ANALYTICAL MECHANICS

Lectures Notes 2018/2019

A. Ponno

January 10, 2019

Contents

Chapter 1

An overview of Lagrangian mechanics

1.1 From Newton to Lagrange equations

The mathematical structure of Newtonian mechanics is the theory of ordinary differential equations (ODEs) of second order of the form

$$M\ddot{X} = F(X, \dot{X}, t) \ , \tag{1.1}$$

where $t \mapsto X(t) \in \mathbb{R}^N$ is the unknown vector valued function (curve), M is a $N \times N$ constant, symmetric, positive definite matrix, $F : \mathbb{R}^{2N+1} \to \mathbb{R}$ is a given vector field[1]. The great success of such a theory consists in the solution of important problems in celestial mechanics, starting with the two body problems. Modern results of non-perturbative celestial mechanics (n-body problem) are still based on the study of the Newton equations (1.1), and in particular on their well posedness (existence, uniqueness and regularity of the solution).

Newtonian mechanics displays some limits, most of them technical in character and due to the fact that one has to work with differential equations. Thus, for example, determining the presence of symmetries and/or first integrals and using them to reduce the dimensionality of the problem can be quite cumbersome. Also the treatment of perturbation problems can be very difficult. However, the most important shortcomings of Newtonian mechanics is the difficulty in treating constrained systems. Indeed, when geometrical constraints are imposed on a given system, on the right hand side of the Newton equation (1.1) one has to add to the "active force" F a force, or reaction R, which is necessary to render the motions compatible with the constraint (e.g., one needs a force to constrain a mass point to move on a sphere subject to gravity). The problem is that such a reaction, at variance with the active force, is not known, *in general*, as a function of X, $\dot{X}$ and t. This is obvious, since $R = M\ddot{X} - F$, so that R displays a dependence on the acceleration. As a matter of fact, it was such a second obstruction that led to the formulation of Lagrangian mechanics.

Lagrangian mechanics is the mathematical theory describing the dynamics of mechanical systems subject to constraints. In the case of holonomic constraints the mechanical system is

[1] In the specific case of n mass points moving in the d-dimensional physical space one has $N = nd$, and M is block diagonal with each of the n diagonal blocks of the form $m_i\mathbb{I}_d$, m_i being the mass of the i-th particle ($i = 1, \ldots, n$) and $\mathbb{I}_d$ denoting the $d \times d$ identity.

supposed to move on a given L-dimensional manifold (smooth surface of co-dimension $N-L$ in $\mathbb{R}^N$) given by an implicit equation of the form $\Phi(X,t)=0$, where $\Phi:\mathbb{R}^N\times\mathbb{R}\to\mathbb{R}^{N-L}$ (here $1\le L\le N-1$; the limit case $L=N$ means unconstrained systems). Under the usual hypotheses of the implicit function theorem such a manifold is given in the parametric form $\mathbb{R}^L\ni q\mapsto X(q,t)\in\mathbb{R}^N$. The free parameters $q=(q_1,\dots,q_L)$ are the so-called generalized or Lagrangian coordinates. The fundamental hypothesis that allows to close the problem is that of ideal holonomic constraint, i.e. the hypothesis that the field of reactions R be orthogonal to the constraint manifold (the so-called D'Alambert principle). Such a condition is locally expressed by requiring

$$R\cdot\frac{\partial X}{\partial q_i}=0\ \ \forall i=1,\dots,L\ \ ,\ \forall t\ , \tag{1.2}$$

where $X(q(t),t)$ is understood in the above expression. Upon scalar multiplication of the Newton equation $M\ddot{X}=F+R$ by $\partial X/\partial q_i$ the, due to condition (1.2), the contribution of the reaction disappears and one is left with the L projections of the Newton equation onto the tangent space to the constraint manifold, namely

$$M\ddot{X}\cdot\frac{\partial X}{\partial q_i}=Q_i\ \ ;\ \ Q_i:=F\cdot\frac{\partial X}{\partial q_i}\ . \tag{1.3}$$

The generalized force Q_i on the right hand is just the definition of the projection of F along the i-th coordinate direction. The left hand side is instead worked out by writing it as

$$M\ddot{X}\cdot\frac{\partial X}{\partial q_i}=\frac{\mathrm{d}}{\mathrm{d}t}\left(M\dot{X}\cdot\frac{\partial X}{\partial q_i}\right)-M\dot{X}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial X}{\partial q_i}$$

and taking into account the two identities

$$\frac{\partial X}{\partial q_i}=\frac{\partial\dot{X}}{\partial\dot{q}_i}\ \ ;\ \ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial X}{\partial q_i}=\frac{\partial\dot{X}}{\partial q_i}\ . \tag{1.4}$$

Defining the kinetic energy $K:=(\dot{X}\cdot M\dot{X})/2$ of the system, one easily shows that the equations (1.3) take on the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial K}{\partial\dot{q}}-\frac{\partial K}{\partial q}=Q\ , \tag{1.5}$$

to be meant by components. This is the most general form of Lagrange equations. If the force F admits a conservative component, i.e. if a function $U(X,t)$ exists such that $F=-\nabla_X U+\mathscr{F}$, then one easily checks that the Lagrange equations (1.5) read

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial\dot{q}}-\frac{\partial L}{\partial q}=\mathscr{Q}\ , \tag{1.6}$$

where $\mathscr{Q}_i:=\mathscr{F}\cdot\partial X/\partial q_i$, and

$$L(q,\dot{q},t):=K(q,\dot{q},t)-U(q,t) \tag{1.7}$$

is the Lagrange function, or Lagrangian of the system. In the definition (1.7) we have set, with abuse of notation

$$K(q,\dot{q},t) := \frac{1}{2}\dot{X}\cdot M\dot{X} \quad ; \quad \dot{X} = \sum_j \frac{\partial X}{\partial q_j}\dot{q}_j + \frac{\partial X}{\partial t}$$

and

$$U(q,t) := U(X(q,t),t) \ .$$

Once one has solved the Lagrange equations determining the motion $t \mapsto q(t)$ of the system, one inserts $X(q(t),t)$ into the original Newton equation and solves for $R = M\ddot{X} - F$, which allows to completely close the problem, in principle.

