. W. Lindeberg, 1876–1932 [cj] P. Lévy, 1886–1971

$$\begin{aligned}
p_{Z_N}(z) &= \int p_{X_1}(x_1) \cdots p_{X_N}(x_N) \, \exp\!\left[\mathrm{i}k \left(\frac{1}{N} \sum_{n=1}^{N} x_n - z \right) \right] \mathrm{d}x_1 \cdots \mathrm{d}x_N \, \frac{\mathrm{d}k}{2\pi} \\
&= \int \mathrm{e}^{-\mathrm{i}kz} \prod_{n=1}^{N} \left(\int p_X(x_n) \, \mathrm{e}^{\mathrm{i}kx_n/N} \, \mathrm{d}x_n \right) \frac{\mathrm{d}k}{2\pi} = \int \mathrm{e}^{-\mathrm{i}kz} \left[G_X\!\left(\frac{k}{N} \right) \right]^N \frac{\mathrm{d}k}{2\pi},
\end{aligned}$$

where G_X is the characteristic function (B.6a). A Taylor expansion of the latter at $k = 0$ yields

$$G_X\!\left(\frac{k}{N} \right) = 1 + \frac{\mathrm{i}k\mu}{N} - \frac{k^2 \langle X^2 \rangle}{2N^2} + \mathcal{O}\!\left(\frac{1}{N^3} \right),$$

i.e.

$$N \ln G_X\!\left(\frac{k}{N} \right) = \mathrm{i}k\mu - \frac{k^2 \sigma^2}{2N} + \mathcal{O}\!\left(\frac{1}{N^2} \right).$$

This eventually gives

$$p_{Z_N}(z) \underset{N \gg 1}{\sim} \int \exp\!\left[-\frac{k^2 \sigma^2}{2N} - \mathrm{i}k(z - \mu) \right] \frac{\mathrm{d}k}{2\pi} = \frac{1}{\sqrt{2\pi\sigma^2/N}} \exp\!\left[-\frac{(z - \mu)^2}{2\sigma^2/N} \right]. \qquad \square$$

Appendix C

Basic notions on stochastic processes

Similar to the previous one, this Appendix introduces further notions of probability theory, namely now some basic definitions and results on stochastic processes.

C.1 Definitions

Consider a random variable X with sample space Ω and probability distribution p_X. Let t denote an additional variable, which takes an infinite number (countable or not) of values in some set $\mathcal{I}$. Any function f on the product $\mathcal{I} \times \Omega$ defines an infinite number of stochastic variables

$$Y_X(t) = f(t, X) \tag{C.1}$$

labeled by t. Such a quantity is referred to as a *random function* of the variable t. In case the latter stands for time, $Y_X(t)$ is called a *stochastic process*.
Taking at every t a realization x of the random variable X, one obtains a *realization* of the process or *sample function*

$$Y_x(t) = f(t, x), \tag{C.2}$$

which is a function in the usual sense of analysis. In turn, fixing $t \in \mathcal{I}$, $Y_X(t)$ is a random variable in the sense of App. B.

The random function can be multidimensional, $\mathbf{Y}_X(t) = Y_X^1(t), Y_X^2(t), \dots, Y_X^D(t)$. This is in particular often the case when the random variable itself is multidimensional, $\mathbf{X}$.

> In this appendix, the random functions we shall consider will take their values in (subsets of) $\mathbb{R}$ (in the one-dimensional case) or $\mathbb{R}^D$ with $D > 1$. The results can be extended to further spaces, provided a product can be defined on them, so that e.g. the integrand of Eq. (C.4) makes sense.

In turn, there might be more than one additional variable, that is the random function is labeled by a multidimensional variable. In physics, this corresponds for instance to the case of random fields, whose value is a stochastic variable at each instant and at each point in space.

For the sake of brevity, we shall hereafter refer to the variable t as "time", and assume that it takes its values in (an interval of) $\mathbb{R}$: we thus consider *continuous-time stochastic processes*. The points of the t-axis will be referred to as "instants".

C.1.1 Averages and moments

Using the probability distribution $p_X(x)$ of the random variable, one easily defines averages as in Sec. B.2. For instance, the (single-time) *sample average* or *ensemble average* of the stochastic process $Y_X(t)$ is given by

$$\langle Y_X(t) \rangle \equiv \int_\Omega Y_x(t)\, p_X(x)\, \mathrm{d}x, \tag{C.3a}$$

where the integration runs over the sample space Ω.

It is sometimes helpful to view this sample average differently: let $Y_{x^{(r)}}(t)$ with $r = 1, 2, \ldots, N$ denote different realizations of the process at the same instant t, where the corresponding realizations $x^{(r)}$ of the random variable X are distributed according to p_X. Then the sample average is given by

$$\big\langle Y_X(t)\big\rangle = \lim_{N\to\infty}\frac{1}{N}\sum_{r=1}^{N} Y_{x^{(r)}}(t), \tag{C.3b}$$

in accordance with the definition of the probability distribution p_X.

More generally, one can define multi-time averages, or higher moments, as follows. Let $n \in \mathbb{N}^*$, and consider n (not necessarily different) values $t_1, t_2, \ldots, t_n$ of the time variable. The n-th moment is then defined as

$$\big\langle Y_X(t_1)Y_X(t_2)\cdots Y_X(t_n)\big\rangle \equiv \int_\Omega Y_x(t_1)Y_x(t_2)\cdots Y_x(t_n)\, p_X(x)\,\mathrm{d}x. \tag{C.4}$$

By combining first and second moments, one obtains the *autocorrelation function*

$$\kappa(t_1,t_2) \equiv \big\langle\big[Y_X(t_1) - \big\langle Y_X(t_1)\big\rangle\big]\big[Y_X(t_2) - \big\langle Y_X(t_2)\big\rangle\big]\big\rangle \tag{C.5a}$$

$$= \big\langle Y_X(t_1)Y_X(t_2)\big\rangle - \big\langle Y_X(t_1)\big\rangle\big\langle Y_X(t_2)\big\rangle. \tag{C.5b}$$

In case the random function is multidimensional, this autocorrelation function is replaced by the *correlation matrix*

$$\kappa^{ij}(t_1,t_2) \equiv \big\langle\big[Y_X^i(t_1) - \big\langle Y_X^i(t_1)\big\rangle\big]\big[Y_X^j(t_2) - \big\langle Y_X^j(t_2)\big\rangle\big]\big\rangle, \tag{C.6}$$

whose diagonal coefficients are *autocorrelations*, while the off-diagonal elements are referred to as *cross-correlations*.

Generalizing the concept of characteristic function for a random variable [see Eqs. (B.6a) and (B.16)], one defines for a given stochastic process $Y_X(t)$ the *characteristic functional*

$$G_{Y_X}[k(t)] \equiv \left\langle \exp\left[\mathrm{i}\int k(t)\,Y_X(t)\,\mathrm{d}t\right]\right\rangle, \tag{C.7}$$

where the integral runs over the space over which the time variable t takes its values, while $k(t)$ is a test function defined over this space. One easily checks that expanding this characteristic functional in powers of k yields the n-time averages (C.4) as functional derivatives.

