Chapter I

Symmetry transformations

A. Basic symmetries

A-1. Definition

Consider a physical system which, at time t_0, is in the state $S(t_0)$. For a classical system containing N particles for example, $S(t_0)$ could be defined by the $2N$ values of the vectors giving their positions and velocities at time t_0. Once it has evolved, the system is at time t in the state $S(t)$.

Introduction to Continuous Symmetries: From Space–Time to Quantum Mechanics, First Edition. Franck Laloë.

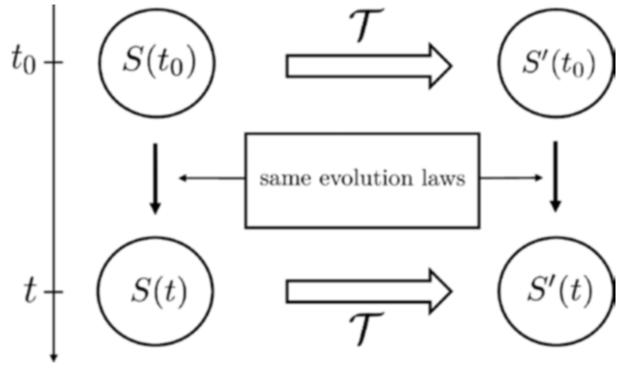

Figure 1: Consider a transformation $\mathcal{T}$ that changes an arbitrary physical system state $S(t)$ into another state $S'(t)$. If the sequence of the states $S(t)$ and that of the states $S'(t)$ follow the same evolution laws, $\mathcal{T}$ is said to be a symmetry transformation.

Imagine now a transformation $\mathcal{T}$ that changes a system in a given state S into another system in a state S' (figure 1). Obviously, the kind of transformations $\mathcal{T}$ one can think of are numerous: translation of the positions, rotation through a constant or time-dependent angle, multiplication by a factor 2 of the inter-particle distances, changing the sign of the electric charges, etc.

Applying this transformation $\mathcal{T}$ to a sequence of states $S(t)$ that describe a possible motion of the system yields another sequence of states $S'(t)$. By definition, $\mathcal{T}$ is said to be a symmetry transformation if that latter sequence of states $S'(t)$ also describes a possible motion of the system, following the same evolution laws. This condition must be satisfied for any initial state $S(t_0)$. Therefore:

A transformation $\mathcal{T}$ is said to be a symmetry transformation if it transforms any possible motion into another possible motion.

This means that in figure 1 one can, at any instant t and for any motion, "finish closing the square" with a transformation $\mathcal{T}$.

The definition of a symmetry transformation $\mathcal{T}$ not only concerns a particular state of the system $S(t)$ at a given instant (as, for example, when a given geometrical figure is said to be symmetrical), but all its history as well, i.e. the whole set of states the system successively goes through as time passes.

Comment:

One can also consider $\mathcal{T}$ transformations that are not instantaneous, such as translations or expansions of the time scale, or Lorentz transformations of a field that is extended in space (its different points then undergo different time transformations), etc. In the left and right parts of figure 1, different time scales should then be used. In fact, for an extended system, a Lorentz transformation is not a succession of transformations, each acting at a single time t, but a global transformation acting on the entire time evolution of the system.

A-2. Examples

A transformation $\mathcal{T}$ is not necesssarily a symmetry transformation. For instance, space dilation by a factor 2 is not in general a symmetry transformation in classical mechanics, when the forces between particles depend on their distance. Neither is the one associated with the rotation of the system by an angle proportional to time (change to a non Galilean reference frame, where inertia effects are different). On the other hand, and this will be explained in detail, the translation or rotation by a fixed quantity of an *isolated* system is a symmetry operation (space homogeneity and isotropy).

Here are a number of so-called fundamental symmetries:

– space translations ;

– space rotations ;

– time translations ;

– Lorentz (or Galilean) “relativistic” transformations;

– P (parity, meaning space symmetry with respect to the origin), C (charge conjugation), and T (time reversal);

– exchange between identical particles.

Among these transformations, all are at the present considered as symmetry transformations for the whole set of physical laws governing isolated systems[1], with the exception of P, C, and T. These latter are symmetry transformations if the interactions within the system are electromagnetic (or strong), but no longer if weak interactions play a role (the product CPT nevertheless remains a symmetry operation).

Note that the translation invariance of the evolution of an isolated system is a concept that would be hard to abandon completely as it is almost the definition of a so-called “isolated physical system”. As for the time translations, the basis of physics or even of the scientific method would be destroyed if they were not, at least approximately[2], symmetry transformations for isolated systems: the same experiment performed today or tomorrow would yield different results.

As it is considered at present that all the physical laws known or to be discovered must satisfy all these symmetries, these latter can be viewed as fundamental “*superlaws*” (Wigner), well worth studying.

[1] Obviously, Galilean transformations must also be excluded since they only qualify as symmetry transformations when they are approximations of the Lorentz transformations, in the “non-relativistic” limit (all the velocities $\ll$ light velocity, all the distances $\Delta x \ll c\Delta t$).

[2] In certain cosmological theory, “fundamental” physical constants change (Dirac) as the Universe expands. This changes the group of transformations for which the physical laws are invariant.

Comments:

(i) If both $\mathcal{T}$ and $\mathcal{T}'$ are symmetry transformations, so is their product $\mathcal{T}'\mathcal{T}$ (transformation obtained by applying $\mathcal{T}$ first, and then $\mathcal{T}'$): a group structure is thus expected for the set of symmetry transformations.

(ii) In the case of the time-reversal symmetry, one must actually change in figure 1 $S'(t)$ into $S'(-t)$ and invert the sense of the vertical arrow on the right, which represents the system evolution (*cf.* appendix and its figure 2).

(iii) For an isolated physical system, all the translations and rotations are symmetry transformations. If the physical system is subjected to an external potential (and hence no longer isolated), some of these transformations may still keep their symmetry character. This happens, for example, for the rotations around the origin O of a system subjected to a central potential around O.

A-3. Active and passive points of views

A transformation $\mathcal{T}$ can be defined from two points of views. The first one concerns a single observer in a given reference frame. To any motion of the physical system, the observer associates another motion obtained by a transformation that can be a translation, a rotation, a time delay, etc. As we saw above, $\mathcal{T}$ is a symmetry transformation if the observer can describe both motions with the same dynamical equations. The two motions only differ by their initial conditions. This is the so-called "active" point of view that gives the observer the role of applying the transformation.

One can also adopt another "passive" point of view where a single motion of the system is described by two observers, each using its own reference frame deduced from one another by the transformation $\mathcal{T}$. Each observer will give the system, for example, a different position, or orientation, or velocity, etc. and hence a different mathematical description. $\mathcal{T}$ is a symmetry transformation if the evolutions of the coordinates in the two reference frames are solutions of the same dynamical equations.

In short, the system changes in the active point of view, whereas the reference frame (the axes) changes in the passive point of view. Depending on the case, one or the other point of view will seem more natural. For example, for a time translation it is easy to imagine two different motions delayed in time as in the active point of view. On the other hand, when studying relativity where several Galilean reference frames are used by different observers, the passive point of view is often preferred.

Comments:

(i) As the definition of a translation, rotation, etc. amounts to a change of the space (or time) coordinates of the physical system, and as these coordinates define the relative position of the physical system with respect to the reference frame, it is clear that the active and passive points of view are in fact equivalent. Mathematically, the operations that must be performed on the equations to account for the effects of the transformation $\mathcal{T}$ are identical in both cases. From a physics point of view, one can describe the transformation in the passive point of view with a single motion of the system, but seen from two different reference frames; but one can also switch to the active point of view and introduce a new motion that is seen, in the first reference frame, as the initial motion in the second.

(ii) The physical distinction between these two points of views becomes meaningful if one of the reference frames is prominent with respect to the other. This happens, for example, if there implicitly exists a third reference frame, independent both from $Oxyz$ and from the system under study, like the laboratory reference frame ($Oxyz$ could then be the reference frame of the measuring devices, which, like the system, are mobile with respect to the laboratory). The difference between the two points of views then becomes clear: one moves either the system or the measuring devices.

Another example where the two points of view are not equivalent is when $Oxyz$ is an inertial reference frame, but not its transformed $O'x'y'z'$ since the transformation $\mathcal{T}$ is time-dependent (for instance, a rotation at constant angular velocity). In such a case, the passive point of view is to be preferred in quantum mechanics[3], and we shall use it most of the time.

B. Symmetries in classical mechanics

Let us show that, in classical mechanics, symmetries enforce certain forms for the physical laws, and, in addition, impose constants of motion. We begin with a particularly simple example using Newton's equation, that is, $\boldsymbol{F} = m\boldsymbol{\gamma}$.

[3] A simple example shows the difficulties of the active point of view in that case. It is known that the circulation of an electron's "velocity" (probability current) in a central potential is quantized. An arbitrary increase of the electron's angular velocity in a hydrogen atom is therefore forbidden in quantum mechanics (no quantum family of states exist where this velocity varies continuously).

B-1. Newton's equations

Consider two particles with mass m_1 and m_2, positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, and interacting through a potential $U(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t)$. The equations of motion read:

$$\begin{cases} m_1\, \ddot{\boldsymbol{r}}_1 = \boldsymbol{f}_1\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) = -\boldsymbol{\nabla}_{\boldsymbol{r}_1}\, U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \\ m_2\, \ddot{\boldsymbol{r}}_2 = \boldsymbol{f}_2\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) = -\boldsymbol{\nabla}_{\boldsymbol{r}_2}\, U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \end{cases} \tag{I-1}$$

where $\ddot{\boldsymbol{r}}_1$ represents the second derivative of $\boldsymbol{r}_1$ and $\boldsymbol{\nabla}_{\boldsymbol{r}_1}$ the gradient with respect to the coordinates $\boldsymbol{r}_1$.

• **Translational invariance**

Let $\boldsymbol{b}$ be an arbitrary constant vector. Consider the transformation of the positions at every time t:

$$\begin{cases} \boldsymbol{r}_1 \overset{\mathcal{T}}{\Longrightarrow} \boldsymbol{r}_1 + \boldsymbol{b} \\ \boldsymbol{r}_2 \overset{\mathcal{T}}{\Longrightarrow} \boldsymbol{r}_2 + \boldsymbol{b} \end{cases} \tag{I-2}$$

(the two velocities are then unchanged). It transforms a possible motion into another possible motion (with the same potential U). Since the transformation does not change the accelerations, we have:

$$\begin{cases} \boldsymbol{f}_1\left(\boldsymbol{r}_1 + \boldsymbol{b}, \boldsymbol{r}_2 + \boldsymbol{b}\,; t\right) = \boldsymbol{f}_1\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \\ \boldsymbol{f}_2\left(\boldsymbol{r}_1 + \boldsymbol{b}, \boldsymbol{r}_2 + \boldsymbol{b}\,; t\right) = \boldsymbol{f}_2\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \end{cases} \tag{I-3}$$

This means that, when the two variables increase by a quantity $\boldsymbol{b}$, the gradients of the potential function U with respect to these two variables remains constant. It follows that, in this change of the two variables, U only varies by a constant. This constant could be time-dependent, but has no consequence on the particle's motion since it does not depend on their positions. If we furthermore require U to go to zero at infinity, this constant is necessarily equal to zero. This means that the potential U is invariant in the translation of the two variables, and is therefore only a function of $\boldsymbol{r}_1 - \boldsymbol{r}_2$:

$$U\left(\boldsymbol{r}_1 + \boldsymbol{b}, \boldsymbol{r}_2 + \boldsymbol{b}\,; t\right) \equiv U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \tag{I-4}$$

This restricts the possible potentials, thus imposing a constraint on the form of the physical laws:

$$U \equiv U\left(\boldsymbol{r}_1 - \boldsymbol{r}_2\,; t\right) \tag{I-5}$$

In addition, it is easy to show that $\boldsymbol{f}_1 = -\boldsymbol{f}_2$, which leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m_1\, \dot{\boldsymbol{r}}_1 + m_2\, \dot{\boldsymbol{r}}_2\right) = 0 \tag{I-6}$$

Accordingly, when translations are symmetry transformations, the total momentum is a constant of the motion.

• **Rotational invariance**

If, in addition, rotations are symmetry transformations, other properties appear. Following the same line of reasoning as above, we see that the field of forces $\boldsymbol{f}_1$ (or $\boldsymbol{f}_2$), considered as a function of $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, is invariant under any rotation of the vector $\boldsymbol{r}$. It is thus a field such that $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are parallel to $\boldsymbol{r}_1 - \boldsymbol{r}_2$ and have a modulus that only depends on $|\boldsymbol{r}|$. It follows:

$$U \equiv U\left(|\boldsymbol{r}|\,; t\right) \tag{I-7}$$

and:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left[m_1\ \boldsymbol{r}_1 \times \dot{\boldsymbol{r}}_1 + m_2\ \boldsymbol{r}_2 \times \dot{\boldsymbol{r}}_2\right] &= \boldsymbol{r}_1 \times \boldsymbol{f}_1 + \boldsymbol{r}_2 \times \boldsymbol{f}_2 \\ &= (\boldsymbol{r}_1 - \boldsymbol{r}_2) \times \boldsymbol{f}_1 = 0 \end{aligned} \tag{I-8}$$

The conservation over time of the total angular momentum thus comes from the rotational invariance of all the possible motions.