Actually, it is possible to close the problem of ideal holonomic constraints in another equivalent way, without introducing the Lagrange equations (1.5) or (1.6). Indeed, the L conditions (1.2) are equivalent to the requirement

$$R = \sum_{r=1}^{N-L} c_r(X,t)\nabla_X\Phi_r(X,t) \ , \tag{1.8}$$

where the Φ_r are the $N-L$ components of Φ, whereas the c_r are unknown coefficients. Notice that the above mentioned hypothesis of the implicit function theorem, allowing the parametric representation of the constraint manifold of dimension L, is that of linear independence of the $N-L$ gradients $\nabla\Phi_r$, so that $R = 0$ iff $c_r = 0$ for any r. Now, by explicitly computing the two identities

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(X(t),t) = 0 \quad ; \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi(X(t),t) = 0 \ ,$$

and making use of the Newton equation $\ddot{X} = M^{-1}(F+R)$ one shows that the constraint reaction R can be explicitly computed a priori as a function of X, $\dot{X}$ and t (i.e. one determines the coefficients c_r entering (1.8)). Finally, one can reduce the dimension of the Newton equations from N to L. As a matter of fact, such an equivalent approach reveals much more involved in practice.

1.2 General properties of Lagrange equations

The Lagrange equations for conservative systems, i.e. for system subject to purely conservative forces ($F = -\nabla_X U$), the Lagrange equations take on the simple standard form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \ , \tag{1.9}$$

with $L = K - U$ as specified above. All the properties of the Lagrange equations reported in the sequel are easily checked by means of direct simple computations, and are *independent of the mechanical origin of the equations* themselves.

1. The L-equations are left invariant in form by any (possibly) time dependent change of coordinates
$$q \mapsto Q(q,t) \ \ ; \ \ \dot{q} \mapsto \dot{Q}(q,t) \ ,$$
where here and in what follows $\dot{f}(q(t),t) := \sum_j (\partial f/\partial q_j)\dot{q}_j + \partial f/\partial t$.

2. For any constant $c \neq 0$ and any function $F(q,t)$, the Lagrangians L and $L' := cL + \dot{F}$ are equivalent, in the sense that their associated L-equations are the same. The change of Lagrangian $L \to L'$ is referred to as a gauge transformation.

3. If $\partial L/\partial t = 0$ then the energy function
$$\mathscr{H}(q,\dot{q}) := \frac{\partial L}{\partial \dot{q}} \cdot \dot{q} - L(q,\dot{q}) \tag{1.10}$$
is a first integral, i.e. is constant along the solutions of the L-equations. The function $\mathscr{H}$ is known as the Jacobi integral; in the mechanical case it is the total energy of the system.

4. Consider a one-parameter family of coordinate transformations $q \mapsto Q = \Phi^s(q,t)$, $s \in I \subseteq \mathbb{R}$, such that $\Phi^0(q,t) = q$. Let $u(q,t) := \partial\Phi^s(q,t)/\partial s|_{s=0}$. Then, the Nöther theorem holds:
$$\frac{\partial}{\partial s} L\left(\Phi^s(q,t), \frac{\mathrm{d}}{\mathrm{d}t}\Phi^s(q,t), t\right)\Big|_{s=0} = 0 \ \Rightarrow \ J_u := u \cdot \frac{\partial L}{\partial \dot{q}} = \text{const.} \tag{1.11}$$
along the solutions of the L-equations (i.e. J_u is a first integral). The one-parameter family Φ^s satisfying the condition on the left of (1.11) is said to be an admissible (infinitesimal) symmetry for the Lagrangian at hand. Observe that for mechanical Lagrangians L is a quadratic form in $\dot{q}$ so that J_u is always linear in $\dot{q}$. A particular case of the Nöther theorem is that of an ignorable or cyclic coordinate, e.g. q_1, such that $\partial L/\partial q_1 = 0$. In this case one can choose $\Phi^s(q) = (q_1 + s, q_2, \ldots, q_L)$, and the quantity $J_1 := \partial L/\partial \dot{q}_1$ is constant, as also follows directly from the L-equations. One can prove that if Φ^s is a one-parameter group (with respect to the composition) of symmetry for L, then there exists a change to new coordinates such that one of them is ignorable (this follows by rectifying the vector field u generating the group).