C.1.2 Distribution functions

Consider a stochastic process $Y_X(t)$. The probability density that $Y_X(t)$ takes the value y at time t, also called *single-time density*, is trivially given by

$$p_{Y,1}(t,y) \equiv \int_\Omega \delta\big(y - Y_x(t)\big)\,p_X(x)\,\mathrm{d}x. \tag{C.8}$$

Introducing now *different* instants $t_1, t_2, \ldots, t_n$ with $n > 1$, the joint probability for Y_X to take the value y_1 at t_1, the value y_2 at t_2, …, and the value y_n at t_n is given by

$$p_{Y,n}(t_1,y_1;t_2,y_2;\ldots;t_n,y_n) \equiv \int_\Omega \delta\big(y_1 - Y_x(t_1)\big)\,\delta\big(y_2 - Y_x(t_2)\big)\cdots\delta\big(y_n - Y_x(t_n)\big)\,p_X(x)\,\mathrm{d}x. \tag{C.9}$$

$p_{Y,n}$ is referred to as *n-time density* or *n-point density*. With its help, the n-th moment (C.4) can be rewritten as

$$\boxed{\big\langle Y_X(t_1)Y_X(t_2)\cdots Y_X(t_n)\big\rangle = \int y_1y_2\cdots y_n\,p_{Y,n}(t_1,y_1;t_2,y_2;\ldots;t_n,y_n)\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n,} \tag{C.10}$$

where the integral runs over (n copies of the) space on which the realizations of $Y_X(t)$ take their values.

One easily checks that the n-point densities satisfy the following four properties:

- $p_{Y,n}(t_1,y_1;t_2,y_2;\dots;t_n,y_n)\geq 0$ for every $n\geq 1$ and for all $(t_1,y_1),\dots,(t_n,y_n)$. (C.11a)
- $p_{Y,n}$ is symmetric under the exchange or two pairs (t_j,y_j) and (t_k,y_k) for all j,k. (C.11b)
- The densities obey for all $m<n$ and for all $(t_1,y_1),\dots,(t_m,y_m)$ the consistency conditions

$$\boxed{\begin{aligned}&p_{Y,m}(t_1,y_1;t_2,y_2;\dots;t_m,y_m)=\\ &\qquad\int p_{Y,n}(t_1,y_1;\dots;t_m,y_m;t_{m+1},y_{m+1};\dots;t_n,y_n)\,\mathrm{d}y_{m+1}\cdots\,\mathrm{d}y_n.\end{aligned}}\tag{C.11c}$$

  That is, every $p_{Y,n}$ encompasses all information contained in all $p_{Y,m}$ with $m<n$.
- The single-time density $p_{Y,1}$ is normalized to unity: $\int p_{Y,1}(t,y)\,\mathrm{d}y=1$. (C.11d)

Remarks:

∗ Property (C.11b) allows one to order the time arguments at will.

∗ The definition of the densities need not be extended to the case where two or more of the time arguments, say t_j and t_k, are equal, since it in that case, only $y_j=y_k$ is meaningful—the probability that the process takes two different values at the same instant is obviously zero.

∗ Relation (C.11c) expresses $p_{Y,m}$ as a marginal distribution of $p_{Y,n}$, cf. Eq.(B.19).

∗ Starting from the normalization (C.11d) and using Eq. (C.11c), one easily proves recursively that every n-point density $p_{Y,n}$ is normalized to unity as well

$$\int p_{Y,n}(t_1,y_1;t_2,y_2;\dots;t_n,y_n)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\cdots\,\mathrm{d}y_n=1.\tag{C.12}$$

Together with the positivity condition (C.11a), this means that the n-point densities possess the "good properties" (B.1) of probability distributions.

Conditional n-point densities

One also introduces *conditional probability densities*, by considering the probability density that Y_X takes the value y_1 at t_1, the value y_2 at t_2, …, and the value y_m at t_m, knowing that it takes the value y_{m+1} at t_{m+1}, the value y_{m+2} at t_{m+2}, …, and the value y_n at t_n:

$$p_{Y,m|n-m}(t_1,y_1;\dots;t_m,y_m|\,t_{m+1},y_{m+1};\dots;t_n,y_n)=\frac{p_{Y,n}(t_1,y_1;\dots;t_m,y_m;t_{m+1},y_{m+1};\dots;t_n,y_n)}{p_{Y,n-m}(t_{m+1},y_{m+1};\dots;t_n,y_n)}\tag{C.13}$$

[cf. Bayes' theorem (B.21)].

Remarks:

∗ Working recursively, one finds that every n-point density can be expressed as the product of conditional probability densities $p_{Y,1|m}$, with m ranging from $n-1$ to 1, and of a single-time density:

$$\begin{aligned}p_{Y,n}(t_1,y_1;\dots;t_n,y_n)=\;&p_{Y,1|n-1}(t_n,y_n|\,t_1,y_1;\dots;t_{n-1},y_{n-1})\\ &\times p_{Y,1|n-2}(t_{n-1},y_{n-1}|\,t_1,y_1;\dots;t_{n-2},y_{n-2})\cdots p_{Y,1|1}(t_2,y_2|\,t_1,y_1)\\ &\times p_{Y,1}(t_1,y_1),\end{aligned}\tag{C.14}$$

which is easily interpreted.

$*$ Writing down the previous identity for $n = 3$, integrating over y_2 under consideration of the consistency condition (C.11c), and eventually dividing by $p_{Y,1}(t_1, y_1)$, one finds

$$p_{Y,1|1}(t_3, y_3 \,|\, t_1, y_1) = \int p_{Y,1|2}(t_3, y_3 \,|\, t_1, y_1; t_2, y_2)\, p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)\, \mathrm{d}y_2. \tag{C.15}$$

Again, this identity has an intuitive meaning. Mathematically, it is an integral-functional equation for the conditional probability $p_{Y,1|1}$, involving the integration kernel $p_{Y,1|2}$. Similarly, one can write down an analogous equation for $p_{Y,1|2}$, with $p_{Y,1|3}$ as integration kernel; and more generally, a whole hierarchy of integral-functional relations, where the equation for $p_{Y,1|n}$ admits $p_{Y,1|n+1}$ as integration kernel.

C.2 Some specific classes of stochastic processes

The knowledge of all n-point probability densities $p_{Y,n}$ for a random function $Y_X(t)$ allows the computation of all n-point averages and thus replaces the knowledge of the probability density p_X. Accordingly, we shall from now on drop any reference to the random variable X and denote a stochastic process more simply as $Y(t)$, and its realizations as $y(t)$.

C.2.1 Centered processes

A stochastic process $Y(t)$ is called *centered* if its single-time average $\langle Y(t)\rangle$ is identically vanishing, for any time t.

Given an arbitrary stochastic process $Y(t)$, the process $Z(t) \equiv Y(t) - \langle Y(t)\rangle$ is obviously centered. One checks at once that $Y(t)$ and the associated process $Z(t)$ share the same autocorrelation function $\kappa(t_1, t_2)$.

C.2.2 Stationary processes

A stochastic process $Y(t)$ is said to be *stationary* when all its moments are invariant under arbitrary shifts of the origin of times, that is when for all $n \in \mathbb{N}^*$, $\Delta t \in \mathbb{R}$ and n-uplets t_1, t_2, ..., t_n, one has the identity

$$\big\langle Y(t_1+\Delta t)Y(t_2+\Delta t)\cdots Y(t_n+\Delta t)\big\rangle = \big\langle Y(t_1)Y(t_2)\cdots Y(t_n)\big\rangle. \tag{C.16a}$$

In particular, the single-time average $\langle Y(t)\rangle$ is time-independent, so that it is convenient to work with the associated centered process $Y(t) - \langle Y\rangle$.