• **Time translation**

The correspondence between the motions is given by:

$$\begin{cases} \boldsymbol{r}_1(t) \stackrel{\mathcal{T}}{\Longrightarrow} \boldsymbol{r}_1(t+\tau) \\ \boldsymbol{r}_2(t) \stackrel{\mathcal{T}}{\Longrightarrow} \boldsymbol{r}_2(t+\tau) \end{cases} \tag{I-9}$$

where τ is an arbitrary constant (the new motion has a temporal advance of $+\tau$ compared to the initial motion). This time translation is a symmetry transformation if:

$$\begin{cases} \boldsymbol{f}_1\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t+\tau\right) = \boldsymbol{f}_1\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \\ \boldsymbol{f}_2\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t+\tau\right) = \boldsymbol{f}_2\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) \qquad \forall\ \tau \end{cases} \tag{I-10a}$$

The forces depend explicitly on the position but not on the time. It follows that:

$$U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t+\tau\right) = U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t\right) + G(t,\tau) \tag{I-10b}$$

The function G does not have any physical meaning since it does not depend on the positions, and does not change the accelerations. We shall ignore it and, since τ is arbitrary in (I-10b), choose a time-independent potential energy W that describes the possible motions:

$$W\left(\boldsymbol{r}_1, \boldsymbol{r}_2\right) = U\left(\boldsymbol{r}_1, \boldsymbol{r}_2\,; t=0\right) \tag{I-11}$$

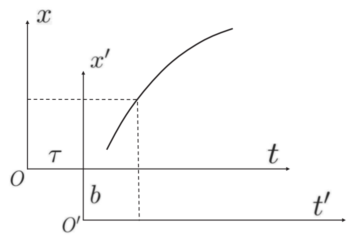

Figure 2: The trajectories of a physical system are symbolized in this figure by the trajectory of a single position as a function of time. Two space–time reference systems Oxt and $O'x't'$ are used to describe this motion; they have a space offset b and a time offset τ. The figure shows a motion described in the second reference system as being in advance, in both time and space, with respect to the description in the first reference system. Nevertheless, the origin O has been moved to O' by a positive amount along the time axis, but by a negative amount along the position axis.

We now compute:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[m_1\,\dot{\boldsymbol{r}}_1^2 + m_2\,\dot{\boldsymbol{r}}_2^2\right] = 2\left(\dot{\boldsymbol{r}}_1\cdot\boldsymbol{f}_1 + \dot{\boldsymbol{r}}_2\cdot\boldsymbol{f}_2\right) = -2\left(\dot{\boldsymbol{r}}_1\cdot\boldsymbol{\nabla}_{\boldsymbol{r}_1}W + \dot{\boldsymbol{r}}_2\cdot\boldsymbol{\nabla}_{\boldsymbol{r}_2}W\right)$$
$$= -2\frac{\mathrm{d}}{\mathrm{d}t}\,W \tag{I-12}$$

which leads to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{m_1}{2}\,\dot{\boldsymbol{r}}_1^2 + \frac{m_2}{2}\,\dot{\boldsymbol{r}}_2^2 + W\right] = 0 \tag{I-13}$$

The total energy (kinetic and potential) is thus a constant of motion.

Comment:

Space and time translations belong to the same category of transformations, even though they show some differences. In the preceding equations, time was not considered as a system's dynamical variable like the position $\boldsymbol{r}(t)$, but as a parameter the dynamical variables depend on. In relativity, however, time also plays the role of a coordinate that is needed, in addition to the three components of $\boldsymbol{r}$, to define an event in space–time. One should then be careful about signs, which may differ in both cases.

To see why in a simple case, let us consider a one-dimensional space. Figure 2 shows the motion of a particle seen by two different observers placed in reference frames Oxt and $O'x't'$. These reference frames have a space offset b and a time offset τ. The two offsets are counted positive if the second observer sees a motion that has progressed in space and time with respect to the first observer – this is the case shown in the figure. The same event is then described, either by the coordinates (x, t), or by the coordinates (x', t') given by:

$$
\begin{aligned}
x' &= x + b \\
t' &= t - \tau
\end{aligned}
\tag{I-14a}
$$

We notice an opposite sign between the space and time variations. The minus sign for the time is not surprising: if the event "the train arrives at the station" occurs at ten to twelve instead of twelve, the train is ten minutes early.

The sign of τ is also different in (I-9) and (I-14a). The reason lies in the difference between the two points of view mentioned above: in the first case, time was considered as a parameter the dynamical variables depend on, in the second case, as the coordinate of a space–time event.

Starting from equation (I-14a), we can recover the plus sign of (I-9) . The same motion is described in the first reference frame by a function $x(t)$ of time t, and in the second reference frame by a different function $x'(t')$. If times t and t' are such that $t' = t - \tau$, they relate to the same event, so that $x' = x + b$, and therefore:

$$
x'(t') = x(t) + b = x(t' + \tau) + b \tag{I-14b}
$$

Time τ now follows a plus sign. The conclusion is that, when time plays the role of a parameter the position depends on, one has to add (and not subtract) τ to obtain a time progression, in agreement with (I-9).

We shall go no further in the symmetry studies around Newton's equations as these are not a convenient starting point for the quantization of a physical system. It is best to use either the Lagrangian or the Hamiltonian formalism.

B-2. Lagrange's equations

B-2-a. General formalism

The system is described[4] by a set of generalized coordinates q_i, with $i = 1, 2, \ldots, N$, which define its configuration $\mathscr{C}$ (distinct from its state S, which also contains velocities); it can also depend on a number of parameters λ_α (particles' masses, charges, etc.). One associates with that system a function called

[4] A more detailed presentation can be found, for example, in the appendix III of reference [9], or in a standard book on analytical mechanics.

"*Lagrangian*", which depends on all the coordinates q_i and q'_i:

$$L \equiv L(q_i, \dot{q}_i \,; t \,; \lambda_\alpha) \tag{I-15}$$

and such that the equations of motion are given by the Lagrangian equations:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \qquad [i = 1, 2, \ldots, N]} \tag{I-16}$$

In these equations, $\dot{q}_i$ stands for the time derivative of q_i. Remember that the total time derivative of a function $K\,(q_i, \dot{q}_i \,; t)$ is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, K\,(q_i, \dot{q}_i \,; t) = \frac{\partial K}{\partial t} + \sum_i \dot{q}_i \frac{\partial K}{\partial q_i} + \sum_i \ddot{q}_i \frac{\partial K}{\partial \dot{q}_i} \tag{I-17}$$

Like Newton's equations, the Lagrangian equations are of second-order with respect to the time.

They are equivalent to a principle of least action that yields, in a global way (rather than local in time), the possible motions of the physical system. This principle states that, among all the possible motions leading the system at time t_1 from the configuration $\mathscr{C}_1$ (symbolizing the set of all q_i) to the configuration $\mathscr{C}_2$ at time t_2, the only realizable motion (that which satisfies the equations of motion) must have a stationary action $\mathscr{A}$:

$$\mathscr{A} = \int_{t_1}^{t_2} \mathrm{d}t \;\, L\,[q_i(t), \dot{q}_i(t) \,; t] \tag{I-18}$$

Figure 3, where the set of coordinates q_i is symbolized by a single ordinate axis q, shows with dashed lines several a priori possible paths, and with a solid line the path actually followed by the system, as it minimizes $\mathscr{A}$.

The primary advantage of the Lagrangian point of view is its great generality, as, with a proper choice of the function L, the equations of motion of a large number of physical systems can be put in the form (I-16). Furthermore, whatever the variables q_i chosen to describe the system (for example, taking spherical instead of Cartesian coordinates, etc.), the system's equations of motion keep the same form (I-16), which is not the case for Newton's equations.

For a system of particles of mass m_n, with an interaction potential $V(\boldsymbol{r}_n \,; t)$, the Lagrangian can be written:

$$L = T\,(\dot{\boldsymbol{r}}_1, ..., \dot{\boldsymbol{r}}_n, ...) - V(\boldsymbol{r}_1, ..., \boldsymbol{r}_n \, ...; t) \tag{I-19}$$

Here, the 3 components of the particles' position vectors $\boldsymbol{r}_n$ play the role of the q_i, and T is the kinetic energy:

$$T\,(\dot{\boldsymbol{r}}_1, ..., \dot{\boldsymbol{r}}_n, ...) = \frac{1}{2} \sum_n m_n \, \dot{\boldsymbol{r}}_n^2 \tag{I-20}$$

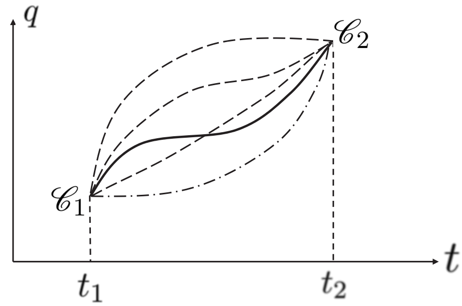

Figure 3: Two successive configurations of a physical system, $\mathscr{C}_1$ and $\mathscr{C}_2$, are represented in a symbolic way by the value of a single position variable q on the vertical axis. These states being fixed, various "paths" (set of states at all the intermediary times) going from $\mathscr{C}_1$ to $\mathscr{C}_2$ are shown with dashed lines. None of these virtual paths will be followed by the system for real, except the one, drawn with a solid line, that makes action stationary.

In this example, the forces can be obtained by taking the derivative of a potential, but Lagrangian formalism can be used in more general cases. One can, for example, compute the evolution of a system of particles interacting with a given electromagnetic field described by a scalar potential V and a vector potential $\boldsymbol{A}$. A possible Lagrangian is then (*cf.*, for example, Appendix III, § 4.b, of [9]):

$$L = \sum_n \left\{ \frac{1}{2} m_n \dot{\boldsymbol{r}}_n^2 + q_n^{\mathrm{e}} \boldsymbol{A}(\boldsymbol{r}_n, t) \cdot \dot{\boldsymbol{r}} - q_n^{\mathrm{e}} V(\boldsymbol{r}_n, t) \right\} \tag{I-21}$$

(where q_n^{e} is the electric charge of the n^{th} particle). As mentioned above, many other equations of motion (where the values of the fields at each point of space become dynamical variables, playing the role of the q_i and $\dot{q}_i$) can be obtained from Lagrange's equations and a variational principle. This is the case, for example, of Maxwell's equations.

Comments:

(i) It should not be assumed that a unique Lagrangian corresponds to the equations of motion of a given physical system. There are actually many "equivalent" Lagrangians. For example, if Λ is an arbitrary function $\Lambda(q_i, t)$, one can, starting from a Lagrangian L obtain another[5] Lagrangian L':

$$L'(q_i, \dot{q}_i\,; t) = L(q_i, \dot{q}_i; t) + \frac{\mathrm{d}}{\mathrm{d}t}\,\Lambda(q_i\,; t) \tag{I-22}$$

The variation δL of the Lagrangian is:

$$\delta L = L' - L = \frac{\partial \Lambda}{\partial t} + \sum_j \dot{q}_j \frac{\partial \Lambda}{\partial q_j} \tag{I-23}$$

and:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial}{\partial \dot{q}_i}\delta L\right) = \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial \Lambda}{\partial q_i} = \dfrac{\partial^2 \Lambda}{\partial q_i \partial t} + \displaystyle\sum_j \dot{q}_j \dfrac{\partial^2 \Lambda}{\partial q_i \partial q_j} \\ \dfrac{\partial}{\partial q_i}\delta L = \dfrac{\partial^2 \Lambda}{\partial q_i \partial t} + \displaystyle\sum_j \dot{q}_j \dfrac{\partial^2 \Lambda}{\partial q_i \partial q_j} \end{cases} \tag{I-24}$$

This means that the contributions of δL to each side of equation (I-16) are identical and hence compensate each other. Lagrangians that lead to the same differential equations of motion are called "equivalent Lagrangians".

Another way to check that L and L' are equivalent is to note that the corresponding variation of the action is written:

$$\delta\mathscr{A} = \int_{t_1}^{t_2} \mathrm{d}t\,\frac{\mathrm{d}\Lambda}{\mathrm{d}t} = \Lambda\Big[q_i(t_2)\,; t_2\Big] - \Lambda\Big[q_i(t_1)\,; t_1\Big] \tag{I-25}$$

The variation $\delta\mathscr{A}$ only depends on S_1, S_2 and t_1, t_2, and not on the path followed by the system between S_1 and S_2. It follows that $\mathscr{A}$ and $\mathscr{A} + \delta\mathscr{A}$ will always be stationary for the same paths.

(ii) Reciprocally, it is sometimes suggested that the difference between two equivalent Lagrangians must be a total derivative with respect to time. This is simply not true.

A first simple counter-example is to multiply the Lagrangian by an arbitrary constant. Actually, the ensemble of the equivalent Lagrangians is in general much larger. For example, for a free particle, one can choose either L or L' given by:

$$\begin{aligned} L &= \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \\ L' &= \alpha\dot{x}^2 + \beta\dot{y}^2 + \gamma\dot{z}^2 \end{aligned}$$

where α, β, and γ are arbitrary constants.

It might be interesting to establish a general procedure to find all the Lagrangians

[5] Note that Λ should not depend on the $\dot{q}_i$ or else second derivatives of the q_i would appear, which is not possible for a Lagrangian.

equivalent to a given Lagrangian. This would enable one to find the necessary and sufficient conditions for a given transformation $\mathcal{T}$ to be a symmetry transformation. Furthermore, it would allow seeing if the quantization obtained from these Lagrangians leads to the same physical results. This general problem does not seem to have been solved at the present.

B-2-b. Constants of motion; Noether's theorem

α. Simple cases

The invariance properties of L can lead to the existence of constants of motion. For example, imagine the Lagrangian (I-19) is invariant under the translation of the whole set of particles by a given value $\boldsymbol{b}$ (substitution $\boldsymbol{r}_n \Longrightarrow \boldsymbol{r}_n + \boldsymbol{b}$). In that case:

$$\sum_n \boldsymbol{\nabla}_{\boldsymbol{r}_n} L = 0 \tag{I-26}$$

leads to, according to (I-16):

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_n \boldsymbol{p}_n(t) = 0 \tag{I-27}$$

where $\boldsymbol{p}_n$ is defined as:

$$\boldsymbol{p}_n(t) = \boldsymbol{\nabla}_{\dot{\boldsymbol{r}}_n} L = m_n \dot{\boldsymbol{r}}_n(t) \tag{I-28}$$

Conservation of the total momentum is thus a consequence of space translation invariance of the Lagrangian[6].