Exercise 1.1. *Show that if $\mathscr{H}$ is the Jacobi integral corresponding to $L(q,\dot{q})$, the Jacobi integral $\mathscr{H}'$ corresponding to the gauge-equivalent Lagrangian $L' = cL + \dot{F}$, $\partial F/\partial t = 0$, is given by $\mathscr{H}' = c\mathscr{H}$ (hint: $p' = cp + \nabla_q F$ is the Lagrangian momentum corresponding to L').*

Exercise 1.2. *Consider the Lagrangian of a particle of mass m and charge q moving in a given electromagnetic field, namely*
$$L(x,\dot{x},t) = m\frac{|\dot{x}|^2}{2} + \frac{q}{c}A(x,t)\cdot\dot{x} - q\phi(x,t) \ , \tag{1.12}$$
being A and ϕ the vector and scalar potential, respectively, and c the velocity of light.

1. *Show that the Lagrange equation read*

$$m\ddot{x} = q\left(E + \frac{1}{c}\dot{x} \wedge B\right) \ , \tag{1.13}$$

where, in the Lorentz force on the right hand side, the electric field E and the magnetic field B are defined in terms of the potentials by

$$E := -\nabla_x \phi - \frac{1}{c}\frac{\partial A}{\partial t} \quad ; \quad B := \nabla_x \wedge A \ . \tag{1.14}$$

2. *Observe that E and B are invariant with respect to the gauge transformation of the potentials*

$$A \mapsto A' = A + \nabla\chi \quad ; \quad \phi \mapsto \phi' = \phi - \frac{1}{c}\frac{\partial \chi}{\partial t} \ , \tag{1.15}$$

so that the Lagrange equation (1.13) is invariant as well.

3. *Show that under the gauge transformation (1.15) $L \mapsto L' = L + \dot{F}$, with $F := (q/c)\chi$.*

4. *Show that in the autonomous case (i.e. $\partial A/\partial t = 0$, $\partial\phi/\partial t = 0$), the Jacobi integral is given by*

$$\mathscr{H}(x, \dot{x}) = \frac{m|\dot{x}|^2}{2} + q\phi(x) \ , \tag{1.16}$$

and explain why this is obviously invariant under gauge transformations of the e.m. potentials that are independent of time.

1.3 Hamilton first variational principle

Let us denote by $C_{a,b}$ the space of smooth curves $[t_1, t_2] \ni t \mapsto q(t) \in \mathbb{R}^L$ with fixed ends $q(t_1) := a$, $q(t_2) := b$. Then, for any given Lagrangian, the so-called *action functional* $A_L : C_{a,b} \to \mathbb{R}$ is defined by

$$A_L[q] := \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)\ \mathrm{d}t \ . \tag{1.17}$$

Notice that the curve $t \mapsto q(t)$ appearing on the right hand side of the latter formula denotes any element of $C_{a,b}$ (containing also the curve solution of the Lagrange equations with fixed ends, if any). Pay attention to such abuses of notation made here and in the sequel.

The weak, or Gateaux differential of A_L in $q \in C_{a,b}$ with increment $\delta q \in C_{0,0}$ is defined by

$$\delta A_L := \frac{\mathrm{d}}{\mathrm{d}\epsilon} A_L[q + \epsilon\delta q]\Big|_{\epsilon=0} = \lim_{\epsilon\to 0} \frac{A_L[q + \epsilon\delta q] - A[q]}{\epsilon} \ . \tag{1.18}$$

Such a quantity is also known as Lagrange first variation of A_L, and is the linear part of the increment of the functional. Notice that the increment δq is a finite curve with fixed ends set to zero.

Remark 1.1. *A more precise notation for δA_L would be $\mathrm{d}A_L[q]h := \mathrm{d}A_L[q+\epsilon h]/\mathrm{d}\epsilon|_{\epsilon=0}$, where h denotes the increment curve (δq). However, for later convenience, we keep on making use of the simpler δ-notation introduced above, also because it is the most widespread one in theoretical physics.*

Now, according to definition (1.18), a simple calculation shows that

$$\delta A_L = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} \right) \cdot \delta q \ \mathrm{d}t \ . \tag{1.19}$$

The critical points of A_L are those points of $C_{a,b}$ where $\delta A_L = 0$ independently of the increment. The following proposition, characterizing the critical points of the action A_L, is known as the Hamilton first variational principle.

Proposition 1.1.

$$\delta A_L = 0 \ \ \forall \delta q \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \ . \tag{1.20}$$

$\lhd$ PROOF. The implication $\Leftarrow$ is obvious. The opposite one follows by contradiction (reductio ad absurdum): if the Lagrange equations do not hold, then one can properly choose the increment δq in such a way that $\delta A_L \neq 0$. $\rhd$

In words, the critical points of the action A_L are the solutions of the Lagrange equations in $C_{a,b}$. We stress that the latter is a boundary value problem that, depending on the interval $[t_1, t_2]$, may have no solution, unique solution, or infinitely many solutions.

Exercise 1.3. *For the harmonic oscillator with unit frequency $L=(\dot{q}^2 - q^2)/2$, and the Lagrange equation is $\ddot{q} = -q$. Take $t_1 = 0$ and $t_2 = T$, $q(0) = a$ and $q(T) = b$. Show that*

- *if $T \neq k\pi$, $k \in \mathbb{Z}$, then the boundary value problem has unique solution (write it);*
- *if $T = k\pi$ and $b = (-1)^k a$ the problem has infinitely many solutions, namely a one parameter family of them (write it);*
- *if $T = k\pi$ and $b \neq (-1)^k a$ there is no solution.*