Remark: An equivalent definition is that all n-point densities of the process are invariant under arbitrary time translations:

$$p_{Y,n}(t_1+\Delta t, y_1; t_2+\Delta t, y_2; \ldots; t_n+\Delta t, y_n) = p_{Y,n}(t_1, y_1; t_2, y_2; \ldots; t_n, y_n). \tag{C.16b}$$

The autocorrelation function $\kappa(t_1, t_2)$ of a stationary process only depends on the time difference $\tau \equiv t_2 - t_1$, and is an even function of τ (i.e. it only depends on $|\tau|$):

$$\kappa(\tau) = \big\langle Y(t)Y(t+\tau)\big\rangle - \big\langle Y\big\rangle^2. \tag{C.17}$$

A widespread case in physics is that of processes whose autocorrelation function only takes significant values over some scale $|\tau| \lesssim \tau_c$—the *autocorrelation time*—, and become negligible for $|\tau| \gg \tau_c$.

For a centered stationary multidimensional stochastic process $\mathbf{Y}(t)$ with components $Y^1(t)$, $Y^2(t)$, ..., defining [cf. the correlation matrix (C.6)]

$$\kappa^{ij}(\tau) \equiv \big\langle Y^i(t)Y^j(t+\tau)\big\rangle, \tag{C.18a}$$

which is not necessarily even in τ, one has the obvious property

$$\kappa^{ij}(\tau) = \kappa^{ji}(-\tau). \tag{C.18b}$$

Stationary processes are conveniently characterized by their spectral properties, which follow from considering their (discrete) Fourier transform. This idea will be further discussed in Sec. C.3.

C.2.3 Ergodic processes

A stationary stochastic process $Y(t)$ is called *ergodic*, when any single realization $y(t)$ contains all statistical information on the whole process, i.e. allows one to compute all possible n-point averages.

Let $y(t)$ denote a given realization. The time average of the stationary process over the finite interval $[t-\frac{\mathcal{T}}{2}, t+\frac{\mathcal{T}}{2}]$, where $\mathcal{T} > 0$, is defined as

$$\overline{Y(t)}^{\mathcal{T}} \equiv \frac{1}{\mathcal{T}} \int_{t-\frac{\mathcal{T}}{2}}^{t+\frac{\mathcal{T}}{2}} y(t')\,\mathrm{d}t'. \tag{C.19}$$

This average depends on t, $\mathcal{T}$ and the realization y. In the limit of large $\mathcal{T}$, the average becomes the *time average* $\overline{Y}$ of the process,

$$\overline{Y} = \lim_{\mathcal{T}\to+\infty} \overline{Y(t)}^{\mathcal{T}} \equiv \lim_{\mathcal{T}\to+\infty} \frac{1}{\mathcal{T}} \int_{t-\frac{\mathcal{T}}{2}}^{t+\frac{\mathcal{T}}{2}} y(t')\,\mathrm{d}t'. \tag{C.20}$$

As hinted at by the notation, $\overline{Y}$ no longer depends on t and $\mathcal{T}$, thanks to the assumed stationarity; yet it still depends on the specific realization of the process. If it is independent of the realization, then the time average $\overline{Y}$ is equal to the (time-independent, since $Y(t)$ is stationary) ensemble average $\langle Y \rangle$.

A stochastic process is ergodic when the identity between time average and sample average holds for all products of $Y(t)$ at different times, i.e. for all moments.

C.2.4 Gaussian processes

A stochastic process $Y(t)$ is called *Gaussian process* if all its n-point densities (C.9) are Gaussian distributions. Equivalently, for each integer n and each choice of arbitrary instants t_1, t_2, …, t_n, the n-dimensional random variable with components $Y(t_1)$, …, $Y(t_n)$ is Gaussian-distributed.

The corresponding characteristic functional reads

$$G_Y[k(t)] = \exp\left[\mathrm{i}\int k(t)\big\langle Y(t)\big\rangle \mathrm{d}t - \frac{1}{2}\int k(t_1)k(t_2)\,\kappa(t_1,t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\right], \tag{C.21}$$

so that the process is entirely determined by its single-time average $\langle Y(t)\rangle$ and its autocorrelation function $\kappa(t_1,t_2)$—or equivalently, by the single- and two-time densities $p_{Y,1}$, $p_{Y,2}$. For instance, one can show(100) that for even n, the n-point moment is given by

$$\big\langle Y(t_1)Y(t_2)\cdots Y(t_n)\big\rangle = \sum \big\langle Y(t_j)Y(t_k)\big\rangle \cdots \big\langle Y(t_l)Y(t_m)\big\rangle,$$

where the sum runs over all possible pairings of the indices 1, 2, …, n, while the product for a given pairing involves all $n/2$ corresponding pairs.

If $Y(t)$ is a Gaussian process, then the associated centered process $Z(t) \equiv Y(t) - \langle Y(t)\rangle$ is also Gaussian, and all moments of odd order of $Z(t)$ vanish.

---

(100) This is (part of) the Isserlis'(ck) theorem, better known in physics as Wick's(cl) theorem.

(ck) L. Isserlis, 1881–1966 (cl) G.-C. Wick, 1909–1992

C.2.5 Markov processes

We now introduce a class of stochastic processes which are often encountered in physics—or, one should rather say, which are often used to model physical phenomena due to their simplicity, since they are entirely determined by the two densities $p_{Y,1}$ and $p_{Y,1|1}$.

C.2.5 a Markov property

A *Markov*$^{(101)}$ *process* is a stochastic process $Y(t)$ for which for all $n \in \mathbb{N}^*$ and arbitrary ordered times $t_1 < t_2 < \cdots < t_{n-1} < t_n < t_{n+1}$, the conditional probability densities obey the *Markov property*

$$p_{Y,1|n}(t_{n+1}, y_{n+1} \,|\, t_1, y_1; t_2, y_2; \ldots; t_{n-1}, y_{n-1}; t_n, y_n) = p_{Y,1|1}(t_{n+1}, y_{n+1} \,|\, t_n, y_n). \tag{C.22}$$

Viewing t_n as being “now”, this property means that the (conditional) probability that the process takes a given value y_{n+1} in the future (at t_{n+1}) only depends on its present value y_n, not on the values it took in the past.

An even more drastically “memoryless” class of processes is that of *fully random processes*, for which the value taken at a given time is totally independent of the past values. For such a process, the conditional probability densities equal the joint probability densities—i.e. $p_{Y,n|m} = p_{Y,n}$ for all m, n—, and repeated applications of Bayes’ theorem (C.13) show that the n-point density factorizes into the product of n single-time densities,

$$p_{Y,n}(t_1, y_1; \ldots; t_n, y_n) = p_{Y,1}(t_1, y_1) \cdots p_{Y,1}(t_n, y_n).$$

One can check that a Markov process is entirely determined by the single-time probability density $p_{Y,1}(t_1, y_1)$ and by the *transition probability* $p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)$, or equivalently by $p_{Y,1}(t_1, y_1)$ and the two-time density $p_{Y,2}(t_1, y_1; t_2, y_2)$.
For instance, the 3-time probability density can be rewritten as

$$\begin{aligned} p_{Y,3}(t_1, y_1; t_2, y_2; t_3, y_3) &= p_{Y,1|2}(t_3, y_3 \,|\, t_1, y_1; t_2, y_2)\, p_{Y,2}(t_1, y_1; t_2, y_2) \\ &= p_{Y,1|1}(t_3, y_3 \,|\, t_2, y_2)\, p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)\, p_{Y,1}(t_1, y_1), \end{aligned} \tag{C.23}$$

where we have used twice Bayes’ theorem (C.13) and once the Markov property (C.22).