Similarly, if L is invariant under a time translation:

$$\frac{\partial L}{\partial t} = 0 \tag{I-29}$$

Defining the function H as:

$$H(t) = \sum_n \dot{\boldsymbol{r}}_n(t) \cdot \boldsymbol{p}_n(t) - L\Big(\boldsymbol{r}_1(t), ..., \boldsymbol{r}_n(t)...; \dot{\boldsymbol{r}}_1(t), ..., \dot{\boldsymbol{r}}_n(t), ...\Big) \tag{I-30}$$

it is easy to show, using (I-16) and (I-28) that, as the system evolves, H remains constant:

$$\frac{\mathrm{d}}{\mathrm{d}t} H(t) = \sum_n \left\{ \ddot{\boldsymbol{r}}_n \cdot \boldsymbol{p}_n + \dot{\boldsymbol{r}}_n \cdot \boldsymbol{\nabla}_{\boldsymbol{r}_n} L - \dot{\boldsymbol{r}}_n \cdot \boldsymbol{\nabla}_{\boldsymbol{r}_n} L - \ddot{\boldsymbol{r}}_n \cdot \boldsymbol{p}_n \right\} = 0 \tag{I-31}$$

[6] The total momentum conservation results from the invariance of L when all the interacting particles undergo the same translation. An invariance of L under the translation of a single particle would lead to the conservation of that particle's momentum (particle without interaction).

Time translation invariance thus leads to the energy H being a constant of motion.

In the two examples above, we have assumed that L is invariant, but the invariance of the equations of motion in a transformation $\mathcal{T}$ does not forcibly imply that the Lagrangian itself is invariant under this transformation. Another possibility (*cf.* comment above) is that the transformation $\mathcal{T}$ adds to L a total derivative with respect to time. Emily Noether has shown in 1918 that, in this case as well, the invariance of the transformation $\mathcal{T}$ also leads to the existence of a constant of motion.

β. *General demonstration of the theorem*

Consider a transformation $\mathcal{T}$ of the generalized coordinates q_i of an arbitrary system, that can be written as:

$$q_i \overset{\mathcal{T}}{\Longrightarrow} q_i + \delta q_i \tag{I-32}$$

The transformation $\mathcal{T}$ is supposed to be infinitesimal and:

$$\delta q_i = \delta\varepsilon \; f_i\left(q_j, \dot{q}_j \,; t\right) \tag{I-33a}$$

where $\delta\varepsilon$ is an infinitely small quantity. The variations $\delta\dot{q}_i$ of the time derivatives of the q_i are given by:

$$\delta\dot{q}_i = \delta\varepsilon \; g_i\left(q_j, \dot{q}_j, \ddot{q}_j \,; t\right) \tag{I-33b}$$

Note that g_i, contrary to f_i, may depend on $\ddot{q}_j$, as can be seen from the definition of $\delta\dot{q}_i$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(q_i + \delta q_i\right) = \dot{q}_i + \delta\dot{q}_i \tag{I-34}$$

which leads to:

$$g_i = \frac{\mathrm{d}}{\mathrm{d}t}\, f_i = \frac{\partial f_i}{\partial t} + \sum_j \left(\dot{q}_j \,\frac{\partial f}{\partial q_j} + \ddot{q}_j \,\frac{\partial f}{\partial \dot{q}_j}\right) \tag{I-35}$$

In the transformation $\mathcal{T}$, the infinitesimal variation δL of the Lagrangian is written:

$$\delta L = \sum_i \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right) = \delta\varepsilon \;\sum_i \left(f_i \frac{\partial L}{\partial q_i} + g_i \frac{\partial L}{\partial \dot{q}_i}\right) \tag{I-36}$$

If the transformation $\mathcal{T}$ is chosen in such a way that δL is proportional to the total time derivative of a function Λ:

$$\delta L = \delta\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\;\Lambda \tag{I-37}$$

Noether's theorem states that the function:

$$\boxed{F = \sum_i f_i \frac{\partial L}{\partial \dot{q}_i} - \Lambda} \tag{I-38}$$

is a constant of motion; $\mathrm{d}F/\mathrm{d}t$ is zero along all the possible trajectories of the physical system.

Demonstration: the total time derivative of $\sum_i f_i \, \partial L/\partial \dot{q}_i$ is:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} \sum_i \left[f_i \frac{\partial L}{\partial \dot{q}_i} \right] &= \sum_i \left[g_i \frac{\partial L}{\partial \dot{q}_i} + f_i \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right] \\ &= \sum_i \left(g_i \frac{\partial L}{\partial \dot{q}_i} + f_i \frac{\partial L}{\partial q_i} \right)\end{aligned} \tag{I-39}$$

(the second equality, obtained from the equations of motion, is valid along any possible trajectory of the system). It follows, using (I-36) and (I-37):

$$\frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_i f_i \frac{\partial L}{\partial \dot{q}_i} \right] = \frac{\delta L}{\delta \varepsilon} = \frac{\mathrm{d}}{\mathrm{d}t} \Lambda \tag{I-40}$$

Inserting this result into the time derivative of (I-38) leads to $\mathrm{d}F/\mathrm{d}t = 0$, which proves the theorem.

Comments:

(i) If Λ depends only on the coordinates q_i and of the time t, the variation δL of the Lagrangian depends on the q_i, the $\dot{q}_i$, and of time; $L + \delta L$ then provides a Lagrangian that is equivalent to L. But, if Λ depends on the $\dot{q}_i$, the function $L + \delta L$ also depends on the $\ddot{q}_i$, which is no longer compatible with the standard definition of a Lagrangian. The Noether theorem nevertheless remains valid in this case; see, for instance, the first example below.

(ii) One may have to use Lagrange's equations to go from (I-36) to (I-37). As an example, they can be used to express the $\ddot{q}_j$ appearing in general in (I-36), as a function of the q_i and $\dot{q}_i$; one can then look for a function $\Lambda(q_i, t)$ independent of the $\dot{q}_i$ (and whose total time derivative does not contain the $\ddot{q}_i$).

(iii) Obviously, Noether's theorem is only interesting if it yields a non-trivial constant of motion F, and not, for example, a zero or a constant independent of the generalized coordinates!

If one accepts functions such as $\Lambda(q_i, \dot{q}_i, t)$, there is a greater chance of obtaining zero information. A trivial case is to take completely arbitrary functions

f_i , and write δL in the form (I-37) choosing the function $\Lambda = \sum_i f_i \, \partial L/\partial \dot{q}_i$ (this is always possible since (I-40) simply states that, along the trajectories, δL is always proportional to the total derivative of that function Λ). We then get the trivial result $F \equiv 0$.

γ. *Examples*

Here are a few simple examples of the application of Noether's theorem; the theorem is also valid in field theory, where it has important applications – *cf.* complement B_I.

• Consider first an arbitrary system whose Lagrangian is not explicitly time-dependent. As a transformation law, we choose:

$$q_i(t) \stackrel{\mathcal{T}}{\Longrightarrow} q_i(t+\delta t) \tag{I-41a}$$

where δt is a constant that plays the role of $\delta\varepsilon$. Operation $\mathcal{T}$ shifts the time evolution of the system; we are thus dealing with a time translation. As δt is infinitesimal, we can write:

$$\delta q_i(t) = \dot{q}_i \delta t \tag{I-41b}$$

so that:

$$\begin{cases} f_i = \dot{q}_i \\ g_i = \ddot{q}_i \end{cases} \tag{I-42}$$

On the other hand:

$$\delta L = \delta t \sum_i \left[\dot{q}_i \frac{\partial L}{\partial q_i} + \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right] = \delta t \frac{\mathrm{d}L}{\mathrm{d}t} \tag{I-43}$$

(since by hypothesis $\partial L/\partial t = 0$). We find that the function Λ is none other than the Lagrangian itself. According to (I-38), the constant of motion is:

$$F = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L \tag{I-44}$$

which is simply the usual definition of the Hamiltonian H (energy).

• Let us now go back to the example discussed in § B-2-b above, and find which constants of motion derive from the invariance under space translation, or under a change of Galilean reference frame. The Lagrangian is written as:

$$L = \sum_n \frac{m_n}{2} \dot{\boldsymbol{r}}_n^2 - V(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n \ldots) \tag{I-45}$$

where the interaction potential V is invariant under the translation of all the positions (as usual, we have replaced the q_i by the $\boldsymbol{r}_n$, or more precisely by the three components of these vectors). The invariance of L under the transformation $\boldsymbol{r}_n \Longrightarrow \boldsymbol{r}_n + \delta\boldsymbol{b}$ readily yields the conservation of the total momentum. The 3 components of the infinitesimal vector $\delta\boldsymbol{b}$ now play the role of 3 infinitely small $\delta\varepsilon$; setting $f_i = 1$, $g_i = 0$, $\delta L = 0$, yield as constants of motion the 3 components of the vector:

$$\boldsymbol{P} = \sum_n \boldsymbol{\nabla}_{\dot{\boldsymbol{r}}_n} L = \sum_n \boldsymbol{p}_n \tag{I-46}$$

- We now introduce a change of Galilean reference frame by:

$$\boldsymbol{r}_n \Longrightarrow \boldsymbol{r}_n + t\,\delta\boldsymbol{v} \tag{I-47}$$

Relations (I-33) then become:

$$\delta\boldsymbol{r}_n = t\delta\boldsymbol{v} \qquad\qquad \delta\dot{\boldsymbol{r}}_n = \delta\boldsymbol{v} \tag{I-48}$$

so that we have $f_i = t$ and $g_i = 1$ for $i = 1, 2, 3$; the 3 components of the vector $\delta\boldsymbol{v}$ now play the role of $\delta\varepsilon$. In this case, δL is written:

$$\delta L = \delta\boldsymbol{v} \cdot \sum_n \left(m_n\,\dot{\boldsymbol{r}}_n - t\,\boldsymbol{\nabla}_{\boldsymbol{r}_n} V\right) \tag{I-49}$$

If V is translation invariant as in (I-26), the sum of all its gradients is equal to zero, and:

$$\delta L = \delta\boldsymbol{v} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{G} \tag{I-50}$$

where:

$$\boldsymbol{G}(t) = \sum_n m_n\,\boldsymbol{r}_n(t) \tag{I-51}$$

$\boldsymbol{G}$ is the position of all the particles' center of mass, multiplied by the sum of masses. Replacing Λ by $\boldsymbol{G}$ in (I-38) yields the (vectorial) constant of the motion:

$$\boldsymbol{G}_0 = t\sum_n m_n\,\dot{\boldsymbol{r}}_n - \sum_n m_n\,\boldsymbol{r}_n = \boldsymbol{P}t - \boldsymbol{G}(t) \tag{I-52}$$

Hence:

$$\boldsymbol{G}(t) = \boldsymbol{P}t - \boldsymbol{G}_0 \tag{I-53}$$

Dividing this equation by the sum of masses we check that, as expected, the particles' center of mass moves at constant velocity.

Finally, imagine that V is not translation invariant, but that the sum of the (outside) forces acting on the particles is a constant vector $\boldsymbol{F}$. The variation of L is then:

$$\delta L = \delta L = \delta \boldsymbol{v} \cdot \Big(\sum_n m_n \, \dot{\boldsymbol{r}}_n + t \, \boldsymbol{F} \Big) = \delta \boldsymbol{v} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\Lambda} \tag{I-54a}$$

with:

$$\boldsymbol{\Lambda}(t) = \boldsymbol{G}(t) + \frac{1}{2} \, t^2 \, \boldsymbol{F} \tag{I-54b}$$

The constant of motion is now:

$$\boldsymbol{G}_0 = \boldsymbol{P} t - \boldsymbol{G}(t) - \frac{1}{2} \, \boldsymbol{F} \, t^2 = \boldsymbol{P}_0 \, t + \frac{1}{2} \, \boldsymbol{F} \, t^2 - \boldsymbol{G}(t) \tag{I-55}$$

where, in the second equality, we have taken into account the linear time variation of the total momentum $\boldsymbol{P} = \boldsymbol{P}_0 + \boldsymbol{F} t$. As a result, we have to add a term $(t^2/2)\boldsymbol{F}$ to the right-hand side of (I-53), which shows that the center of mass now moves with a constant acceleration.

Exercise:

Consider a particle of mass m, position $\boldsymbol{r}$, momentum $\boldsymbol{p} = m\dot{\boldsymbol{r}}$, and angular momentum $\boldsymbol{\ell} = \boldsymbol{r} \times \boldsymbol{p}$. This particle is subjected to a central potential $V = -\alpha/r^n$ (n is a positive integer). Introducing the transformation

$$\delta \boldsymbol{r} = \delta \boldsymbol{\varepsilon} \times \boldsymbol{\ell} \tag{I-56}$$

show that:

$$\delta L = -m \frac{n\alpha}{r^{n+2}} \left[\boldsymbol{r}^2 \left(\delta \boldsymbol{\varepsilon} \cdot \dot{\boldsymbol{r}} \right) - \left(\delta \boldsymbol{\varepsilon} \cdot \boldsymbol{r} \right) \left(\boldsymbol{r} \cdot \dot{\boldsymbol{r}} \right) \right] \tag{I-57}$$

In the case of a Coulomb or Newton potential ($n = 1$), show that:

$$\delta L = \frac{\mathrm{d}}{\mathrm{d}t} \left[-m\alpha \, \delta \boldsymbol{\varepsilon} \cdot \frac{\boldsymbol{r}}{r} \right] \tag{I-58}$$

and that, consequently, the vector $\boldsymbol{M}$ (Runge–Lenz vector):

$$\boldsymbol{M} = \boldsymbol{p} \times (\boldsymbol{r} \times \boldsymbol{p}) - m\alpha \frac{\boldsymbol{r}}{r} \tag{I-59}$$

is a constant. Physical interpretation: the points of the particle's (plane) trajectory where its velocity is perpendicular to $\boldsymbol{r}$ are fixed. Instead of following a "rosette pattern", the trajectory is a closed curve. It is actually an ellipse whose major axis is parallel to $\boldsymbol{M}$ and whose eccentricity equals $|\boldsymbol{M}|/\alpha m$.