Exercise 1.4. *Consider the boundary value problem in general. The Lagrange equations can be written in second order form $\ddot{q} = g(q, \dot{q}, t)$. Denote the unique solution of the associate* initial value problem *with initial conditions $q(t_1) = a$ and $\dot{q}(t_1) = v$ by $q(t) = \phi(t, t_1; a, v)$. Now, in the boundary value problem the initial velocity v is not known. Find under which condition v is uniquely determined by the boundary data t_1, t_2 and a, b. Check such a condition on the harmonic oscillator case.*

Remark 1.2. *The condition of the exercise above is violated, in general for special choices of the end time t_2 when t_1 and the initial point (a, v) of the phase space are fixed.*

We observe that by means of the Hamilton principle some of the properties of Lagrangian systems listed above become obvious. For example, the gauge invariance of the L-equations is immediately proven: the action associated to $cL + \dot{F}$ (for any constant $c \neq 0$ and any function $F(q,t)$) is $cA_L + \Delta F$, where $\Delta F := F(b,t_2) - F(a,t_1)$ is a constant that vanishes under differentiation with fixed ends. Also the invariance in form of the L-equations under point transformations $q \mapsto Q(q,t)$ is easily proven: the critical points of the action $A_{L'}$ associated to the transformed Lagrangian $L'(Q,\dot{Q},t) := L(q(Q,t),\dot{q}(Q,t),t)$ are the solutions of the L-equations in the new variables.

1.4 Maupertuis-Jacobi variational principle

In the autonomous case, i.e. $\partial L/\partial t = 0$, the flow of the L-equations preserves the Jacobi integral $\mathscr{H} = p \cdot \dot{q} - L$, where $p := \partial L/\partial \dot{q}$. In such a case, one can formulate a variational principle taking into account such a fact, which means restricting to curves $t \mapsto (q(t),\dot{q}(t))$ and on the surface $\Sigma_E := \{(q,\dot{q}) :\ \mathscr{H}(q,\dot{q}) = E\}$. In this case, the action reads

$$A_L[q] = \int_{t_1}^{t_2} L\ \mathrm{d}t = \int_{t_1}^{t_2} (p \cdot \dot{q} - \mathscr{H})\ \mathrm{d}t = \int_{t_1}^{t_2} (p \cdot \dot{q})\ \mathrm{d}t - E(t_2 - t_1) := A_E[q] - E(t_2 - t_1)\ ,$$

where the reduced action $A_E := \int_{t_1}^{t_2} (p \cdot \dot{q}) \mathrm{d}t$ has been defined. The constant contribution vanishes under differentiation, which suggests the formulation of the following variational principle, named after Maupertuis and Jacobi.

Proposition 1.2.

$$\delta A_E = 0 \quad \forall \delta q \qquad \Longleftrightarrow \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0\ . \tag{1.21}$$

$\triangleleft$ PROOF. Observe that the displaced curves must belong to Σ_E, namely

$$\mathscr{H}(q + \epsilon \delta q, \dot{q} + \epsilon \delta \dot{q}) = E$$

which, upon taking the derivative with respect to ϵ at $\epsilon = 0$, implies

$$\frac{\partial \mathscr{H}}{\partial q} \cdot \delta q + \frac{\partial \mathscr{H}}{\partial \dot{q}} \cdot \delta \dot{q} = 0\ . \tag{1.22}$$

Thus, observing that $p \cdot \dot{q} = L + \mathscr{H}$, one gets

$$\delta A_E = \delta \int_{t_1}^{t_2} (L + \mathscr{H})\ \mathrm{d}t = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} \right) \cdot \delta q\ \mathrm{d}t + \int_{t_1}^{t_2} \left(\frac{\partial \mathscr{H}}{\partial q} \cdot \delta q + \frac{\partial \mathscr{H}}{\partial \dot{q}} \cdot \delta \dot{q} \right)\ \mathrm{d}t\ .$$

The statement (1.21) follows now from relation (1.22). $\triangleright$

Chapter 2

From Lagrangian to Hamiltonian mechanics

2.1 Hamilton equations

The L-equations in second order form $\ddot{q} = g(q, \dot{q}, t)$ can always be written as an equivalent system of first order: $\dot{q} = v$, $\dot{v} = g(q, v, t)$. This is not the only way of doing that. Another approach is to make use of the Lagrangian momentum

$$p := \frac{\partial L}{\partial \dot{q}}(q, \dot{q}, t) \tag{2.1}$$

in place of the velocity $\dot{q}$ as a variable. This is possible if one can express $\dot{q}$ in terms of p. From (2.1), it follows that if the Hessian of L with respect to the velocities $(\partial^2 L/\partial \dot{q}^2)$ is non singular, then there exists a function f such that

$$\dot{q} = f(q, p, t)\ . \tag{2.2}$$

Definition 2.1. *The function*

$$H(q, p, t) := p \cdot \dot{q} - L\Big|_{\dot{q}=f} = p \cdot f(q, p, t) - L(q, f(q, p, t), t) \tag{2.3}$$

is called Hamilton function, or Hamiltonian, of the given Lagrangian system.

Observe that H is the Legendre (or Donkin, to some authors) transformation of L. The following proposition holds.

Proposition 2.1. *The Lagrange equations $\dot{p} = \partial L/\partial q$ are equivalent to the Hamilton (H) equations*

$$\boxed{\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q}}\ . \tag{2.4}$$

Moreover, $\partial H/\partial t = -\partial L/\partial t$.