Remarks:

∗ The Markov property (C.22) characterizes the n-point densities for ordered times. The value for arbitrary t_1, t_2, …, t_n follows from the necessary invariance [property (C.11b)] of $p_{Y,n}$ when two pairs (t_j, y_j) and (t_k, y_k) are exchanged.

∗ The single-time probability density $p_{Y,1}$ and the transition probability $p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)$ are not fully independent of each other, since they have to obey the obvious identity

$$p_{Y,1}(t_2, y_2) = \int p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)\, p_{Y,1}(t_1, y_1)\, \mathrm{d}y_1. \tag{C.24}$$

C.2.5 b Chapman–Kolmogorov equation

Integrating Eq. (C.23) over the intermediate value y_2 of the stochastic process, while taking into account the consistency condition (C.11c), gives

$$p_{Y,2}(t_1, y_1; t_3, y_3) = p_{Y,1}(t_1, y_1) \int p_{Y,1|1}(t_3, y_3 \,|\, t_2, y_2)\, p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)\, \mathrm{d}y_2,$$

where $t_1 < t_2 < t_3$.

---

(101) … or *Markoff* in the older literature.

Dividing by $p_{Y,1}(t_1, y_1)$, one obtains the *Chapman*(cm)*–Kolmogorov*(cn) *equation*

$$\boxed{p_{Y,1|1}(t_3, y_3 \,|\, t_1, y_1) = \int p_{Y,1|1}(t_3, y_3 \,|\, t_2, y_2)\, p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)\, \mathrm{d}y_2 \quad \text{for } t_1 < t_2 < t_3,} \tag{C.25}$$

which gives a relation—a nonlinear integral-functional equation—fulfilled by the transition probability of a Markov process.

Reciprocally, two arbitrary nonnegative functions $p_{Y,1}(t_1, y_1)$, $p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1)$ obeying the two identities (C.24) and (C.25) entirely define a unique Markov process.

Remark: The Chapman–Kolmogorov equation follows quite obviously when invoking the Markov property in the more generic relation (C.15), which holds for every stochastic process. In contrast to the latter, Eq. (C.25) is closed, i.e. does not depend on another function.

C.2.5 c Examples of Markov processes

Wiener process

The stochastic process defined by the "initial condition" $p_{Y,1}(t=0, y) = \delta(y)$ for $y \in \mathbb{R}$ and the Gaussian transition probability $(0 < t_1 < t_2)$

$$p_{Y,1|1}(t_2, y_2 \,|\, t_1, y_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp\left[-\frac{(y_2 - y_1)^2}{2(t_2 - t_1)}\right] \tag{C.26a}$$

is called *Wiener*(co) *process*.

One easily checks that the transition probability (C.26a) satisfies the Chapman–Kolmogorov equation (C.25), and that the probability density at time $t > 0$ is given by

$$p_{Y,1}(t, y) = \frac{1}{\sqrt{2\pi t}}\, \mathrm{e}^{-y^2/2t}. \tag{C.26b}$$

The Wiener process is obviously not a stationary process, since for instance the second moment $\langle [Y(t)]^2 \rangle = t$ depends on time.

Remark: The single-time probability density (C.26b) is solution of the diffusion equation

$$\frac{\partial f}{\partial t} = \frac{1}{2} \frac{\partial^2 f}{\partial y^2} \tag{C.27}$$

with diffusion coefficient $D = \frac{1}{2}$.

Poisson process

Consider now the integer-valued stochastic process $Y(t)$ defined by the Poisson-distributed [cf. Eq. (B.10)] transition probability $(0 \leq t_1 \leq t_2)$

$$p_{Y,1|1}(t_2, n_2 \,|\, t_1, n_1) = \frac{(t_2 - t_1)^{n_2 - n_1}}{(n_2 - n_1)!}\, \mathrm{e}^{-(t_2 - t_1)} \quad \text{for } n_2 \geq n_1 \tag{C.28}$$

and 0 otherwise, and by the single-time probability density $p_{Y,1}(t=0, n) = \delta_{n,0}$. That is, a realization $y(t)$ is a succession of unit steps taking place at arbitrary instants, whose number between two given times t_1, t_2 obeys a Poisson distribution with parameter $t_2 - t_1$.

$Y(t)$ is a non-stationary Markov process, called *Poisson process*.

Remark: In both Wiener and Poisson processes, the probability density of the *increment* ($y_2 - y_1$ resp. $n_2 - n_1$) between two successive instants t_1, t_2 only depends on the time difference $t_2 - t_1$, not on t_1 (or t_2) alone. Such increments are called *stationary*. Since in addition successive increments are independent, both processes are instances of *Lévy*(cp) *processes*.

(cm) S. Chapman, 1888–1970 (cn) A. N. Kolmogorov, 1903–1987 (co) N. Wiener, 1894–1964 (cp) P. Lévy, 1886–1971

C.2.5 d Stationary Markov processes

An interesting case in physics is that of stationary Markov processes. For such processes, the transition probability $p_{Y,1|1}(t_2,y_2\,|\,t_1,y_1)$ only depends on the time difference $\tau\equiv t_2-t_1$, which is hereafter reflected in the use of the special notation

$$\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)\equiv p_{Y,1|1}(t_1+\tau,y_2\,|\,t_1,y_1). \tag{C.29}$$

Using this notation, the Chapman–Kolmogorov equation (C.25) takes the form (both τ and τ' are taken to be nonnegative)

$$\boxed{\mathcal{T}_{Y;\tau+\tau'}(y_3\,|\,y_1)=\int\mathcal{T}_{Y;\tau'}(y_3\,|\,y_2)\,\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)\,\mathrm{d}y_2.} \tag{C.30}$$

If a Markov process is also stationary, the single-time probability density $p_{Y,1}(y)$ does not depend on time. Invoking a setup in which the probability density would first be time-dependent, i.e. in which the stochastic process Y is not (yet) stationary, $p_{Y,1}$ characterizes the large-time "equilibrium" distribution, reached after a sufficiently large τ, irrespective of the "initial" distribution $y(t)$ at some time $t=t_0$. Taking as initial condition $p_{Y,1}(t{=}t_0,y)=\delta(y-y_0)$, where y_0 is arbitrary, one finds

$$p_{Y,1}(y)=\lim_{\tau\to+\infty}\mathcal{T}_{Y;\tau}(y\,|\,y_0).$$

This follows from the identities

$$p_{Y,1}(t_0+\tau,y)=\int p_{Y,2}(t_0+\tau,y;t_0,y')\,\mathrm{d}y'=\int p_{Y,1|1}(t_0+\tau,y\,|\,t_0,y')\,p_{Y,1}(t_0,y')\,\mathrm{d}y'$$
$$=\int\mathcal{T}_{Y;\tau}(y\,|\,y')\,p_{Y,1}(t_0,y')\,\mathrm{d}y',$$

which with the assumed initial distribution $p_{Y,1}(t_0,y')$ gives the result. □

Ornstein–Uhlenbeck process

An example of stationary Markov process is the *Ornstein*(cq)*–Uhlenbeck*(cr) *process* [45] defined by the (time-independent) single-time probability density

$$p_{Y,1}(y)=\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-y^2/2} \tag{C.31a}$$

and the transition probability

$$\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)=\frac{1}{\sqrt{2\pi(1-\mathrm{e}^{-2\tau})}}\,\exp\left[-\frac{(y_2-y_1\mathrm{e}^{-\tau})^2}{2(1-\mathrm{e}^{-2\tau})}\right]. \tag{C.31b}$$

One can show that the Ornstein–Uhlenbeck process is also Gaussian and that its autocorrelation function is $\kappa(\tau)=\mathrm{e}^{-\tau}$.