Comment:

In (I-33), we assumed that the transformation only concerned the q_i (and the $\dot{q}_i$), but not the time. This restriction can be lifted by introducing, in addition to the variations δq_i and $\delta \dot{q}_i$ written in (I-33), a time variation:

$$\delta t = \delta\varepsilon\, h\,(q_j, \dot{q}_j\,; t) \tag{I-60}$$

Relation (I-35) is no longer valid in this case. Writing:

$$(\dot{q}_i + \delta\dot{q}_i)\; \mathrm{d}\,(t + \delta t) = \mathrm{d}\,[q_i + \delta q_i] \tag{I-61}$$

we readily obtain, to first-order in $\delta\varepsilon$:

$$g_i\, \mathrm{d}t + \dot{q}_i\, \mathrm{d}h = \mathrm{d}f_i \tag{I-62}$$

meaning:

$$g_i = \frac{\mathrm{d}f_i}{\mathrm{d}t} - \dot{q}_i\, \frac{\mathrm{d}h}{\mathrm{d}t} \tag{I-63}$$

We also have:

$$\delta L = \delta\varepsilon \left\{ \sum_i \left(f_i\, \frac{\partial L}{\partial q_i} + g_i\, \frac{\partial L}{\partial \dot{q}_i} \right) + h\, \frac{\partial L}{\partial t} \right\} \tag{I-64}$$

Writing δL as:

$$\delta L = \delta\varepsilon \left[\frac{\mathrm{d}}{\mathrm{d}t}\Lambda + L\, \frac{\mathrm{d}h}{\mathrm{d}t} \right] \tag{I-65a}$$

it can be shown that the function:

$$F = \sum_i \frac{\partial L}{\partial \dot{q}_i}\, [f_i - h\dot{q}_i] + hL - \Lambda \tag{I-65b}$$

is a constant of motion. Using Lagrange's equation to get the total time derivative of $\partial L/\partial \dot{q}_i$ and (I-65a) to get that of Λ, one shows that:

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}t} = &\sum_i \frac{\partial L}{\partial q_i}\, [f_i - h\dot{q}_i] + \sum_i \frac{\partial L}{\partial \dot{q}_i} \left[\frac{\mathrm{d}f_i}{\mathrm{d}t} - \dot{q}_i\, \frac{\mathrm{d}h}{\mathrm{d}t} - h\ddot{q}_i \right] \\ &+ L\, \frac{\mathrm{d}h}{\mathrm{d}t} + h \left[\frac{\partial L}{\partial t} + \sum_i \left(\dot{q}_i\, \frac{\partial L}{\partial q_i} + \ddot{q}_i\, \frac{\partial L}{\partial \dot{q}_i} \right) \right] \quad - \frac{\delta L}{\delta\varepsilon} - L\, \frac{\mathrm{d}h}{\mathrm{d}t} \end{aligned} \tag{I-66}$$

Using relation (I-64) to compute $\delta L/\delta\varepsilon$, and finally replacing g_i by its value (I-63), all the terms in the right-hand side of this equality cancel each other two by two. A is indeed a constant of motion.

Exercise: Setting $f_i = g_i = 0$, $h = 1$, show, as in (I-44), that H is a constant of the motion if $\partial L/\partial t = 0$.

B-3. Hamilton's equations

Before leaving classical mechanics, let us quickly review the Hamiltonian formalism, which is often used for the "quantization" of a physical system. Note that another quantization procedure starts from the Lagrangian formalism [10] (Feynman's postulates), and is often used, especially in field theory.

B-3-a. General formalism

To each generalized coordinate of the system q_i we associate the conjugate momentum:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{I-67}$$

We assume that all the $\dot{q}_i$ can be expressed in terms of the p_i and q_i, so that the dynamical state of the system can be defined either by the set of all the q_i and $\dot{q}_i$, or by the set of q_i and p_i. In Hamilton's point of view, the q_i and p_i are chosen as the independent variables. They are the coordinates of a point defining the state of the physical system in a space called "phase space", with a large dimension if there are a large number of particles. We introduce the Hamiltonian:

$$H\left(p_i, q_i\right) = \sum_i p_i\, \dot{q}_i\left(q_j, p_j\right) - L\left[q_i, \dot{q}_i\left(q_j, p_j\right)\right] \tag{I-68}$$

where, in the right-hand side, the notation $\dot{q}_i\left(q_j, p_j\right)$ explicits the fact that the derivatives of the q_i are functions of the variables q_j and p_j for any j. From now on we shall simply write $\dot{q}_i$. The differential of H is written:

$$\mathrm{d}H\left(p_i, q_i\right) = \sum_i \left[\dot{q}_i\, \mathrm{d}p_i + p_i\, \mathrm{d}\dot{q}_i\right] - \frac{\partial L}{\partial q_i}\mathrm{d}q_i - \frac{\partial L}{\partial \dot{q}_i}\mathrm{d}\dot{q}_i \tag{I-69a}$$

Using definition (I-67) for p_i, we see that the $\mathrm{d}\dot{q}_i$ terms on the right-hand side cancel out. Using Lagrange's equations, we can replace $\partial L/\partial q_i$ by $\dot{p}_i$ and write:

$$\mathrm{d}H\left(p_i, q_i\right) = \sum_i \left[\dot{q}_i\, \mathrm{d}p_i - \dot{p}_i \mathrm{d}q_i\right] \equiv \sum_i \left[\frac{\partial H}{\partial p_i}\, \mathrm{d}p_i + \frac{\partial H}{\partial q_i}\mathrm{d}q_i\right] \tag{I-69b}$$

Identifying the terms in $\mathrm{d}p_i$ and $\mathrm{d}q_i$ leads to the Hamilton's equations of motion:

$$\boxed{\begin{cases} \dot{q}_i = \dfrac{\partial H}{\partial p_i} \\[2ex] \dot{p}_i = -\dfrac{\partial H}{\partial q_i} \end{cases}} \tag{I-70}$$

Instead of N second-order differential equations (Lagrangian point of view), we now have $2N$ equations between $2N$ variables, but these are first-order differential equations.

B-3-b. Poisson bracket

Consider two physical quantities $A(q_i, p_i)$ and $B(p_i, q_i)$ defined using Hamilton's point of view. The Poisson bracket $\{A, B\}$ is defined as:

$$\{A, B\} = \sum_i \left[\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i}\frac{\partial A}{\partial p_i}\right] \tag{I-71}$$

It's easy to show that:

$$\begin{aligned} &\{q_i, p_j\} = \delta_{ij} \\ &\{q_i, q_j\} = \{p_i, p_j\} = 0 \end{aligned} \tag{I-72}$$

as well as:

$$\{A, A\} = 0 \tag{I-73a}$$

$$\{A, B\} = -\{B, A\} \tag{I-73b}$$

$$\{A, BC\} = \{A, B\}\, C + B\,\{A, C\} \tag{I-73c}$$

(they follow the same rules as the commutators).

According to (I-70), the time evolution of any given physical quantity $F(q_i, p_i\,; t)$ can be written as:

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial F}{\partial t} + \sum_i \left(\dot{q}_i\,\frac{\partial F}{\partial q_i} + \dot{p}_i\,\frac{\partial F}{\partial p_i}\right) = \frac{\partial F}{\partial t} + \{F, H\} \tag{I-74}$$

The Poisson bracket thus yields directly the time evolution of any function F of the position and momentum. A constant of motion is therefore described by a function F whose Poisson bracket $\{F, H\}$ with the Hamiltonian vanishes.

In certain cases, the Poisson brackets also yield new constants of motion: using (I-73c) and (I-74) we see that, if both F and G are constants of the motion, their Poisson bracket is also a constant of the motion (Poisson's theorem)[7].

Exercise: Show that expression:

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C\,\{A, B\}\} = 0 \tag{I-75}$$

equals zero (Jacobi identity). Infer from this result Poisson's theorem.

[7] This Poisson bracket may simply be equal to zero or to a constant independent of the q and p, or else be proportional to F or G; in such cases, Poisson's theorem does not bring any new information.

B-3-c. Infinitesimal transformations

The Poisson brackets can also yield the transformation of a system's motion in an infinitesimal operation $\mathcal{T}$. Let us show this on the concrete example of a particle of mass m whose motion is given by the values of its 4 space–time coordinates as a function of a parameter u: $x = f_1(u),\ y = f_2(u),\ z = f_3(u),\ t = f_4(u)$.

• Spatial translation by an infinitesimal constant vector $\delta\boldsymbol{b}$. By definition of this translation:

$$x' = x + \delta b_x \qquad y' = y + \delta b_y \qquad z' = z + \delta b_z \qquad t' = t \qquad \boldsymbol{p}' = \boldsymbol{p} \tag{I-76}$$

Now:

$$\{p_x, x\} = -1 \qquad \{p_y, x\} = \{p_z, x\} = 0 \tag{I-77}$$

and consequently:

$$\{\delta\boldsymbol{b}\cdot\boldsymbol{p}, x\} = -\delta b_x \tag{I-78}$$

We can thus write:

$$\begin{cases} \boldsymbol{r}' = \boldsymbol{r} - \{\delta\boldsymbol{b}\cdot\boldsymbol{p}, \boldsymbol{r}\} \\ \boldsymbol{p}' = \boldsymbol{p} - \{\delta\boldsymbol{b}\cdot\boldsymbol{p}, \boldsymbol{p}\} = \boldsymbol{p} \end{cases} \tag{I-79}$$

which shows that the Poisson brackets of $\delta\boldsymbol{b}\cdot\boldsymbol{p}$ with the dynamical variables yield infinitesimal translations.

• Spatial rotation of an infinitesimal angle $\delta\varphi$ around the unit vector $\boldsymbol{u}$. This rotation is associated with an infinitesimal (time-independent) vector $\delta\boldsymbol{a}$:

$$\delta\boldsymbol{a} = \boldsymbol{u}\,\delta\varphi \tag{I-80}$$

We can write:

$$\begin{cases} \delta\boldsymbol{r} = \boldsymbol{r}' - \boldsymbol{r} = \delta\boldsymbol{a}\times\boldsymbol{r} + \dots \\ \delta\boldsymbol{p} = \boldsymbol{p}' - \boldsymbol{p} = \delta\boldsymbol{a}\times\boldsymbol{p} + \dots \end{cases} \tag{I-81}$$

We now introduce the (angular momentum) vector:

$$\boldsymbol{\ell} = \boldsymbol{r}\times\boldsymbol{p} \tag{I-82}$$

This yields:

$$\begin{aligned} \{\delta\boldsymbol{a}\cdot\boldsymbol{\ell}\ ,\ x\} &= -\delta a_y\, z + \delta a_z\, y \\ \{\delta\boldsymbol{a}\cdot\boldsymbol{\ell}\ ,\ p_x\} &= -\delta a_y\, p_z + \delta a_z\, p_y \end{aligned} \tag{I-83}$$

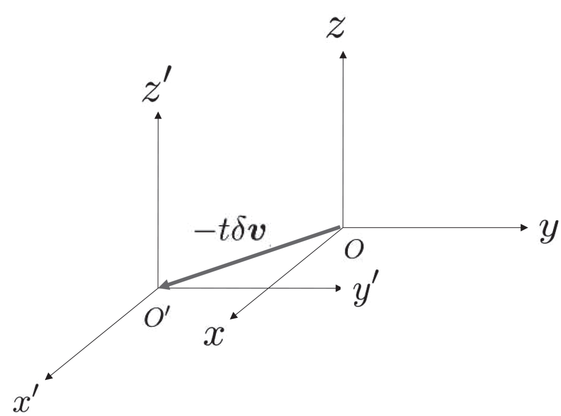

Figure 4: Two Galilean reference frames $Oxyz$ and $O'x'y'z'$ have parallel axes, and their origins O and O' are linked by a vector equal to $-t\delta\boldsymbol{v}$, where $\delta\boldsymbol{v}$ is an infinitesimal constant vector.

and, consequently:

$$\begin{cases} \boldsymbol{r}' = \boldsymbol{r} - \{\delta\boldsymbol{a}\cdot\boldsymbol{\ell}\ ,\ \boldsymbol{r}\} + \dots \\ \boldsymbol{p}' = \boldsymbol{p} - \{\delta\boldsymbol{a}\cdot\boldsymbol{\ell}\ ,\ \boldsymbol{p}\} + \dots \end{cases} \tag{I-84}$$

The angular momentum $\boldsymbol{\ell}$ thus plays for the rotations the same role as $\boldsymbol{p}$ for the translations.

• Time translation

Consider two motions defined by $\boldsymbol{r}(t)$ and $\boldsymbol{r}'(t)$, the second being ahead of the first one by a time interval δt (figure 2), as in relation (I-9). We then have:

$$\begin{cases} \boldsymbol{r}'(t) = \boldsymbol{r}(t+\delta t) \simeq \boldsymbol{r}(t) + \delta t\,\dot{\boldsymbol{r}}(t) \\ \boldsymbol{p}'(t) = \boldsymbol{p}(t+\delta t) \simeq \boldsymbol{p}(t) + \delta t\,\dot{\boldsymbol{p}}(t) \end{cases} \tag{I-85}$$

According to (I-74), and to first-order in δt:

$$\begin{cases} \boldsymbol{r}' = \boldsymbol{r} - \{\delta t\,H, \boldsymbol{r}\} \\ \boldsymbol{p}' = \boldsymbol{p} - \{\delta t\,H, \boldsymbol{p}\} \end{cases} \tag{I-86}$$

H is thus associated with time translations

• Pure Galilean transformation

Consider (figure 4) two Galilean reference frames $Oxyz$ and $O'x'y'zy$ with parallel axes, and such that the vector $\overrightarrow{OO'}$ linking their origins is equal to $-t\delta\boldsymbol{v}$ ($\delta\boldsymbol{v}$ is an infinitesimal constant vector). If the position vector of a given particle is $\boldsymbol{r}(t) = \overrightarrow{OM}(t)$ in the first reference frame, it will become $\boldsymbol{r}'(t) = \overrightarrow{O'M}(t)$ in the second (passive point of view, § A-3), with:

$$\boldsymbol{r}' = \boldsymbol{r} + t\delta\boldsymbol{v} \tag{I-87}$$

In a similar way:

$$\boldsymbol{p}' = \boldsymbol{p} + m\delta\boldsymbol{v} \tag{I-88}$$

As in (I-52), we introduce the vector:

$$\boldsymbol{G}_0(t) = t\boldsymbol{p} - m\boldsymbol{r} \tag{I-89}$$

We get:

$$\begin{cases} \{x\ ,\ \boldsymbol{G}_0 \cdot \delta\boldsymbol{v}\} = t\delta v_x \\ \{p_x\ ,\ \boldsymbol{G}_0 \cdot \delta\boldsymbol{v}\} = m\delta v_x \end{cases} \tag{I-90}$$

and hence:

$$\begin{aligned} \boldsymbol{r}' &= \boldsymbol{r} - \{\delta\boldsymbol{v} \cdot \boldsymbol{G}_0, \boldsymbol{r}\} \\ \boldsymbol{p}' &= \boldsymbol{p} - \{\delta\boldsymbol{v} \cdot \boldsymbol{G}_0, \boldsymbol{p}\} \end{aligned} \tag{I-91}$$

Using Poisson brackets, the vector $\boldsymbol{G}_0$ generates infinitesimal Galilean transformations. For an ensemble of several particles, we can reason in a similar way by introducing the position of the system's center of mass – *cf.* (I-52).