◁ PROOF. Recalling that $p = \partial L/\partial \dot{q}$, one finds

$$\mathrm{d}H = -\frac{\partial L}{\partial q} \cdot \mathrm{d}q + f \cdot \mathrm{d}p - \frac{\partial L}{\partial t} \ ,$$

which in turn implies

$$\frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} \ \ ; \ \ \frac{\partial H}{\partial p} = f \ \ ; \ \ \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \ .$$

The statement of the proposition follows now by taking into account that $\dot{q} = f$ (equivalent to the definition of p) and the L-equations $\dot{p} = \partial L/\partial q$. ▷

One easily checks that in the autonomous case ($\partial H/\partial t = 0$) the Hamiltonian H is a constant of motion, i.e. $\dot{H} = 0$ along the solution of the H-equations. On the other hand, by the identity $\partial H/\partial t = -\partial L/\partial t = 0$, the Jacobi integral $\mathscr{H}$ is constant along the solution of the L-equations. From the definitions (2.3) of the Hamiltonian and (1.10) of the Jacobi integral, one immediately finds that

$$H(q,p) = \mathscr{H}(q, f(q,p)) \ . \tag{2.5}$$

where $\dot{q} = f(q,p)$ as explained above.

Example 2.1. *Consider a mechanical Lagrangian $L = K(q,\dot{q}) - U(q)$, with kinetic energy given by $K(q,\dot{q}) = (\dot{q} \cdot M(q)\dot{q})/2$, $M(q)$ being the mass or kinetic matrix (symmetric and positive definite). In this case $\mathscr{H} = K + U$. On the other hand, $p = M(q)\dot{q}$, so that $\dot{q} = f(q,p) := M^{-1}(q)p$, and the Hamiltonian is given by $H = (p \cdot M^{-1}(q)p)/2 + U(q)$.*

Exercise 2.1. *Consider the Lagrangian (1.12) of a charged particle in a given e.m. field. Show that the corresponding Hamiltonian is*

$$H(x,p,t) = \frac{|p - (q/c)A(x,t)|^2}{2m} + q\phi(x,t) \ . \tag{2.6}$$

Show that, in the autonomous case (i.e. $\partial A/\partial t = 0$, $\partial \phi/\partial t = 0$), formula (2.5) holds with the Jacobi integral (1.16).

Exercise 2.2. *Show that, if $H(q,p,t)$ is the Hamiltonian associated to $L(q,\dot{q},t)$, the Hamiltonian $H'(q,p,t)$ associated to the gauge equivalent Lagrangian $L' = aL + \dot{F}$ (a a constant) is given by*

$$H'(q,p,t) = aH\left(q, \frac{p - \nabla_q F}{a}, t\right) - \frac{\partial F}{\partial t} \ . \tag{2.7}$$

Notice that the momentum conjugated to q is $p = a\partial L/\partial \dot{q} + \nabla_q F$. Compute the Hamiltonian H' corresponding to the gauge displaced Lagrangian L' of a charged particle in an e.m. field (where $a = 1$ and $F = (q/c)\chi$).

The procedure described above fails if the Hessian of L with respect to the velocities is singular. This happens for example in the case of Lagrangians that are linear in the velocity $\dot{q}$. However, a Hamiltonian formulation of the dynamics may exist even in such pathological cases.

Example 2.2. *Consider a particle of zero mass and charge q moving in the (x, y) plane subject to a constant, uniform magnetic field B_0 orthogonal to the plane and to an electric potential $\phi(x, y)$. The Lagrangian of the system is*

$$L = \frac{q|B_0|}{2c}(x\dot{y} - y\dot{x}) - q\phi(x, y) \ ,$$

which can be obtained by (1.12) setting $m = 0$, $A = (B_0 \wedge x)/2$ and $\phi = \phi(x, y)$. The components of the momentum p are $p_x = -(q\gamma/2)y$ and $p_y = (q\gamma/2)x$, where $\gamma := |B_0|/c$. The L-equations read

$$\dot{y} = \frac{1}{\gamma}\frac{\partial\phi}{\partial x} \quad ; \quad \dot{x} = -\frac{1}{\gamma}\frac{\partial\phi}{\partial y} \ . \tag{2.8}$$

The Jacobi integral in this case is $\mathscr{H} = q\phi(x, y)$. The Legendre transformation here obviously fails, and one can check that, for example, $\dot{p}_x \neq -\partial\mathscr{H}/\partial x$. On the other hand, upon setting $y = q$, $x = p$ and $h(q, p) := \phi(p, q)/\gamma$, the L-equations (2.8) take on the standard Hamiltonian form $\dot{q} = \partial h/\partial p$, $\dot{p} = -\partial h/\partial p$.

2.2 Hamilton second variational principle

The definition of the the Hamiltonian (2.3) and the Proposition 2.1 lead to the following variational principle.

Proposition 2.2. *The solutions of the Hamilton equations $\dot{q} = \partial H/\partial p$, $\dot{p} = -\partial H/\partial q$ are the critical points (i.e. curves) of the Hamiltonian action functional*

$$A_H[q, p] := \int_{t_1}^{t_2} [p \cdot \dot{q} - H(q, p, t)] \, \mathrm{d}t \tag{2.9}$$

in the space of the smooth curves $t \mapsto (q(t), p(t))$ such that $q(t_1)$ and $q(t_2)$ are fixed.