Doob's(cs) theorem actually states that the Ornstein–Uhlenbeck process is, up to scalings or translations of the time argument, the only nontrivial(102) process which is Markovian, Gaussian and stationary.

C.2.5 e Master equation for a Markov process

For an *homogeneous* Markov process $Y(t)$, i.e. a process for which the probability transition $p_{Y,1|1}(t_2,y_2\,|\,t_1,y_1)$ only depends on the time difference $\tau\equiv t_2-t_1$, one can derive under minimal assumptions a linear integrodifferential equation for the transition probability, which constitutes the differential form of the Chapman–Kolmogorov equation for the process.

Remark: The assumption on the probability transition does not automatically imply that the process is stationary; yet in analogy with Eq. (C.29) we shall denote it by $\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)$.

---

(102) The fully random process mentioned below the introduction of the Markov property (C.22) may also be Gaussian and stationary.

(cq) L. Ornstein, 1880–1941 (cr) G. Uhlenbeck, 1900–1988 (cs) J. L. Doob, 1910–2004

Master equation

Let us assume that for time differences τ much smaller than some time scale τ_c, the transition probability is of the form

$$\mathcal{T}_{Y;\tau}(y_2\,|\,y_1) = [1-\gamma(y_1)\tau]\,\delta(y_2-y_1) + \Gamma(y_2\,|\,y_1)\tau + o(\tau), \tag{C.32a}$$

where $o(\tau)$ denotes a term which is much smaller than τ in the limit $\tau \to 0$. The nonnegative quantity $\Gamma(y_2\,|\,y_1)$ is readily interpreted as being the transition rate from y_1 to y_2, and $\gamma(y_1)$ is its integral over y_2

$$\gamma(y_1) = \int \Gamma(y_2\,|\,y_1)\,\mathrm{d}y_2, \tag{C.32b}$$

thereby ensuring that the integral of the transition probability $\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)$ over all possible final states y_2 gives unity.

Consider now the Chapman–Kolmogorov equation (C.30). Rewriting $\mathcal{T}_{Y;\tau'}(y_3\,|\,y_2)$ with the help of Eq. (C.32a), i.e. under the assumption that $\tau' \ll \tau_c$, and leaving aside the negligible term $o(\tau')$, one finds

$$\mathcal{T}_{Y;\tau+\tau'}(y_3\,|\,y_1) = \big[1-\gamma(y_3)\,\tau'\big]\,\mathcal{T}_{Y;\tau}(y_3\,|\,y_1) + \tau'\!\int \Gamma(y_3\,|\,y_2)\,\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)\,\mathrm{d}y_2.$$

Note that we need not assume anything on τ here. Taylor-expanding $\mathcal{T}_{Y;\tau+\tau'}$ with respect to τ' and dividing both sides by τ', this gives in the limit $\tau' \to 0$ the integrodifferential equation

$$\frac{\partial\mathcal{T}_{Y;\tau}(y_3\,|\,y_1)}{\partial\tau} = -\gamma(y_3)\mathcal{T}_{Y;\tau}(y_3\,|\,y_1) + \int \Gamma(y_3\,|\,y_2)\,\mathcal{T}_{Y;\tau}(y_2\,|\,y_1)\,\mathrm{d}y_2,$$

where the derivative on the left-hand side has to be taken with a grain of salt in case τ' may (for physical reasons pertaining to the system being considered) actually not become vanishingly small.

Using Eq. (C.32b) for $\gamma(y_3)$ and relabeling the variables ($y_1 \to y_0$, $y_2 \to y'$, $y_3 \to y$), this can be rewritten as

$$\boxed{\frac{\partial\mathcal{T}_{Y;\tau}(y\,|\,y_0)}{\partial\tau} = \int \big[\Gamma(y\,|\,y')\,\mathcal{T}_{Y;\tau}(y'\,|\,y_0) - \Gamma(y'\,|\,y)\,\mathcal{T}_{Y;\tau}(y\,|\,y_0)\big]\,\mathrm{d}y'.} \tag{C.33}$$

This evolution equation—which is fully equivalent to the Chapman–Kolmogorov equation—for the transition probability $\mathcal{T}_{Y;\tau}$ is called *master equation*. It has the structure of a balance equation, with a gain term, involving the rate $\Gamma(y\,|\,y')$, and a loss term depending on the rate $\Gamma(y'\,|\,y)$. It is a linear integrodifferential equation, of first order with respect to τ.

Evolution equation for the single-time probability density

From the master equation (C.33) for the transition probability, one can deduce an equation governing the dynamics of the single-time probability density, which turns out to possess exactly the same structure.

Rewriting Eq. (C.24) as

$$p_{Y,1}(\tau,y) = \int \mathcal{T}_{Y;\tau}(y\,|\,y_0)\,p_{Y,1}(t\!=\!0,y_0)\,\mathrm{d}y_0, \tag{C.34}$$

and differentiating with respect to τ, one obtains with the help of the master equation

$$\frac{\partial p_{Y,1}(\tau,y)}{\partial\tau} = \int \big[\Gamma(y\,|\,y')\,\mathcal{T}_{Y;\tau}(y'\,|\,y_0) - \Gamma(y'\,|\,y)\,\mathcal{T}_{Y;\tau}(y\,|\,y_0)\big]\,p_{Y,1}(t\!=\!0,y_0)\,\mathrm{d}y_0\,\mathrm{d}y'.$$

Performing the integration over y_0 and using relation (C.34), this yields

$$\boxed{\frac{\partial p_{Y,1}(\tau,y)}{\partial\tau} = \int \big[\Gamma(y\,|\,y')\,p_{Y,1}(\tau,y') - \Gamma(y'\,|\,y)\,p_{Y,1}(\tau,y)\big]\,\mathrm{d}y',} \tag{C.35}$$

formally identical to the equation for $\mathcal{T}_{Y;\tau}$, and accordingly also referred to as *master equation*.

Remark: When applied to a physical system, Eq. (C.35) allows the computation of the single-time probability density at any time from an initial distribution $p_{Y,1}(t\!=\!0,y)$ and the transition rates Γ.

C.3 Spectral decomposition of stationary processes

In this section, we focus on stationary stochastic processes and introduce an alternative description of their statistical properties, based on Fourier transformation (Sec. C.3.1). This approach in particularly leads to the Wiener–Khinchin theorem relating the spectral density to the autocorrelation function (Sec. C.3.2).