B-3-d. Symmetry transformations and constants of motion

Consider an infinitesimal transformation generated by Poisson brackets containing a function $G(t)$:

$$q_i' = q_i - \{G(t),\, q_i\}\,\delta\varepsilon \qquad p_j' = p_j - \{G(t),\, p_j\}\,\delta\varepsilon \tag{I-92}$$

This transformation remains the same if one adds to $G(t)$ any function of time t, leading to a whole family of equivalent functions $G'(t)$, associated with the same transformation. We now show that:

> The transformation generated by Poisson brackets containing a function $G(t)$ is a symmetry transformation provided either $G(t)$, or one of the equivalent functions $G'(t)$, is a constant of motion.

As we will see in § C-2, this result has a direct analogue in quantum mechanics, where Poisson brackets are replaced with commutators between operators.

To prove this result, we come back to the scheme of figure 1, and take as an example the transformation associated with changing the system's description from reference frame $Oxyzt$ to another frame $O'x'y'z't'$. In a first case, the dynamical variables of the physical system are changed at an initial time t_0 by the transformation associated with $G(t_0)$. The system then evolves during an infinitesimal time δt under the effect of its Hamiltonian H. In the second case, we first let the system evolve during the same time δt, and then change the reference frame at time $t_0 + \delta t$. Since the relative position of the two reference frames may have changed during δt, a different transformation has to be applied, associated with $G(t_0 + \delta t)$. If the same final physical state is obtained in both cases, the function $G(t)$ is associated with a symmetry transformation.

In the first case, the successive variations of the dynamical variables are:

$$\begin{aligned}
q_i &\Rightarrow q_i - \{G(t_0),\, q_i\}\, \delta\varepsilon \\
&\Rightarrow q_i - \{G(t_0), q_i\}\, \delta\varepsilon - \{H,\, q_i\}\, \delta t + \{H, \{G(t_0),\, q_i\}\}\, \delta t \delta\varepsilon + \dots \\
p_j &\Rightarrow p_j - \{G(t_0),\, p_j\}\, \delta\varepsilon \\
&\Rightarrow p_j - \{G(t_0), p_j\}\, \delta\varepsilon - \{H,\, p_j\}\, \delta t + \{H, \{G(t_0),\, p_j\}\}\, \delta t \delta\varepsilon + \dots
\end{aligned} \tag{I-93a}$$

while, in the second case, they become:

$$\begin{aligned}
q_i &\Rightarrow q_i - \{H,\, q_i\}\, \delta t \\
&\Rightarrow q_i - \{H,\, q_i\}\, \delta t - \{G(t_0), q_i\}\delta\varepsilon + \Big[- \{\frac{\partial G}{\partial t}, q_i\} + \{G(t_0), \{H,\, q_i\}\}\Big]\, \delta t \delta\varepsilon + \dots \\
p_j &\Rightarrow p_j - \{H,\, p_j\}\, \delta t \\
&\Rightarrow p_j - \{H,\, p_j\}\, \delta t - \{G(t_0), p_j\}\delta\varepsilon + \Big[- \{\frac{\partial G}{\partial t}, p_i\} + \{G(t_0), \{H,\, p_j\}\}\Big]\, \delta t \delta\varepsilon + \dots
\end{aligned} \tag{I-93b}$$

The transformation is a symmetry transformation if, at time t_0:

$$\{H, \{G(t),\, q\}\} + \{\frac{\partial G}{\partial t}, q\} + \{G(t), \{q,\, H\}\} = 0 \tag{I-94}$$

where q is any of the dynamical variables, position or momentum. The Jacobi identity (I-75) then allows us to write this condition as:

$$-\{q, \{H,\, G(t)\}\} + \{q, \frac{\partial G}{\partial t}\} = 0 \tag{I-95}$$

It follows that the function $\{H, G(t)\} - \partial G / \partial t$ has a zero Poisson bracket with all the dynamical variables, which means that this function is independent of all of them. It is therefore a function of time only:

$$-\{H,\, G(t)\}\} + \frac{\partial G}{\partial t} = F(t) \tag{I-96}$$

The left-hand side of this equation is the total time derivative $\mathrm{d}G/\mathrm{d}t$ of G – *cf.* (I-74). We therefore have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[G(t) - \int \mathrm{d}t\; F(t)\Big] = 0 \tag{I-97}$$

The function between brackets, which is equivalent to $G(t)$ since it defines the same transformation when inserted in (I-92), is a constant of motion.

C. Symmetries in quantum mechanics

In quantum mechanics, symmetry transformations play at least as big a role as in classical mechanics, as will be shown below.

C-1. Quantization standard procedure

In quantum mechanics, the set of q_i and p_i are replaced by a ("ket") vector $|\psi\rangle$ belonging to the state space $\mathscr{E}$ of the system. For a spinless particle, one can choose one of two possible "bases" , either the basis $\{|\boldsymbol{r}\rangle\}$ of the eigenvectors of the position operator $\boldsymbol{R}$, or the basis $\{|\boldsymbol{p}\rangle\}$ of the eigenvectors of the momentum $\boldsymbol{P}$. To each of these bases correspond the functions:

$$\psi(\boldsymbol{r}) = \langle\boldsymbol{r}|\psi\rangle \tag{I-98a}$$

$$\overline{\psi}(\boldsymbol{p}) = \langle\boldsymbol{p}|\psi\rangle \tag{I-98b}$$

called the wave functions in the $\boldsymbol{r}$ and $\boldsymbol{p}$ representations respectively. For a particle with non-zero spin, several wave functions are necessary to characterize a quantum state $|\psi\rangle$.

The classical quantities $\mathscr{A}(q_i, p_i)$ become linear Hermitian operators acting in $\mathscr{E}$:

$$\mathscr{A}(q_i, p_i) \Longrightarrow A \tag{I-99}$$

In general, any ket $|\psi\rangle$ may be expanded on eigenvectors of A; A is then called an "observable". To the classical Poisson brackets correspond commutators between observables. Relations (I-72) become:

$$[R_i, P_j] = i\hbar\,\delta_{ij} \tag{I-100}$$

(where i and j stand for x, y or z).

Comment:

Quantization rules are often defined by a replacement of classical functions $\mathscr{A}$ by operators A, obtained by simply substituting an operator Q for every variable q, an operator P for every variable p, and finally performing a symmetrization of the products of operators to ensure the hermiticity of A. The operator value of a commutator is obtained in the same way from the Poisson bracket (Dirac rule [37]). This method is nevertheless rather ill-defined over the ensemble of classical quantities, and may even sometimes lead to contradictory results [38, 39]. Moreover, if one changes the classical variables in the Lagrangian (for instance, in order to use angle and distance variables), the method does not provide a unique and well-defined quantization. In this book, we proceed differently, and start only from general considerations on space–time and symmetries.

The time evolution of $|\psi\rangle$ is the Schrödinger equation:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H(t)\, |\psi(t)\rangle \tag{I-101}$$

where the Hamiltonian $H(t)$ is the quantum observable associated with the classical Hamiltonian function.

The equation is equivalent to:

$$|\psi(t)\rangle = U(t, t_0)\; |\psi(t_0)\rangle \tag{I-102}$$

In this equation, $|\psi(t)\rangle$ represents the state vector at time t, $|\psi(t_0)\rangle$ its value at the initial time t_0, and $U(t, t_0)$ is the unitary operator obeying:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} U(t, t_0) = H(t)\, U(t, t_0) \tag{I-103a}$$

with:

$$U(t_0, t_0) = \mathbb{1} \tag{I-103b}$$

When H is time-independent, we simply have:

$$U(t, t_0) = \exp\left\{-\frac{i}{\hbar} H \times (t - t_0)\right\} \tag{I-104}$$

The operator $U(t)$ is called an "evolution operator". More details on its properties can be found for instance in complement $\mathrm{F}_{\mathrm{III}}$ and in § A of chapter XX of [9].

C-2. Symmetry transformations

Consider a transformation $\mathcal{T}$ of the physical system. Before the transformation, the system is described at time t by a ket $|\psi(t)\rangle$, and after the transformation by a ket $|\psi'(t)\rangle$. We define the operator T (acting in the state space $\mathscr{E}$) that transforms $|\psi\rangle$ into $|\psi'\rangle$:

$$|\psi'(t)\rangle = T(t)\, |\psi(t)\rangle \tag{I-105}$$

(in many cases, $T(t)$ is a linear unitary operator). The diagram of figure 1 now becomes that of figure 5. $\mathcal{T}$ will be a symmetry transformation if, at any time t, we have:

$$U(t, t_0)T(t_0)|\psi(t_0)\rangle = T(t)|\psi(t)\rangle = T(t)U(t, t_0)|\psi(t_0)\rangle \tag{I-106}$$

meaning, since this is true for any $|\psi(t_0)\rangle$:

$$\boxed{U(t, t_0)\, T(t_0) = T(t)\, U(t, t_0)} \tag{I-107}$$

$$
\begin{tikzcd}
|\psi(t_0)\rangle \arrow[r, Rightarrow, "T(t_0)"] \arrow[d] & |\psi'(t_0)\rangle = T(t_0)\,|\psi(t_0)\rangle \arrow[d] \\
|\psi(t)\rangle = U(t,t_0)\,|\psi(t_0)\rangle \arrow[r, Rightarrow, "T(t)\ ?"] & U(t,t_0)\ \left[T(t_0)\,|\psi(t_0)\rangle\right]
\end{tikzcd}
$$

Figure 5: Diagram expliciting the condition for the transformation $\mathcal{T}$, represented in quantum mechanics by the operator $T(t)$ acting in the state space, to be a symmetry transformation. Arrows pointing down represent the time propagation due to the evolution operator $U(t,t_0)$, the double horizontal arrows the effect of $T(t)$. The transformation $\mathcal{T}$ is a symmetry transformation if $T(t)$ transforms a possible evolution of the system into another possible evolution. This is indeed the case if one can "close the square" with the horizontal arrow at the bottom of the figure, just below the question mark.

If this condition is met, we can write:

$$U^\dagger(t,t_0)\,T(t)\,U(t,t_0) = T(t_0) \tag{I-108}$$

This equality means that the operator $T(t)$ corresponds to a time-independent operator in Heisenberg's point of view. T is thus a constant of motion.

If T is time-independent, (I-107) becomes:

$$\boxed{[T\,,\,U(t,t_0] = 0} \tag{I-109}$$

which shows that, at any time, T and the evolution operator U commute. If, in addition, H is also time-independent (conservative system), U is given by (I-104), and (I-109) is written as:

$$\boxed{[T,H] = 0} \tag{I-110}$$

This shows again that T is a constant of motion.

Comment:

Two kets that differ by a global phase factor correspond to identical physical states. Relation (I-107) may thus be replaced by the more general equality:

$$U(t,t_0)\,T(t_0)|\psi(t_0)\rangle = \mathrm{e}^{\mathrm{i}\alpha(t)}T(t)\,U\,(t,t_0)\,|\psi(t_0)\rangle \qquad \forall\ |\psi(t_0)\rangle \tag{I-111a}$$

where $\alpha(t)$ is a function of time, which could, a priori, also depend on $|\psi(t_0)\rangle$. This is actually not the case as we now show. Setting:

$$T(t)\,U(t,t_0)|\psi(t_0)\rangle = |\chi\rangle \tag{I-111b}$$

relation (I-111a) can be written (operator T can be inverted):

$$U(t,t_0)\,T(t_0)\,U^{-1}\,(t_0,t)\ T^{-1}(t)\ |\chi\rangle = \mathrm{e}^{\mathrm{i}\alpha(t)}\,|\chi\rangle \tag{I-111c}$$

where α could be a function of $|\chi\rangle$. But the operator acting on the left of this equality is a linear operator having as eigenvector any ket $|\chi\rangle$, which means it corresponds to a diagonal matrix in any base; this is possible only if all its eigenvalues are equal (if two of them were different, one can easily see that the sum of the two corresponding eigenvectors would not be an eigenvector of the operator). As a result, α is necessarily independent of the ket.

Consequently we have:

$$U(t,t_0)\,T(t_0) = \mathrm{e}^{\mathrm{i}\alpha(t)}\,T(t)\,U(t,t_0) \tag{I-111d}$$

where $\alpha(t)$ is a real function of time, equal to zero for $t = t_0$. Setting:

$$T'(t) = \mathrm{e}^{\mathrm{i}\alpha(t)}\,T(t) \tag{I-112}$$

we find for T' an equality of the type (I-107). We shall assume that T' is defined [8] in such a way that $\alpha(t) = 0$.

C-3. General consequences

Relations (I-107), (I-109), or (I-110) imply that, if $\mathcal{T}$ is a symmetry transformation, operators U and H cannot take an arbitrary form.