$\lhd$ PROOF. The differential of A_H at (q, p) with increments $(\delta q, \delta p)$ satisfying $\delta q(t_1) = 0$ and $\delta q(t_2) = 0$, is

$$\delta A_H = \int_{t_1}^{t_2} \left[\left(\dot{q} - \frac{\partial H}{\partial p} \right) \cdot \delta p + \left(-\dot{p} - \frac{\partial H}{\partial q} \right) \cdot \delta q \right] \mathrm{d}t + p \cdot \delta q \Big|_{t_1}^{t_2} \ . \tag{2.10}$$

The boundary term vanishes, and $\delta A_H = 0 \ \forall \ \delta q, \delta p$ iff the Hamilton equations hold. $\rhd$

Remark 2.1. *The critical points of the action A_H are the solutions of the Hamilton equations with fixed ends on $q(t)$ and no boundary condition on $p(t)$. On the other hand, since the Hamilton equations are of first order, the associated boundary value problem with $p(t_1)$ and $p(t_2)$ also fixed has no solution, in general (try to understand why).*

2.2.1 Properties of the action

Let us regard the action (2.9) as a function of the arrival time $t_2 := t$ and of the arrival coordinate $q(t_2) := q$, namely let us define

$$S(q,t) := \int_{t_1}^{t} (y \cdot \dot{x} - H(x,y,t))\ \mathrm{d}t\ , \tag{2.11}$$

the integral on the right being computed on curves $t \mapsto (x,y)(t)$ such that $x(t_1) = a$ is fixed and $x(t) = q$ is an independent variable. The action function $S(q,t)$ is the value taken on by the action functional A_H on curves that have the arrival point $q(t)$ in common at time t. Taking the differential of (2.11) one gets

$$\mathrm{d}S = \frac{\partial S}{\partial q} \cdot \mathrm{d}q + \frac{\partial S}{\partial t}\mathrm{d}t = p \cdot \mathrm{d}q - H\mathrm{d}t\ ,$$

which is equivalent to

$$p = \frac{\partial S}{\partial q}\ \ ;\ \ \frac{\partial S}{\partial t} + H(q,p,t) = 0\ . \tag{2.12}$$

Remark 2.2. *Observe that the variation of $y(t_2) = p$ does not affect S, which means $\partial S/\partial p = 0$.*

The two relations (2.12) together give rise to a partial differential equation, namely the *Hamilton-Jacobi* equation

$$\frac{\partial S}{\partial q} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0\ , \tag{2.13}$$

on which we will come back later. We here notice only that the action function S satisfies (i.e. is a solution of) the Hamilton-Jacobi equation, and its deep meaning will be clarified below, in the framework of the theory of the transformations of H-systems.

2.3 General properties of the Hamilton equations

The following general properties of the H-equations (2.4) can be easily checked to hold, independently of the fact that the H-system at hand correspond to a L-system via a Legendre transformation.

1. Along the solutions of the H-equations
$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}\ , \tag{2.14}$$
so that H is a first integral iff H does not depend (explicitly) on time.

2. The Hamiltonians H and $H' = H + \psi(t)$ are equivalent, i.e. the H-equations are invariant under any time-dependent translation of the Hamiltonian.

3. Given a non-autonomous Hamiltonian $H(q,p,t)$, $(q,p) \in \mathbb{R}^{2n}$ it is always possible to associate to it the equivalent autonomous problem defined by
$$K(q,\xi,p,\eta) = H(q,p,\xi) + \eta\ \ ;\ \ \xi(0) = 0\ ,$$
where $(\xi,\eta) \in \mathbb{R}^2$ is a pair coordinate-momentum (so that $\dot{\xi} = \partial K/\partial \eta = 1$).

2.3.1 Poisson bracket

Given a Hamiltonian system defined by $H(q,p,t)$, let us denote by $\Phi^t_H(q_0,p_0) := (q(t),p(t))$ the solution of the associated Hamilton equations (2.4) with initial condition (q_0,p_0). The map $\Phi^t_H : \Gamma \to \Gamma$, which is called the *Hamiltonian flow* of (or associated to) H, is a one parameter family of diffeomorphisms of the phase space Γ. If H is independent of time ($\partial H/\partial t = 0$), and the solution exists globally, i.e. for $t \in \mathbb{R}$, then Φ^t_H is a one parameter group[1] of diffeomorphisms of Γ. A condition ensuring that Φ^t_H exists for any t and any initial condition is that Γ is a compact smooth manifold. Since Φ^t_H preserves the value of H, then one can restrict the phase space to the surface of constant energy, namely $\Gamma = \Sigma_E = \{(q,p) : H(q,p) = E\}$. Thus, the compactness of Σ_E ensures that Φ^t_H is a (commutative) group of diffeomorphisms of Σ_E, the group operation being the composition: $\Phi^t_H \circ \Phi^s_H = \Phi^{t+s}_H$; $\Phi^0 = 1_{\Sigma_E}$; $\Phi^{-t} = \Phi^t_H$ $(s,t \in \mathbb{R})$.