C.3.1 Fourier transformations of a stationary process

Consider a stationary process $Y(t)$. In general, a given realization $y(t)$ will not be an integrable function, e.g. because it does not tend to 0 as t goes to infinity. In order to talk of Fourier transformations of the realization, one thus has to first introduce a finite time interval $[0,\mathcal{T}]$ with $\mathcal{T}$ positive, and to work with continuations of the restriction of $y(t)$ to this interval, before letting $\mathcal{T}$ go to infinity.

C.3.1 a Fourier transform

Let first $y_{\mathcal{T}}(t)$ denote the function which coincides with $y(t)$ for $0<t<\mathcal{T}$, and which vanishes outside the interval. $y_{\mathcal{T}}(t)$ may be seen as the realization of a stochastic process $Y_{\mathcal{T}}(t)$.

One can meaningfully define the Fourier transform of $y_{\mathcal{T}}(t)$ with the usual formula

$$\tilde{y}_{\mathcal{T}}(\omega) \equiv \int y_{\mathcal{T}}(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t = \int_0^{\mathcal{T}} y(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t. \tag{C.36a}$$

$\tilde{y}_{\mathcal{T}}(\omega)$ is now the realization of a stochastic function $\tilde{Y}_{\mathcal{T}}(\omega)$. The inverse transform reads

$$y_{\mathcal{T}}(t) = \int \tilde{y}_{\mathcal{T}}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \frac{\mathrm{d}\omega}{2\pi}. \tag{C.36b}$$

Taking the limit $\mathcal{T}\to\infty$ defines for each realization $y(t)$ a corresponding $\tilde{y}(\omega)$. The latter is itself the realization of a process $\tilde{Y}(\omega)$, and one symbolically writes for the stochastic processes themselves

$$\tilde{Y}(\omega) = \int Y(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t, \qquad Y(t) = \int \tilde{Y}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \frac{\mathrm{d}\omega}{2\pi}. \tag{C.37}$$

Remark: The reader can check that thanks to the assumed stationarity of the process, we could have started with restrictions of the realizations to any interval of width $\mathcal{T}$, for instance $[-\frac{\mathcal{T}}{2},\frac{\mathcal{T}}{2}]$, without changing the result after taking the limit $\mathcal{T}\to\infty$.

C.3.1 b Fourier series

Alternatively, one can consider the $\mathcal{T}$-periodic function which coincides with $y(t)$ on the interval $[0,\mathcal{T}]$. This $\mathcal{T}$-periodic function can be written as a Fourier series, which of course equals $y(t)$ for $0<t<\mathcal{T}$:

$$y(t) = \sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{-\mathrm{i}\omega_n t} \quad \text{for } 0<t<\mathcal{T}, \tag{C.38a}$$

where the (angular) frequencies and Fourier coefficients are as usual given by

$$\omega_n = \frac{2\pi n}{\mathcal{T}}, \qquad c_n = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} y(t)\, \mathrm{e}^{\mathrm{i}\omega_n t}\, \mathrm{d}t \quad \text{for } n\in\mathbb{Z}. \tag{C.38b}$$

Again, one considers the limit $\mathcal{T}\to\infty$ at the end of the calculations.

For the stochastic process, one similarly writes

$$Y(t) = \sum_{n=-\infty}^{\infty} C_n\, \mathrm{e}^{-\mathrm{i}\omega_n t} \quad \text{for } 0<t<\mathcal{T}, \tag{C.39a}$$

where C_n is a random variable, of which the Fourier coefficient c_n is a realization

$$C_n = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} Y(t)\, \mathrm{e}^{\mathrm{i}\omega_n t}\, \mathrm{d}t. \tag{C.39b}$$

At fixed $\mathcal{T}$, one has the obvious relationship

$$c_n = \frac{1}{\mathcal{T}} \tilde{y}_{\mathcal{T}}(\omega_n), \tag{C.40a}$$

which for the corresponding stochastic variables reads

$$C_n = \frac{1}{\mathcal{T}} \tilde{Y}_{\mathcal{T}}(\omega_n). \tag{C.40b}$$

An equivalent relation will also holds in the limit $\mathcal{T} \to \infty$.

Remark: Instead of the complex Fourier transform, one can also use real transforms, for instance the sine transform as in Ref. [49].

C.3.1 c Consequences of stationarity

The assumed stationarity of the stochastic process, which allowed us to define the Fourier transformations on an arbitrary interval of width $\mathcal{T}$, has further consequences, some of which we now investigate.

First, the single-time average $\langle Y(t)\rangle$ is independent of time, $\langle Y(t)\rangle = \langle Y\rangle$. Averaging the Fourier coefficient (C.39b) over an ensemble of realizations, the sample average and integration over time can be exchanged, which at once leads to

$$\langle C_0\rangle = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \langle Y\rangle\, \mathrm{d}t = \langle Y\rangle\,, \qquad \langle C_n\rangle = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \langle Y\rangle\, \mathrm{e}^{\mathrm{i}\omega_n t}\, \mathrm{d}t = 0 \text{ for } n \neq 0. \tag{C.41}$$

Consider now a two-time average, which, since the process is stationary, only depends on the time difference. For the sake of simplicity, we assume that the stochastic process is real-valued and centered, $\langle Y\rangle = 0$, so as to shorten the expression of the autocorrelation function. The latter reads

$$\kappa(\tau) = \big\langle Y(t)\, Y(t+\tau)\big\rangle = \sum_{n,n'=-\infty}^{\infty} \langle C_n C_{n'}\rangle\, \mathrm{e}^{-\mathrm{i}(\omega_n+\omega_{n'})t} \mathrm{e}^{-\mathrm{i}\omega_n\tau},$$

which can only be independent of t for all values of τ if $\langle C_n C_{n'}\rangle = 0$ for all values of n and n' such that $\omega_n + \omega_{n'} \neq 0$, i.e. [cf. the frequency (C.38b)] when $n' \neq -n$. Using the classical property $C_{-n} = C_n^*$ of Fourier coefficients, this condition can be written as

$$\langle C_n C_{n'}^*\rangle = \big\langle |C_n|^2\big\rangle\, \delta_{n,n'}, \tag{C.42}$$

with $\delta_{n,n'}$ the Kronecker symbol. Fourier coefficients of different frequencies are thus uncorrelated, and the autocorrelation function reads

$$\kappa(\tau) = \sum_{n=-\infty}^{\infty} \big\langle |C_n|^2\big\rangle\, \mathrm{e}^{-\mathrm{i}\omega_n\tau}. \tag{C.43}$$

C.3.2 Wiener–Khinchin theorem

C.3.2 a Spectral density of a stationary process

Consider a centered stationary stochastic process $Y(t)$. Working first on a finite interval $[0, \mathcal{T}]$, one can introduce the Fourier coefficients C_n or alternatively the Fourier transform $\tilde{Y}_{\mathcal{T}}(\omega)$. Using relation (C.40b), Eq. (C.42) reads

$$\frac{1}{\mathcal{T}^2} \Big\langle \tilde{Y}_{\mathcal{T}}(\omega_n) \tilde{Y}_{\mathcal{T}}(\omega_{n'})^* \Big\rangle = \frac{1}{\mathcal{T}^2} \Big\langle \big|\tilde{Y}_{\mathcal{T}}(\omega_n)\big|^2 \Big\rangle\, \delta_{n,n'}.$$