For example, we cannot have a Hamiltonian proportional to $\boldsymbol{L}\cdot\boldsymbol{R}$ (with $\boldsymbol{L} = \boldsymbol{R}\times\boldsymbol{P}$) if the geometrical parity operation is a symmetry transformation. Similarly, a time-reversal symmetry does not allow a Hamiltonian proportional to $\boldsymbol{R}\cdot\boldsymbol{P}$, etc.

We already noted that (I-108) or (I-110) implies the existence of *constants of motion*. For example, (I-110) shows that the average values of T, T^2, T^3, etc. are constant. Hence, wherever they appear, all the T operators, their products, etc. yield constants of motion. For *conservative* systems, the existence of such constants simplifies the *search for stationary states* (eigenkets of H): one looks for a basis of eigenkets common to H and T.

If H and T commute, we have:

$$\langle\theta_1|H|\theta_2\rangle = 0 \quad \text{if} \quad \theta_1 \neq \theta_2 \tag{I-113}$$

[8] We are discussing here a single transformation $\mathcal{T}$. To eliminate the phase factors for a group of transformations is more delicate, and will be discussed later, in particular in chapter IV.

where $|\theta_1\rangle$ and $|\theta_2\rangle$ are eigenvectors of T, with eigenvalues θ_1 and θ_2. This type of relation yields "*selection rules*" for H; these rules are actually valid for all the *invariant observables* T.

One can also obtain some information concerning the *degeneracy*. Consider for example $|E_0\rangle$, an eigenket of H, with eigenvalue E_0. Relation (I-110) shows that if:

$$|E_0'\rangle = T|E_0\rangle \tag{I-114}$$

$|E_0'\rangle$ is an eigenket of H with the same eigenvalue E_0, since:

$$\begin{aligned} H\,|E_0'\rangle &= H\,T\,|E_0\rangle = T\,H\,|E_0\rangle = E_0\,T\,|E_0\rangle \\ &= E_0\,|E_0'\rangle \end{aligned} \tag{I-115}$$

Unless $|E_0\rangle$ and $|E_0'\rangle$ are proportional, the eigenvaleue E_0 is at least doubly degenerate.

If both T_1 and T_2 obey (I-108), so does their product. The operators T associated with the symmetry transformations form a group. The T do not necessarily commute with each other (non-commutative group), and we shall see the importance of the commutation relations between these operators.

Complement A_I

Eulerian and Lagrangian points of view in classical mechanics

The purpose of this complement is to present two different points of views used in classical statistical mechanics to describe an ensemble of physical systems moving in the phase space. In fluid mechanics, Eulerian or Lagrangian points of view can be used. In the first (Eulerian), one focuses on the time evolution of the fluid flow at each fixed position in space, computing for example the fluid flow rate at every point in space and as a function of time. In other words, one studies successive "snapshots" of the fluid. In the second point of view (Lagrangian), one follows the motion of each fluid element along its own trajectory. To study the motion of these fluid elements, one must define total time derivatives along each of these trajectories.

Extending these ideas from ordinary space to a larger space, that we call "phase space", we now use Eulerian and Lagrangian points of view to describe in classical statistics an ensemble of physical systems. These two points of view present a certain analogy with those of Schrödinger and Heisenberg in quantum mechanics, respectively. Just as Heisenberg's point of view, that of Lagrange is better adapted for the computation of two-time average values (correlation functions).

Consider an ensemble of physical systems, each being described, in the Hamiltonian formalism, by N coordinates q_i and N conjugated momenta p_i. Each dynamical state is defined by the coordinates q_i and p_i of a point that belongs to a $2N$ dimensional space, the so-called "phase space" (in the case of a single particle, the phase space has 6 dimensions). This point's motion obeys Hamilton's equations:

$$\begin{cases} \dot{q}_i(t) = \dfrac{\partial}{\partial p_i} H\left[q_j(t),\, p_j(t)\,;\, t\right] \\ \dot{p}_i(t) = -\dfrac{\partial}{\partial q_i} H\left[q_j(t),\, p_j(t)\,;\, t\right] \end{cases} \tag{1}$$

where $H[q_i, p_i\,;\, t]$ is the Hamiltonian of each system, supposed to be the same for all of them.

1. Eulerian point of view

Consider now a statistical ensemble of such systems, define at time $t = t_0$ by a probability density $\rho(q_i, p_i, t_0)$ in the phase space. A time t, this ensemble is described by a probability density $\rho(q_i, p_i, t)$ that yields the probability $\mathrm{d}\mathcal{P}(t)$:

$$\mathrm{d}\mathcal{P}(t) = \rho\,(q_1, q_2 \ldots q_N, p_1, p_2 \ldots p_N\,;\,t) \times \mathrm{d}q_1 \ldots \mathrm{d}q_N\ \mathrm{d}p_1 \ldots \mathrm{d}p_N \tag{2}$$

for the point characterizing the system to be in the volume element:

$$\mathrm{d}q_1\ \mathrm{d}q_2 \ldots \mathrm{d}q_N\ \ \mathrm{d}p_1\ \mathrm{d}p_2 \ldots \mathrm{d}p_N \tag{3}$$

centered at:

$$q_1, q_2, \ldots, q_N,\ \ p_1, p_2 \ldots p_N \tag{4}$$

We obviously have:

$$\int \mathrm{d}^N q \int \mathrm{d}^N p\ \rho\,(q_1, \ldots, q_N,\ p_1, \ldots, p_N\,;\,t) = 1 \tag{5}$$

where $\int \mathrm{d}^N q$ and $\int \mathrm{d}^N p$ are simplified notations for the pluri-dimensional integrals $\int \mathrm{d}q_1 \int \mathrm{d}q_2 \ldots \int \mathrm{d}q_N$ and $\int \mathrm{d}p_1 \int \mathrm{d}p_2 \ldots \int \mathrm{d}p_N$, respectively. We shall also simplify the notation for the whole set of coordinates (4) into $q_1, ..., p_N$. Remember that the system's evolution is fully determined (the function H is not random), so that probabilities only come into play in the choice of the system's initial state.

In the $2N$-dimensional phase space, the probability density ρ can be assimilated to the density of a fluid; the associated flow density will then be a $2N$-component vector:

$$\boldsymbol{J}(q_1, ..., p_N) \begin{cases} \rho\,(q_j, p_j\,;\,t)\ \dot{q}_i = \rho\,(q_j, p_j\,;\,t)\ \dfrac{\partial}{\partial p_i} H\,(q_j, p_j\,;\,t) \\ \rho\,(q_j, p_j\,;\,t)\ \dot{p}_i = -\rho\,(q_j, p_j\,;\,t)\ \dfrac{\partial}{\partial q_i} H\,(q_j, p_j\,;\,t) \end{cases} \tag{6}$$

where the two i-ranking components of that vector are written on the right-hand side of the bracket. This allows computing the number of systems of the ensemble that enter or exit each volume of the phase space, hence obtaining the local variation $\partial\rho/\partial t$ of the density ρ. The divergence theorem shows:

$$-\frac{\partial \rho}{\partial t} = {}^{2N}\boldsymbol{\nabla} \cdot \boldsymbol{J} \tag{7}$$

where ${}^{2N}\boldsymbol{\nabla} \cdot \boldsymbol{J}$ symbolizes the divergence of the vector $\boldsymbol{J}$ in the phase space:

$$\begin{aligned} {}^{2N}\boldsymbol{\nabla} \cdot \boldsymbol{J} = \sum_i \Bigg\{ &\frac{\partial}{\partial q_i} \left[\rho\,(q_j, p_j\,;\,t)\ \frac{\partial}{\partial p_i} H\,(q_j, p_j\,;\,t)\right] \\ &- \frac{\partial}{\partial p_i} \left[\rho\,(q_j, p_j\,;\,t)\ \frac{\partial}{\partial q_i} H\,(q_j, p_j\,;\,t)\right]\Bigg\} \end{aligned} \tag{8}$$

The $\partial^2 H/\partial q_i\,\partial p_i$ terms cancel out from this expression and, using definition (I-71) for the Poisson brackets, we have:

$$^{2N}\nabla\cdot\boldsymbol{J} = \Big\{\rho\,(q_j,p_j\,;\,t)\,,H\,(q_j,p_j\,;\,t)\Big\} \tag{9}$$

so that relation (7) becomes:

$$\frac{\partial\rho}{\partial t} = \{H\,(q_j,p_j\,;\,t)\,,\rho\,(q_j,p_j\,;\,t)\,\} \tag{10}$$

Consider now a quantity $\mathscr{A}$ that depends on the q_i and p_i, and possibly on time:

$$\mathscr{A}\,(q_i,p_i\,;\,t) \tag{11}$$

$\mathscr{A}$ could be, for example, the kinetic energy of an ensemble of particles, their interaction potential energy, etc. The average value of $\mathscr{A}$ at any given time t is written:

$$\overline{\mathscr{A}}(t) = \int \mathrm{d}^N q \int \mathrm{d}^N p\;\rho\,(q_i,p_i\,;\,t)\;\;\mathscr{A}\,(q_i,p_i\,;\,t) \tag{12}$$

Using (10), we can compute the time evolution of this average value:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\overline{\mathscr{A}}(t) = \int \mathrm{d}^N q \int \mathrm{d}^N p\;\left[\{H,\rho\}\,\mathscr{A} + \rho\frac{\partial}{\partial t}\mathscr{A}\right] \tag{13}$$

Slightly anticipating on the Lagrangian point of view, we can also compute the total time derivative of ρ. We "go along" with the particles and measure at each instant their "density" in the phase space:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\,\rho &= \frac{\partial}{\partial t}\,\rho + \sum_i\left(\dot q_i\,\frac{\partial\rho}{\partial q_i} + \dot p_i\,\frac{\partial\rho}{\partial p_i}\right) = \frac{\partial}{\partial t}\,\rho + \sum_i\left(\frac{\partial H}{\partial p_i}\frac{\partial\rho}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial\rho}{\partial p_i}\right)\\ &= \frac{\partial}{\partial t}\,\rho + \{\rho, H(t)\}\end{aligned} \tag{14}$$

Since $\{\rho,H\} = -\{H,\rho\}$, we are left with:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\rho = 0 \tag{15}$$

This result is called Liouville's theorem. It states that the "fluid" particles flow along their trajectories keeping their "density" constant (which can vary at a fixed point of the phase space). They occupy, as time goes by, a constant volume of the phase space.

2. Lagrangian point of view

In the point of view of § 1 above, the integrals are performed with probability distributions that are time-dependent. It is sometimes easier, in particular when calculating two-time average values, to compute integrals with a sole distribution, the one describing the system at time $t = t_0$, chosen as the initial time. In the Lagrangian point of view, one works with a single probability density ρ^L in the phase space, which is time-independent and equal to:

$$\rho^L\left(q_i^0, p_i^0\right) = \rho\left(q_i = q_i^0, p_i = p_i^0\,;\, t = t_0\right) \tag{16}$$

where ρ is the function already defined in § 1.

This is reminiscent of the Heisenberg point of view in quantum mechanics, where the state vector $|\psi_H\rangle$ is constant, and it is the operators A_H that show the Hamiltonian time dependence. In a similar way, we define a classical function $\mathscr{A}^L(q_i^0, p_i^0\,;\, t)$ that shows the time dependence induced by the Hamiltonian. At $t = t_0$, this function has the same value as in the Eulerian point of view. At time t, it yields by definition the contribution to the quantity $\overline{\mathscr{A}}$ from all the points in the phase space that were at time t_0 at q_i^0 and p_i^0. In other words, if the solutions of equations (1) with the initial conditions $q_j(t = t_0) = q_j^0$ and $p_j(t = t_0) = p_j^0$ are written as:

$$\begin{cases} q_j\left(q_i^0, p_i^0\,;\, t\right) \\ p_j\left(q_i^0, p_i^0\,;\, t\right) \end{cases} \tag{17}$$

the function $\mathscr{A}^L\left(q_i^0, p_i^0\,;\, t\right)$ is defined by:

$$\mathscr{A}^L\left(q_i^0, p_i^0\,;\, t\right) = \mathscr{A}\left[q_i\left(q_i^0, p_i^0\,;\, t\right),\; p_i\left(q_i^0, p_i^0\,;\, t\right)\,;\, t\right] \tag{18}$$

Using the conservation of the probability density (15), we can write:

$$\overline{\mathscr{A}}(t) = \int \mathrm{d}^N q^0 \int \mathrm{d}^N p^0\; \rho^L\left(q_i^0, p_i^0\right)\; \mathscr{A}^L\left(q_i^0, p_i^0\,;\, t\right) \tag{19}$$

We see that $\mathscr{A}^L$ is explicitly time-dependent, even if this is not the case for $\mathscr{A}$, since according to (18) and (1):

$$\begin{aligned}\frac{\partial}{\partial t}\mathscr{A}^L\left(q_i^0, p_i^0\,;\, t\right) &= \sum_i\left\{\frac{\partial\mathscr{A}}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial\mathscr{A}}{\partial p_i}\frac{\partial H}{\partial q_i}\right\} + \frac{\partial\mathscr{A}}{\partial t} \\ &= \left\{\mathscr{A}(t), H(t)\right\} + \frac{\partial\mathscr{A}}{\partial t}\end{aligned} \tag{20}$$

In the right-hand side of the equality, the q_i^0 and p_i^0 dependence is explicit when replacing the q_j and p_j by their expressions (17).