With this in mind, one can consider the evolution of any smooth function (or observable) $F : \Gamma \times \mathbb{R} \to \mathbb{R} : (q,p,t) \mapsto F(q,p,t)$ along the flow of the Hamilton equations, namely

$$\begin{aligned} \dot{F} &:= \frac{\mathrm{d}}{\mathrm{d}t} F(q(t),p(t),t) = \frac{\partial F}{\partial q}\cdot \dot{q} + \frac{\partial F}{\partial p}\cdot \dot{p} + \frac{\partial F}{\partial t} = \\ &= \frac{\partial F}{\partial q}\cdot\frac{\partial H}{\partial p} - \frac{\partial H}{\partial q}\cdot\frac{\partial F}{\partial p} + \frac{\partial F}{\partial t} := \{F,H\} + \frac{\partial F}{\partial t}\ . \end{aligned} \tag{2.15}$$

Such a formula holds even if the flow is not globally defined. In the last step of (2.15) the *Poisson (P) bracket* $\{F,G\}$ of two functions F and G defined on Γ (possibly depending on time explicitly) has been defined:

$$\{F,G\} = \frac{\partial F}{\partial q}\cdot\frac{\partial G}{\partial p} - \frac{\partial G}{\partial q}\cdot\frac{\partial F}{\partial p} = \sum_{i=1}^{n}\left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i}\frac{\partial F}{\partial p_i}\right)\ . \tag{2.16}$$

The P-bracket is a function defined on Γ and one can check by direct inspection that the following properties hold:

1. $\{F,G\} = -\{G,F\}$ (skew-symmetry);
2. $\{aF + bG, H\} = a\{F,G\} + b\{G,H\}$ (left-linearity);
3. $\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} \equiv 0$ (Jacobi identity);
4. $\{FG,H\} = F\{G,H\} + \{F,H\}G$ (Leibniz rule),

[1] We recall that a set G is a group if a binary application, or operation $\circ : G \times G \to G$ is defined on it, such that:

1. $f \circ (g \circ h) = (f \circ g) \circ h$ for all triples $f,g,h \in G$ (associativity);
2. there exists an element $e \in G$ such that $g \circ e = e \circ g = g$ for any $g \in G$ (unit element);
3. for any $g \in G$ there exists $f := g^{-1} \in G$ such that $f \circ g = g \circ f = e$ (inverse element).

Properly speaking, the group is the pair $(G,\circ)$. If for any $f,g \in G$ $f \circ g = g \circ f$ then the $(G,\circ)$ is said to be a commutative or Abelian group.

for any triple of functions F, G, H and any pair of real numbers a, b.

Remark 2.3. *The algebra of functions defined on the phase space, endowed with the operation $\{\ ,\}$ has the structure of a Lie algebra*[2] *(properties 1., 2., and 3.) of Leibniz type (property 4.). A Lie-Leibniz algebra is called a Poisson (P) algebra.*

The solution of equation (2.15) is clearly given by $F(\Phi_H^t(q_0, p_0), t)$. The Hamilton equations are a particular case of equation (2.15) when one considers $F = q_i$ or $F = p_i$. For such a reason, the evolution equation (2.15), defines the class of Hamiltonian dynamical systems on the given phase space Γ (notice that H is one of the possible functions defined on Γ: the privileged role played by the Hamiltonian comes only from physics).

Remark 2.4. *Observe that the conservation of energy for an autonomous system follows from equation (2.15) and by the skew-symmetry of the P-bracket: $\dot{H} = \{H, H\} = 0$.*

2.3.2 Symplectic structure

The Hamilton equations (2.4) can be rewritten in compact form by defining $x = (q, p)$ and $\nabla_x H(x) = (\partial H/\partial q, \partial H/\partial p)$. One then gets

$$\dot{x} = J_{2n} \nabla_x H(x) \ , \tag{2.17}$$

where the $2n \times 2n$ *standard symplectic matrix*

$$J_{2n} := \begin{pmatrix} \mathbb{O}_n & \mathbb{I}_n \\ -\mathbb{I}_n & \mathbb{O}_n \end{pmatrix} \tag{2.18}$$

has been defined. The matrix J_{2n} is immediately checked to satisfy the properties $J_{2n}^T = J_{2n}^{-1} = -J_{2n}$, $J_{2n}^2 = -\mathbb{I}_{2n}$, the latter implying $|\det J_{2n}| = 1$; actually $\det J_{2n} = 1$, as can be easily proven[3].

The matrix J_{2n} is called standard symplectic matrix for the following reason. Consider the skew-symmetric bilinear form $\langle x, J_{2n} y\rangle$, and look for the linear transformations $x \mapsto Sx$, $y \mapsto Sy$ that leave such a form invariant. One gets $S^T J_{2n} S = J_{2n}$, whose solutions S define the so-called symplectic group $Sp(2n, \mathbb{R})$ of $2n \times 2n$ real matrices. One easily checks that $S, S' \in Sp(2n, \mathbb{R})$ implies $SS' \in Sp(2n, \mathbb{R})$, and the matrix product is associative; $\mathbb{I}_{2n} \in Sp(2n, \mathbb{R})$, so that a unit element exists. Moreover, $S \in Sp(2n, \mathbb{R})$ is non singular and, multiplying $S^T J_{2n} S = J_{2n}$ by S^{-1} from the right and by S^{-T} from the left one verifies that $S^{-1} \in Sp(2n, \mathbb{R}$, so that the inverse element exists. This shows that $Sp(2n, \mathbb{R})$ is a group (with respect to the matrix product). A further property that is easily checked is that $S \in Sp(2n, \mathbb{R})$ implies $S^T \in Sp(2n, \mathbb{R})$. Obviously, $J_{2n} \in Sp(2n, \mathbb{R})$.

---

[2] A Lie algebra $L = (V, [\ ,])$ is a vector space V endowed with a "product" $[\ ,] : V \times V \to V$ that is skew-symmetric, bilinear and Jacobi; here V is the vector space of functions on Γ and $[\ ,] = \{\ ,\}$.