The Kronecker symbol can be rewritten under consideration of the expression (C.38b) of the Fourier frequencies as

$$\delta_{n,n'} = \delta(n-n') = \frac{2\pi}{\mathcal{T}}\delta(\omega_n - \omega_{n'}),$$

leading to

$$\left\langle \tilde{Y}_{\mathcal{T}}(\omega_n)\tilde{Y}_{\mathcal{T}}(\omega_{n'})^* \right\rangle = \frac{2\pi}{\mathcal{T}}\left\langle \big|\tilde{Y}_{\mathcal{T}}(\omega_n)\big|^2 \right\rangle \delta(\omega_n - \omega_{n'}).$$

Defining now the *spectral density* $S(\omega)$ as

$$\boxed{S(\omega) = \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\left\langle \big|\tilde{Y}_{\mathcal{T}}(\omega)\big|^2 \right\rangle,} \tag{C.44}$$

the above identity becomes in the limit $\mathcal{T}\to\infty$

$$\boxed{\left\langle \tilde{Y}(\omega)\,\tilde{Y}(\omega')^* \right\rangle = 2\pi\delta(\omega-\omega')\,S(\omega),} \tag{C.45}$$

where the discrete frequencies ω_n, $\omega_{n'}$ have been replaced by values ω, ω' from a continuous interval.

Coming back to Fourier representations defined on a finite-size time interval, let $\mathcal{I}_\omega \equiv [\omega, \omega+\Delta\omega]$ denote an interval in frequency space, over which $\tilde{Y}_{\mathcal{T}}(\omega)$ is assumed to be continuous and to vary only moderately. One introduces a function $\sigma(\omega)$ such that

$$\sigma(\omega)\,\Delta\omega \equiv \sum_{\omega_n\in\mathcal{I}_\omega}\langle |C_n|^2\rangle = \sum_{\omega_n\in\mathcal{I}_\omega}\frac{1}{\mathcal{T}^2}\big\langle \big|\tilde{Y}_{\mathcal{T}}(\omega_n)\big|^2 \big\rangle.$$

From Eq. (C.38b), the number of modes ω_n inside the interval $\mathcal{I}_\omega$ is $\Delta\omega/(2\pi/\mathcal{T}) = \mathcal{T}\,\Delta\omega/2\pi$. Taking the limit $\mathcal{T}\to\infty$, this gives

$$\sigma(\omega) = \lim_{\mathcal{T}\to\infty}\frac{1}{2\pi\mathcal{T}}\left\langle \big|\tilde{Y}(\omega)\big|^2 \right\rangle = \frac{1}{2\pi}S(\omega)$$

which shows the relation between the spectral density and the sum of the (squared) amplitudes of the Fourier modes.

C.3.2 b Wiener–Khinchin theorem

Consider the autocorrelation function (C.43). In the limit $\mathcal{T}\to\infty$, the discrete sum is replaced by an integral, yielding

$$\kappa(\tau) = \lim_{\mathcal{T}\to\infty}\frac{1}{2\pi\mathcal{T}}\int \left\langle \big|\tilde{Y}_{\mathcal{T}}(\omega)\big|^2 \right\rangle \mathrm{e}^{-\mathrm{i}\omega\tau}\,\mathrm{d}\omega.$$

With the help of the spectral density (C.44), this also reads

$$\boxed{\kappa(\tau) = \int S(\omega)\,\mathrm{e}^{-\mathrm{i}\omega\tau}\,\frac{\mathrm{d}\omega}{2\pi}.} \tag{C.46a}$$

That is, the autocorrelation function is the (inverse) Fourier transform of the spectral density, and reciprocally

$$\boxed{S(\omega) = \int \kappa(\tau)\,\mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau.} \tag{C.46b}$$

The relations (C.46) are known as *Wiener–Khinchin*(ct) *theorem*, and show that the autocorrelation function $\kappa(\tau)$ and the spectral density $S(\omega)$ contain exactly the same amount of information on the stochastic process.

Remarks:

∗ In deriving the theorem, we did not use the stationarity of the stochastic process, but only its *wide-sense stationarity* (or *covariance stationarity*), which only requires that the first and second moments be independent of time, not all of them.

(ct) A. Ya. Khinchin, 1894–1959

$\ast$ If the stochastic process $Y(t)$ is not centered, then the Wiener–Khinchin-theorem states that its autocorrelation function $\kappa(\tau)$ and the spectral density $S(\omega)$ of the fluctuations around its average value constitute a Fourier-transform pair.
The spectral density of $X(t)$ itself is given by $S(\omega) + 2\pi|\langle Y\rangle|^2\delta(\omega)$, i.e. it includes a singular contribution at $\omega = 0$.

Bibliography for Appendix C

- Pottier, *Nonequilibrium statistical physics* [6], chapter 1 § 6–10.
- van Kampen, *Stochastic processes in physics and chemistry* [49], chapters III & IV.

Bibliography

[1] S. R. de Groot, P. Mazur, *Non-equilibrium thermodynamics* (North-Holland, Amsterdam, 1962).

[2] R. Kubo, M. Toda, N. Hashitsume, *Statistical physics II*: *Nonequilibrium statistical mechanics* (Springer, Berlin, Heidelberg, 1985).

[3] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 5*: *Statistical physics part 1*, 3rd ed. (Butterworth-Heinemann, Oxford, 1980).

[4] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 9*: *Statistical physics part 2*, 2nd ed. (Butterworth-Heinemann, Oxford, 1980).

[5] L. Landau, E. Lifschitz, L.P.Pitajewski, *Course of Theoretical Physics. Vol. 10*: *Physical kinetics* (Butterworth-Heinemann, Oxford, 1981).

[6] N. Pottier, *Nonequilibrium Statistical Physics* (University Press, Oxford, 2010).

[7] R. Zwanzig, *Nonequilibrium statistical mechanics* (University Press, Oxford, 2001).

[8] H. B. Callen, *Thermodynamics and an introduction to Thermostatistics*, 2nd ed. (John Wiley & sons, New York, 1985).

[9] M. Le Bellac, F. Mortessagne, G. G. Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* (University Press, Cambridge, 2004).

[10] L. Onsager, *Reciprocal relations in irreversible processes. I*, Phys. Rev. **37** (1931) 405–426.

[11] L. Onsager, *Reciprocal relations in irreversible processes. II*, Phys. Rev. **38** (1931) 2265–2279.

[12] D. G. Miller, *Thermodynamics of irreversible processes*, Chem. Rev. **60** (1960) 15–37.

[13] H. B. G. Casimir, *On Onsager's principle of microscopic reversibility*, Rev. Mod. Phys. **17** (1945) 343–350.

[14] A. Fick, *Über Diffusion*, Ann. Phys. **170** (1855) 59–86.

[15] A. Einstein, *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Ann. Phys. **17** (1905) 549–560.

[16] K. Uchida *et al.*, *Observation of the spin Seebeck effect*, Nature **455** (2008) 778–781.

[17] G. E. W. Bauer, E. Saitoh, B. J. van Wees, *Spin caloritronics*, Nature Mat. **11** (2012) 391–399.

[18] H. Adachi, K. Uchida, E. Saitoh, S. Maekawa, *Theory of the spin Seebeck effect*, Rep. Prog. Phys. **76** (2013) 036501 [arXiv:1209.6407].