A two-time average value is easy to compute:

$$\overline{\mathscr{A}(t)\mathscr{B}(t')} = \int \mathrm{d}^N q^0 \int \mathrm{d}^N p^0\; \rho^L\left(q_i^0, p_i^0\right) \times \mathscr{A}^L\left(q_i^0, p_i^0\,;\, t\right)\mathscr{B}^L\left(q_i^0, p_i^0\,;\, t'\right) \tag{21}$$

Comment:

Relation (20) is a "mixed" equality as it yields the time variation of $\mathscr{A}^L$ as a function of a Poisson bracket involving the variables q_i and p_i, introduced in the Eulerian point of view. It is also possible to only use, in the Lagrangian point of view, the variables q_i^0 and p_i^0, and introduce the corresponding Poisson brackets:

$$\left\{\mathscr{A}^L(t), \mathscr{B}^L(t)\right\}_L = \sum_n \left(\frac{\partial \mathscr{A}^L}{\partial q_n^0} \frac{\partial \mathscr{B}^L}{\partial p_n^0} - \frac{\partial \mathscr{A}^L}{\partial p_n^0} \frac{\partial \mathscr{B}^L}{\partial q_n^0} \right) \tag{22}$$

Equality (20) can then be written, as will be shown below:

$$\begin{aligned}\frac{\partial}{\partial t}\mathscr{A}^L\left(q_i^0, p_i^0\,;\,t\right) = &\left\{\mathscr{A}^L(t), H^L(t)\right\}_L \\ &+ \frac{\partial \mathscr{A}}{\partial t}\left[q_j\left(q_i^0, p_i^0\,;\,t\right),\ p_j\left(q_i^0, p_i^0\,;\,t\right)\,;\,t\right]\end{aligned} \tag{23a}$$

or, if $\mathscr{A}$ (in the Eulerian point of view) is time-independent:

$$\frac{\partial}{\partial t}\mathscr{A}^L\left(q_i^0, p_i^0\,;\,t\right) = \left\{\mathscr{A}^L(t), H^L(t)\right\}_L \tag{23b}$$

Demonstration: Let us first compute the simplest Poisson brackets:

$$\{q_i(t), q_j(t)\}_L \qquad \{p_i(t), p_j(t)\}_L \qquad \{q_i(t), p_j(t)\}_L \tag{24}$$

At time $t = t_0$, they are the usual Poisson brackets. At a later time t, one must bear in mind that the $q_i(t)$ and $p_i(t)$ a priori depend on the q_k^0 and p_k^0 for all values of k. After an infinitesimal time dt, we have:

$$\begin{cases} q_i(t_0 + \mathrm{d}t) = q_i^0 + \dfrac{\partial H}{\partial p_i}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \\[2ex] p_j(t_0 + \mathrm{d}t) = p_j^0 - \dfrac{\partial H}{\partial q_j}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \end{cases} \tag{25}$$

In these relations, we can replace the partial derivatives of H with respect to q_i and p_i by those with respect to q_i^0 and p_i^0. A derivation then provides:

$$\begin{aligned}\frac{\partial q_i(t_0+\mathrm{d}t)}{\partial q_l^0} &= \delta_{il} + \frac{\partial^2 H}{\partial q_l^0 \partial p_i^0}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \\ \frac{\partial q_i(t_0+\mathrm{d}t)}{\partial p_l^0} &= \qquad \frac{\partial^2 H}{\partial p_l^0 \partial p_i^0}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \\ \frac{\partial p_j(t_0+\mathrm{d}t)}{\partial q_l^0} &= \quad\ -\frac{\partial^2 H}{\partial q_l^0 \partial q_j^0}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \\ \frac{\partial p_j(t_0+\mathrm{d}t)}{\partial p_l^0} &= \delta_{jl} - \frac{\partial^2 H}{\partial p_l^0 \partial q_j^0}(q_k^0, p_k^0, t_0)\ \mathrm{d}t \end{aligned} \tag{26}$$

We take the product of the first and fourth of these equalities, minus the product of the second and third, and sum over the index l. We then obtain, to first-order in dt:

$$\{q_i(t_0+\mathrm{d}t), p_j(t_0+\mathrm{d}t)\}_L = \delta_{ij} - \frac{\partial^2 H}{\partial p_i^0 \partial q_j^0}\mathrm{d}t + \frac{\partial^2 H}{\partial q_j^0\,\partial p_i^0}\mathrm{d}t = \delta_{ij} \tag{27a}$$

As for the first two equalities (26), they lead to:

$$\{q_i(t_0+\mathrm{d}t), q_j(t_0+\mathrm{d}t)\}_L = \frac{\partial^2 H}{\partial p_i^0 \partial p_j^0}\mathrm{d}t - \frac{\partial^2 H}{\partial p_j^0 \partial p_i^0}\mathrm{d}t = 0 \tag{27b}$$

A similar calculation starting from the two last lines of (26) shows that the Poisson-Lagrange brackets $\{p_i(t_0+\mathrm{d}t), p_j(t_0+\mathrm{d}t)\}_L$ also vanish. The three brackets (24) therefore take the same value at times t_0 and $t_0+\mathrm{d}t$.

In the above equations, we can replace t_0 by t_0+t', q_i^0 and p_i^0 by $q_i(t_0+t')$ and $p_i(t_0+t')$, and see by the same reasoning that the Poisson–Lagrange brackets are still stationary at any time. These brackets are therefore constant, and keep the value of the usual Poisson brackets:

$$\{q_i(t), p_j(t)\}_L = \delta_{ij} \tag{28a}$$

$$\{q_i(t), q_j(t)\}_L = 0 \tag{28b}$$

$$\{p_i(t), p_j(t)\}_L = 0 \tag{28c}$$

We can now evaluate the quantity $\{\mathscr{A}^L(t), \mathscr{B}^L(t)\}_L$ appearing in (22). According to definition (18) and the chain rule for taking the derivative of the composition of two functions, we get:

$$\begin{aligned}
\left\{\mathscr{A}^L(t), \mathscr{B}^L(t)\right\}_L = \sum_{ijn} \Bigg[& \left[\frac{\partial \mathscr{A}}{\partial q_i}\frac{\partial q_i}{\partial q_n^0} + \frac{\partial \mathscr{A}}{\partial p_i}\frac{\partial p_i}{\partial q_n^0}\right] \left[\frac{\partial \mathscr{B}}{\partial q_j}\frac{\partial q_j}{\partial p_n^0} + \frac{\partial \mathscr{B}}{\partial p_j}\frac{\partial p_j}{\partial p_n^0}\right] \\
& - \left[\frac{\partial \mathscr{A}}{\partial q_i}\frac{\partial q_i}{\partial p_n^0} + \frac{\partial \mathscr{A}}{\partial p_i}\frac{\partial p_i}{\partial p_n^0}\right] \left[\frac{\partial \mathscr{B}}{\partial q_j}\frac{\partial q_j}{\partial q_n^0} + \frac{\partial \mathscr{B}}{\partial p_j}\frac{\partial p_j}{\partial q_n^0}\right]\Bigg] \\
= \sum_{ij} \Bigg[& \frac{\partial \mathscr{A}}{\partial q_i}\frac{\partial \mathscr{B}}{\partial q_j}\{q_i(t),\, q_j(t)\}_L + \frac{\partial \mathscr{A}}{\partial p_i}\frac{\partial B}{\partial q_j}\{p_i(t),\, q_j(t)\}_L \\
& + \frac{\partial \mathscr{A}}{\partial q_i}\frac{\partial \mathscr{B}}{\partial p_j}\{q_i(t),\, p_j(t)\}_L + \frac{\partial \mathscr{A}}{\partial p_i}\frac{\partial \mathscr{B}}{\partial p_j}\{p_i(t),\, p_j(t)\}_L\Bigg]
\end{aligned} \tag{29}$$

Using the general relations (28), we obtain:

$$\begin{aligned}
\left\{\mathscr{A}^L(t), \mathscr{B}^L(t)\right\}_L &= \sum_i \left[\frac{\partial \mathscr{A}}{\partial q_i}\,\frac{\partial \mathscr{B}}{\partial p_i} - \frac{\partial \mathscr{A}}{\partial p_i}\,\frac{\partial \mathscr{B}}{\partial q_i}\right] \\
&= \{\mathscr{A}, \mathscr{B}\}
\end{aligned} \tag{30}$$

where, on the right-hand side, the q_j and p_j must be replaced by their expressions (17) as a function of the q_n^0 and p_n^0.

A simple example: One-dimensional particle subjected to a force $mg(t)$:

- Eulerian point of view:

$$H(t) = \frac{p^2}{2m} - mg(t)x$$
$$\frac{\mathrm{d}}{\mathrm{d}t}x = \frac{p}{m} \qquad \frac{\mathrm{d}}{\mathrm{d}t}p = mg(t)$$
$$\begin{cases} p(t) = m\displaystyle\int_0^t \mathrm{d}t'\ g(t') + p_0 \\ x(t) = \displaystyle\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t''\ g(t'') + \frac{p_0}{m}t + x_0 \end{cases}$$

- Lagrangian point of view:

$$\begin{aligned} H^L(x_0, p_0\,;\,t) &= \frac{1}{2m}\left[p_0 + m\int_0^t \mathrm{d}t'\ g(t')\right]^2 \\ &\quad - mg(t)\left[\int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t''\ g(t'') + \frac{p_0}{m}t + x_0\right] \\ &= \frac{p_0^2}{2m} - mg(t)\,x_0 + p_0\left[\int_0^t \mathrm{d}t'\ g(t') - t\,g(t)\right] \\ &\qquad + F(t) \end{aligned}$$

where :

$$F(t) = m\left[\frac{1}{2}\left(\int_0^t \mathrm{d}t'\ g(t')\right)^2 - \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t''\ g(t)\ g(t'')\right]$$

In the Lagrangian point of view, the particle's momentum and position are:

$$p^L(p_0\,;\,t) = p_0 + m\int_0^t \mathrm{d}t'\ g(t')$$
$$x^L(x_0, p_0\,;\,t) = x_0 + \frac{t}{m}p_0 + \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t''\ g(t'')$$

It's easy to show that $\left\{x^L(t), p^L(t)\right\}_L = 1$, and we can compute:

$$\begin{aligned} \left\{x^L(t), H^L(t)\right\}_L &= \frac{p_0}{m} + \int_0^t \mathrm{d}t'\ g(t') - t\ g(t) - \left(\frac{t}{m}\right)(-mg(t)) \\ &= \frac{p_0}{m} + \int_0^t \mathrm{d}t'\ g(t') = \frac{p(t)}{m} \end{aligned}$$
$$\left\{p^L(t), H^L(t)\right\}_L = m\ g(t) = -\frac{\partial H(t)}{\partial x}$$

which shows that equality (30) is satisfied.

Complement B_I

Noether's theorem for a classical field

Noether's theorem was introduced in § B-2-b of chapter I for the case where the physical system is described by dynamical variables q_i whose index i takes a finite number of discrete values $i = 1, 2, \ldots, N$. Actually, the main interest of the theorem occurs when the system under study is a classical or quantum field. As a field varies in space, the $q_i(t)$ are replaced by the components $\phi_j(\boldsymbol{r}, t)$ of the field, where $\boldsymbol{r}$ labels the dynamical variables: this continuous index replaces the discrete index i of the previous case. The discrete index j labels the components of the field: a scalar field has only one component (and the index j is not needed): a vector field in 3-dimensional space has three components, $j = 1, 2, 3$; one can also define a second-order tensor field for which j takes on 9 values, etc.

For the sake of simplicity, we assume in this complement that the field components are real; the case of complex components will be studied at the beginning of complement C_{VI}.

1. Lagrangian density and Lagrange equations for continuous variables

The Lagrangian L depends on all the dynamical variables of the physical system, and is written:

$$L = \int \mathrm{d}^3 r \; \mathcal{L}(\boldsymbol{r}, t) \tag{1}$$

where $\mathcal{L}(\boldsymbol{r}, t)$ is the "Lagrangian density":

$$\mathcal{L} = \mathcal{L}\Big(\phi_j(\boldsymbol{r}, t), \dot{\phi}_j(\boldsymbol{r}, t), \partial_k \phi_j(\boldsymbol{r}, t); \boldsymbol{r}, t\Big) \tag{2}$$

We use the same notation as in complement A_{XVIII} of reference [9]: $\dot{\phi}_j$ is the time derivative of the field ϕ_j, and $\partial_k \phi_j$ its spatial derivative with respect to the k component of its position. The Lagrangian density $\mathcal{L}$ depends on all these functions for various values of j and k and, eventually, directly on the position $\boldsymbol{r}$ in space, as well as on the time t (for example, when describing the effect of an outside time-dependent potential applied to the system).

Consider a possible history of the field, i.e. a path Γ for the field going from the value $\phi_j(\boldsymbol{r}, t_1)$ at an initial time t_1 to the final value $\phi_j(\boldsymbol{r}, t_2)$ at a final time t_2. The action $\mathscr{A}[\Gamma]$ associated with this path is written[1]:

$$\mathscr{A}[\Gamma] = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3 r \; \mathcal{L}\left(\phi_j(\boldsymbol{r}, t), \dot{\phi}_j(\boldsymbol{r}, t), \partial_k \phi_j(\boldsymbol{r}, t); \boldsymbol{r}, t\right) \tag{3}$$

The principle of least action postulates that among all possible paths starting from the same initial state and ending at the same final state, the path(s) actually followed by the system is the one (or are those) for which $\mathscr{A}[\Gamma]$ presents an extremum. To obtain this (or those) path(s), we compute the variation $\delta\mathscr{A}$ of the action for an infinitesimal variation of the path characterized by the infinitesimal variations $\delta\phi_j(\boldsymbol{r}, t)$, $\delta\left(\partial\phi_j(\boldsymbol{r}, t)/\partial t\right)$, and $\delta\left(\partial\phi_j(\boldsymbol{r}, t)/\partial r_k\right)$:

$$\delta\mathscr{A} = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3 r \sum_j \left[\delta\phi_j(\boldsymbol{r}, t)\frac{\partial \mathcal{L}}{\partial \phi_j} + \delta\dot{\phi}_j(\boldsymbol{r}, t)\frac{\partial \mathcal{L}}{\partial \dot{\phi}_j} \right.$$
$$\left. + \sum_k \delta\left(\partial_k \phi_j(\boldsymbol{r}, t)\right) \frac{\partial \mathcal{L}}{\partial(\partial_k \phi_j)}\right] \tag{4}$$

Since:

$$\delta\dot{\phi}_j(\boldsymbol{r}, t) = \frac{\partial}{\partial t}\,\delta\phi_j(\boldsymbol{r}, t) \qquad \text{and:} \qquad \delta\left(\partial_k \phi_j(\boldsymbol{r}, t)\right) = \partial_k\left(\delta\phi_j(\boldsymbol{r}, t)\right) \tag{5}$$

we can perform an integration by parts of the terms containing the time and space derivatives. We find that the integrated terms are zero, because of the boundary conditions for $\delta\phi_j(\boldsymbol{r}, t)$ at the initial and final times, and for $\boldsymbol{r} \to \infty$. The remaining terms are all proportional to $\delta\phi_j(\boldsymbol{r}, t)$, and we get[2]:

$$\delta\mathscr{A} = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3 r \sum_j \delta\phi_j(\boldsymbol{r}, t)\left[\frac{\partial \mathcal{L}}{\partial \phi_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}_j} - \sum_k \partial_k \frac{\partial \mathcal{L}}{\partial(\partial_k \phi_j)}\right] \tag{6}$$

As $\delta\mathscr{A}$ must be zero for any temporal or spatial variations of $\delta\phi_j(\boldsymbol{r}, t)$, we deduce the Lagrange equations for the field:

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}_j} = \frac{\partial \mathcal{L}}{\partial \phi_j} - \sum_k \partial_k \frac{\partial \mathcal{L}}{\partial(\partial_k \phi_j)} = 0} \tag{7}$$

[1] When $\boldsymbol{r}$ goes to infinity, we assume the Lagrangian density is zero or goes to zero sufficiently rapidly for the integral on the right-hand side of (3) to be convergent.