[19] S. D. Brechet, J.-P. Ansermet, *Thermodynamics of a continuous medium with electric and magnetic dipoles*, Eur. Phys. J. B **86** (2013) 1–19 [arXiv:1303.6809].

[20] S. D. Brechet, F. A. Vetro, E. Papa, S. E. Barnes, J.-P. Ansermet, *Evidence for a magnetic Seebeck effect*, Phys. Rev. Lett. **111** (2013) 087205 [arXiv:1306.1001].

[21] R. E. Graves, B. M. Argrow, *Bulk viscosity: Past to present*, J. Thermophys. Heat Transfer **13** (1999) 337–342.

[22] S. J. Blundell, K. M. Blundell, *Concepts in thermal physics* (University Press, Oxford, 2006).

[23] Y. Andoh *et al.*, *All-atom molecular dynamics calculation study of entire poliovirus empty capsids in solution*, J. Chem. Phys. **141** (2014) 165101.

[24] H. Goldstein, *Classical mechanics*, 3rd ed. (Addison-Wesley, Reading, MA, 2001).

[25] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed. (Springer, New York, 1989).

[26] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 1: Mechanics*, 3rd ed. (Butterworth-Heinemann, Oxford, 1976).

[27] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Quantum mechanics, vol. 1* (Wiley-VCH, Weinheim, 1977).

[28] L. Landau, E. Lifshitz, *Course of Theoretical Physics. Vol. 3: Quantum mechanics*, 3rd ed. (Butterworth-Heinemann, Oxford, 1981).

[29] A. Messiah, *Quantum Mechanics, vol. II* (North Holland, Amsterdam, 1961).

[30] D. Griffiths, *Introduction to quantum mechanics* (Prentice Hall, Upper Saddle River, NJ, 1995).

[31] C. E. Shannon, *A mathematical theory of information*, Bell System Tech. J. **27** (1948) 379—423 & 623–656.

[32] M. Toda, R. Kubo, N. Saitô, *Statistical physics I: Equilibrium statistical mechanics* (Springer, Berlin, Heidelberg, 1983).

[33] K. Huang, *Statistical Mechanics*, 2nd ed. (John Wiley & Sons, New York, 1987).

[34] J. G. Kirkwood, *The statistical mechanical theory of transport processes II. Transport in gases*, J. Chem. Phys. **15** (1946) 72–76 [Erratum: Ibid. **15** (1946) 155.].

[35] F. Reif, *Fundamentals of statistical and thermal physics* (McGraw-Hill, New York, 1965).

[36] L. D. Landau, *Theory of Fermi-liquids*, Sov. Phys. JETP **3** (1957) 920 [Zh. Eksp. Teor. Fiz. **30** (1956) 1050].

[37] D. Guéry-Odelin, F. Zambelli, J. Dalibard, S. Stringari, *Collective oscillations of a classical gas confined in harmonic traps*, Phys. Rev. A **60** (1999) 4851–4856.

[38] D. S. Lobser, A. E. S. Barentine, E. A. Cornell, H. J. Lewandowski, *Observation of a persistent non-equilibrium state in cold atoms*, Nature Phys. **11** (2015) 1009–1012.

[39] H. Grad, *On the kinetic theory of rarefied gases*, Commun. Pure Appl. Math. **2** (1949) 331–407.

[40] A. Sommerfeld, *Vorlesungen über theoretische Physik. Band V: Thermodynamik und Statistik*, 2. ed. (Harri Deutsch, Frankfurt am Main, 1962/1977).

[41] O. Sackur, *Die Anwendung der kinetischen Theorie der Gase auf chemische Probleme*, Ann. Phys. **36** (1911) 958–980.

[42] H. Tetrode, *Die chemische Konstante der Gase und das elementare Wirkungsquantum*, Ann. Phys. **38** (1912) 434–442.

[43] P. L. Bhatnagar, E. P. Gross, M. Krook, *A model for collision processes in gases. I. Small amplitude orocesses in charged and neutral one-component systems*, Phys. Rev. **94** (1954) 511–525.

[44] P. Langevin, *Sur la théorie du mouvement brownien*, C. R. Acad. Sci. (Paris) **146** (1908) 530–533 [English translation: Am. J. Phys. **65** (1997) 1079–1081].

[45] G. E. Uhlenbeck, L. S. Ornstein, *On the theory of the Brownian motion*, Phys. Rev. **36** (1930) 823–841.

[46] A. Caldeira, A. Leggett, *Quantum tunnelling in a dissipative system*, Ann. Phys. **149** (1983) 374–456.

[47] V. Hakim, V. Ambegaokar, *Quantum theory of a free particle interacting with a linearly dissipative environment*, Phys. Rev. A **32** (1985) 423–434.

[48] C. Aslangul, N. Pottier, D. Saint-James, *Quantum dynamics of a damped free particle*, J. Phys. **44** (1987) 1871–1880.

[49] N. G. van Kampen, *Stochastic processes in physics in chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).

[50] R. Kubo, *Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems*, J. Phys. Soc. Japan **12** (1957) 570–586.

[51] J. D. Jackson, *Classical electrodynamics*, 3. ed. (John Wiley & Sons, New York, 1999).

[52] H. M. Nussenzveig, *Causality and dispersion relations* (Academic Press, New York & London, 1972).

[53] R. Kubo, *The fluctuation-dissipation theorem*, Rep. Prog. Phys. **29** (1966) 255–284.

[54] J. B. Johnson, *Thermal Agitation of Electricity in Conductors*, Phys. Rev. **32** (1928) 97–109.

[55] H. Nyquist, *Thermal Agitation of Electric Charge in Conductors*, Phys. Rev. **32** (1928) 110–113.

[56] H. B. Callen, T. A. Welton, *Irreversibility and general noise*, Phys. Rev. **83** (1951) 34–40.

[57] M. S. Green, *Markoff random processes and the statistical mechanics of time-dependent phenomena*, J. Chem Phys. **20** (1952) 1281–1295.

[58] M. S. Green, *Markoff random processes and the statistical mechanics of time-dependent phenomena. II. Irreversible processes in fluids*, J. Chem Phys. **22** (1954) 398–4135.

[59] C. Cohen-Tannoudji, *Mouvement Brownien, réponses linéaires, équations de Mori et fonctions de corrélation*, available online at `http://www.phys.ens.fr/~cct/college-de-france/1977-78/1977-78.htm` [checked on Nov.4, 2014].

[60] B. J. Berne, C. D. Harp, in *Advances in Chemical Physics, vol. XVII*, I. Prigogine, S. A. Rice (ed.) (John Wiley & Sons, New York, 1970), S. 64–227.

[61] B. J. Berne, in *Physical Chemistry. An advanced treatise, vol. VIIIB*, D. Henderdon (ed.) (Academic Press, New York & London, 1971), S. 539–716.

[62] R. Zwanzig, *Time-Correlation Functions and Transport Coefficients in Statistical Mechanics*, Ann. Rev. Phys. Chem. **16** (1965) 67–102.

[63] L. P. Kadanoff, P. C. Martin, *Hydrodynamic equations and correlation functions*, Ann. Phys. (NY) **24** (1963) 419–469.

[64] L. Van Hove, *Correlations in space and time and Born approximation scattering in systems of interacting particles*, Phys. Rev. **95** (1954) 249–262.