[2] Even when the function $\mathcal{L}$ does not directly depend on time, it does so indirectly when we replace in (2) the fields and their derivatives by their values for a given history of the field. The notation $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}_j}$ in (6) and (7) designates the derivative of this function with respect to t in the course of this history. This total derivative includes the contributions of all the partial derivatives with respect to all the variables of $\mathcal{L}$ appearing in the right-hand side of (2): the fields, their derivatives, and possibly time.

When dealing with continuous variables, and when the Lagrangian density depends on the spatial derivatives of the fields, we note the presence of the term in $\sum_k$; this term does not appear when the system is described by discrete variables. This equation is written in a more condensed form in the relativistic notations with $\mu = 0, 1, 2, 3$ labeling the space–time components:

$$\frac{\partial \mathcal{L}}{\partial \phi_j} = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_j)} \tag{8}$$

where $\partial_0 = \mathrm{d}/\mathrm{d}(ct)$ and $\partial_k = \mathrm{d}/\mathrm{d}x^k$ with $x^k = x, y, z$; on the right-hand side the index μ is summed from 0 to 3, following Einstein's convention.

Comment:

Adding to $\mathcal{L}$ the total time derivative of any function of the fields does not change the equations of motion; it simply adds to the action a term that depends only on the initial and final states of the system, but not on the actual path leading from one to the other. One can also add any spatial derivative without changing the Lagrangian, after integration over d^3r (provided the functions go to zero at infinity).

2. Symmetry transformations and current conservation

Consider now variations of the ϕ_j components having the form:

$$\delta\phi_j(\boldsymbol{r}, t) = \delta\varepsilon\, f_j\Big(\phi_l(\boldsymbol{r}, t), \dot{\phi}_l(\boldsymbol{r}, t), \partial_m\phi_l(\boldsymbol{r}, t); \boldsymbol{r}, t\Big) \tag{9a}$$

where f_j depends on all the variables inside the parentheses, for any values of l and m. It follows that:

$$\begin{aligned} \delta\dot{\phi}_j(\boldsymbol{r}, t) &= \delta\varepsilon\, \dot{f}_j\Big(\phi_l(\boldsymbol{r}, t), \dot{\phi}_l(\boldsymbol{r}, t), \partial_m\phi_l(\boldsymbol{r}, t); \boldsymbol{r}, t\Big) \\ \delta[\partial_k\phi_j(\boldsymbol{r}, t)] &= \delta\varepsilon\, \partial_k f_j\Big(\phi_l(\boldsymbol{r}, t), \dot{\phi}_l(\boldsymbol{r}, t), \partial_m\phi_l(\boldsymbol{r}, t); \boldsymbol{r}, t\Big) \end{aligned} \tag{9b}$$

The variation of the Lagrangian density is written:

$$\delta\mathcal{L} = \delta\varepsilon \sum_j \Big[f_j \frac{\partial \mathcal{L}}{\partial \phi_j} + \dot{f}_j \frac{\partial \mathcal{L}}{\partial \dot{\phi}_j} + \sum_k (\partial_k f_j)\, \frac{\partial \mathcal{L}}{\partial(\partial_k \phi_j)}\Big] \tag{9c}$$

As in relation (I-37) of chapter I, we assume that $\delta\mathcal{L}$ is proportional to a total time derivative[3] of a function Λ:

$$\delta\mathcal{L} = \delta\varepsilon\, \frac{\mathrm{d}}{\mathrm{d}t}\Lambda\Big(\phi_l(\boldsymbol{r}, t), \dot{\phi}_l(\boldsymbol{r}, t), \partial_m\phi_l(\boldsymbol{r}, t); \boldsymbol{r}, t\Big) \tag{10}$$

[3]When $\boldsymbol{r}$ goes to infinity, we assume that Λ goes to zero fast enough for the integral (1) yielding the Lagrangian L to remain finite. When Λ does not depend on $\dot{\phi}_j$, L and $L + \delta L$ are equivalent Lagrangians – *cf.* note 5 page 12.

The field transformations (9a) are then said to be symmetry transformations.

We now compute the time variation of the function $\sum_j f_j \, \partial\mathcal{L}/\partial\dot{\phi}_j$. Equation (7) for the time evolution of the field yields:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big[\sum_j f_j \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j}\Big] &= \sum_j \Big[\dot{f}_j \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j} + f_j \frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial\mathcal{L}}{\partial\dot{\phi}_j}\big)\Big] \\
&= \sum_j \Big[\dot{f}_j \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j} + f_j \Big(\frac{\partial\mathcal{L}}{\partial\phi_j} - \sum_k \partial_k \frac{\partial\mathcal{L}}{\partial(\partial_k\phi_j)}\Big)\Big] \\
&= \frac{\delta\mathcal{L}}{\delta\varepsilon} - \sum_{j,k} \partial_k \Big(f_j \frac{\partial\mathcal{L}}{\partial(\partial_k\phi_j)}\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\Lambda - \sum_{j,k} \partial_k \Big(f_j \frac{\partial\mathcal{L}}{\partial(\partial_k\phi_j)}\Big)
\end{aligned}
\tag{11}
$$

or:

$$
\frac{\mathrm{d}}{\mathrm{d}t}\Big[\sum_j f_j \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j} - \Lambda\Big] + \sum_{j,k} \partial_k \Big(f_j \frac{\partial\mathcal{L}}{\partial(\partial_k\phi_j)}\Big) = 0 \tag{12}
$$

This result has the form of a local conservation equation. Defining a local density $\rho(\boldsymbol{r},t)$ and the k component of its associated current $\boldsymbol{J}(\boldsymbol{r},t)$:

$$
\boxed{\rho(\boldsymbol{r},t) = \sum_j f_j \frac{\partial\mathcal{L}}{\partial\dot{\phi}_j} - \Lambda \qquad ; \qquad \big(\boldsymbol{J}(\boldsymbol{r},t)\big)_k = \sum_j f_j \frac{\partial\mathcal{L}}{\partial(\partial_k\phi_j)}} \tag{13}
$$

we get[4]:

$$
\frac{\mathrm{d}}{\mathrm{d}t}\rho(\boldsymbol{r},t) + \boldsymbol{\nabla}\cdot\boldsymbol{J}(\boldsymbol{r},t) = 0 \tag{14}
$$

This equation, similar to the one expressing charge conservation, means that the local increase of the density (of a physical quantity to be determined) in a small volume is equal to the incoming flux of a current $\boldsymbol{J}$ across the surface of that small volume. In the discrete case studied in chapter I, Noether's theorem yielded only one constant of motion; in our present case with continuous field variables, there are an infinite number of them as there are an infinite number of points $\boldsymbol{r}$ in space. This is because assuming the invariance of $\mathcal{L}$ at each point in space is much stronger than assuming the invariance of its spatial integral L.

3. Generalization, relativistic notation

Time plays a prominent role in relation (10). It can be modified to become more symmetric with respect to time and space, and we set:

$$
\delta\mathcal{L} = \delta\varepsilon\Big[\frac{\mathrm{d}}{\mathrm{d}t}\Lambda + \sum_k \mathrm{d}_k\Lambda_k\Big] = \delta\varepsilon\ \partial_\mu\Lambda^\mu \tag{15}
$$

[4]This equation is generally written with a partial time derivative $\partial\rho/\partial t$ (instead of a total derivative) when one adopts a point of view where ρ is a function of only $\boldsymbol{r}$ and t, ignoring its possible dependence on the fields ϕ_j.

where the index k is summed from 1 to 3, and μ from 0 to 3. This expression involves four functions Λ^μ similar[5] to Λ, and we set $\Lambda^0 = c\Lambda$ and $\Lambda^{\mu=k} = \Lambda_k$; the ∂_μ have been defined in relation (8). We can follow a reasoning almost the same as in § 2, but we end up with a modified expression for the current:

$$\boxed{\left(\boldsymbol{J}(\boldsymbol{r},t)\right)_k = \sum_j f_j \frac{\partial \mathcal{L}}{\partial(\partial_k \phi_j)} - \Lambda_k} \qquad \text{with: } k = 1,2,3 \tag{16}$$

This is a generalization of Noether's theorem to the case where the variation of the Lagrangian density $\mathcal{L}$ also contains spatial derivatives. In relativistic notation, we write J^μ the components of the four-vector $(c\rho, \boldsymbol{J})$, and relation (14) simply becomes:

$$\partial_\mu J^\mu = 0 \tag{17}$$

Noether's theorem has numerous applications in field theory, including in quantum theory where the conserved quantities and their associated currents play a central role; the interested reader may consult references [24, 25] for example. The applications to the Schrödinger, Klein–Gordon, and Dirac equations will be discussed in complement C_{VI}. We shall only mention here a general and very simple example, the local conservation of energy and momentum.

4. Local conservation of energy

We assume that the Lagrangian density is not explicitly time-dependent, and introduce a variation of the field $\phi(\boldsymbol{r},t)$ that is a shift in time by a quantity $\delta\varepsilon$. For the sake of simplicity, we assume that the field is a scalar so that the index j is no longer needed. We have $\delta\phi = \delta\varepsilon\,\dot{\phi}$, and hence $f = \dot{\phi}$. As a result, $\mathcal{L}$ will vary since it depends on the field and its derivatives; as we assume that $\mathcal{L}$ does not directly depend on time, we have $\delta\mathcal{L} = \delta\varepsilon\ \mathrm{d}\mathcal{L}/\mathrm{d}t$. This yields $\Lambda = \mathcal{L}$ (and $\Lambda^{1,2,3} = 0$), and relations (13) lead to a local energy density ρ_E and an energy current $\boldsymbol{J}_E$ written as:

$$\rho_E(\boldsymbol{r},t) = \dot{\phi}\,\frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \mathcal{L} \quad ; \quad \left(\boldsymbol{J}_E(\boldsymbol{r},t)\right)_k = \dot{\phi}\,\frac{\partial \mathcal{L}}{\partial(\partial_k \phi)} \tag{18}$$

Integrating over space $\rho_E(\boldsymbol{r},t)$ yields the definition of the Hamiltonian; ρ can thus be identified to the Hamiltonian density.

If, for example, the Lagrangian density is written:

$$\mathcal{L} = \frac{1}{2}\left[\dot{\phi}^2 - c^2(\boldsymbol{\nabla}\phi)^2 - \omega^2\phi^2\right] = \frac{1}{2}\left[c^2(\partial^\mu\phi)(\partial_\mu\phi) - \omega^2\phi^2\right] \tag{19}$$

[5] As before, we assume that, when $\boldsymbol{r}$ goes to infinity, the Λ^μ go to zero fast enough for the spatial integral yielding L to remain finite.

where c and $\omega = mc^2/\hbar$ are constants, we get:

$$\rho_E(\boldsymbol{r},t) = \frac{1}{2}\left[\dot{\phi}^2 + \omega^2\phi^2 + c^2(\boldsymbol{\nabla}\phi)^2\right] \quad ; \quad \boldsymbol{J}_E(\boldsymbol{r},t) = -c^2\dot{\phi}\,\boldsymbol{\nabla}\phi \tag{20}$$

The energy density ρ_E is the sum of the kinetic energy density of the field (term in $\dot{\phi}^2$) and its potential energy density (terms in ϕ^2 and $(\boldsymbol{\nabla}\phi)^2$); the energy current includes temporal and spatial derivatives of ϕ.

In electromagnetism, the field has 6 components, those of the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$, often arranged in a second-order antisymmetric tensor. The energy density at each point is proportional to $(\boldsymbol{E}^2 + c^2\boldsymbol{B}^2)$ and the current $\boldsymbol{J}_E$ to the Poynting vector, proportional to $\boldsymbol{E}\times\boldsymbol{B}$.

Exercise: We assume that the Lagrangian density does not depend directly on the position but only through the values of the field and its derivatives. Show that this yields 3 local conservation laws, with the following densities and currents:

$$\rho_{x_i}(\boldsymbol{r},t) = \partial_{x_i}\phi\,\frac{\partial\mathcal{L}}{\partial\dot{\phi}} \quad ; \quad \left(\boldsymbol{J}_{x_i}(\boldsymbol{r},t)\right)_k = \partial_{x_i}\phi\,\frac{\partial\mathcal{L}}{\partial(\partial_k\phi)} - \delta_{k,x_i}\,\mathcal{L} \tag{21}$$

where $x_i = x, y, z$. These relations express a local conservation of the three components of the field's momentum. The various values of i and k yield the 9 components of the current, which form a 3×3 tensor, the so-called "constraint tensor".