[3] Consider the eigenvalue problem $J_{2n}\xi = \lambda\xi$; $\lambda \neq 0$ since $\det J_{2n} \neq 0$. Setting $\xi = (u, v)$ one gets $v = \lambda u$ and $-u = \lambda v$, so that $u = 0$ iff $v = 0$, and $(1 + \lambda^2)u = 0$. Thus $\lambda = \pm\imath$, $\imath$ denoting the imaginary unit. Since J_{2n} is real, the eigenvalues come in complex conjugate pairs, and the product of them is one.

Given a function $F(x)$, one associates to it the *Hamiltonian vector field* $X_F(x) := J_{2n}\nabla_x F(x)$, motivated by the Hamilton equations $\dot{x} = X_H(x)$. One then easily checks that the P-bracket reads

$$\{F,G\} = \langle \nabla_x F, J_{2n}\nabla_x G\rangle = \langle X_F, J_{2n}X_G\rangle \ . \tag{2.19}$$

2.4 Canonical transformations

Anatural question is the following. As stressed above, the L-equations are left invariant in form by any point transformation $q \mapsto Q(q,t)$. Which are the transformations that leave invariant in form the H-equations (2.4)? In order to see that something richer and nontrivial happens in the Hamiltonian formalism, we first consider once again the relevant example of the charged particle in a given e.m. field.

Example 2.3. *Consider the Hamiltonian (2.6), write down explicitly the corresponding H-equations and show that they are not invariant under the gauge transformation of the e.m. (1.15). Show that, upon defining a new momentum $P = p - (q/c)\nabla\chi$ while retaining the same coordinates, i.e. $X = x$, the H-equations do not change in form. Observe however, that to such a transformation there has to correspond the cancellation of the term $-\partial F/\partial t$ from the Hamiltonian (2.7).*

Definition 2.2. *A change of variables, or transformation*

$$(q,p) \mapsto (Q,P) = (V(q,p,t), U(Q,P,t)) \iff (Q,P) \mapsto (q,p) = (v(Q,P,t), u(Q,P,t)) \tag{2.20}$$

is said do be canonical *if to* any *Hamiltonian $H(q,p,t)$ and its H-equations it associates a Hamiltonian $K(Q,P,t)$ and its H-equations.*

The latter definition implies that a notation $(q,p,H) \mapsto (Q,P,K)$ is more appropriate to specify a canonical transformation. Observe that K is not unique, in general (one can alway add a constant or even a function of time to K).

The shortest way to characterize the canonical transformations makes use of the Hamilton second variational principle. Indeed, we recall in short that the H-equations corresponding to a Hamiltonian $H(q,p,t)$ characterize the critical points of the action $A_H = \int_{t_1}^{t_2}(p\cdot\dot{q} - H)\mathrm{d}t$. On the other hand, also the H-equations corresponding to $K(Q,P,t)$ characterize the critical points of the action $A_K = \int_{t_1}^{t_2}(P\cdot\dot{Q} - K)\mathrm{d}t$. Taking into account that in the mentioned variational principle the both the coordinate curves $q(t)$ and $Q(t)$ have fixed ends, one easily realizes that in order to ensure the validity of the H-equations in both the old and the new coordinates, there have to exist a function $F(q,Q,t)$ and a constant $c \neq 0$ such that

$$A_H = c\ A_K + \Delta F \ , \tag{2.21}$$

where $\Delta F := \int_{t_1}^{t_2} \dot{F}\mathrm{d}t = F(q_2,Q_2,t_2) - F(q_1,Q_1,t_1)$. Explicitly, relation (2.21) reads

$$\underbrace{\int_{t_1}^{t_2} [p\cdot\dot{q} - H(q,p,t)]\,\mathrm{d}t}_{A_H} = c\underbrace{\int_{t_1}^{t_2}\left[P\cdot\dot{Q} - K(Q,P,t)\right]\mathrm{d}Tt}_{A_K} + \underbrace{\int_{t_1}^{t_2} F(q(t),Q(t),t)\mathrm{d}t}_{\Delta F} \ . \tag{2.22}$$

Proposition 2.3. *The transformation $(q, p, H) \mapsto (Q, P, K)$ defined by F and $c(\neq 0)$, satisfying (2.21), maps Hamilton equations into Hamilton equations.*

$\triangleleft$ PROOF. $\delta A_H = c\delta A_K + \delta\Delta F$, and the last term vanishes, so that $\delta A_H = 0$ iff $\mathrm{d}A_K = 0$. According to Proposition 2.2, the condition $\delta A_{H/K} = 0$ are equivalent to the Hamilton equations. $\triangleright$

As a matter of jargon, the function $F(q, Q, t)$ and the constant c appearing on the right hand side of (2.21) are called the *generating function* and the *valence* of the transformation.

The differential form of degree one $\delta\omega := p \cdot \mathrm{d}q - H\mathrm{d}t$ is the so-called Poincaré-Cartan 1-form. Denoting by $\delta\Omega := P \cdot \mathrm{d}Q - K\mathrm{d}t$ the transformed 1-form, due to the arbitrariness of the integration interval $[t_1, t_2]$, (2.21) is equivalent to the closure condition $\delta\omega - c\delta\Omega = \mathrm{d}F$, i.e.

$$\mathrm{d}F(q, Q, t) = p \cdot \mathrm{d}q - cP \cdot \mathrm{d}Q + (cK - H)\mathrm{d}t \ , \tag{2.23}$$

The latter relation implies

$$\frac{\partial F}{\partial q} = p \ \ ; \ \ \frac{\partial F}{\partial Q} = -cP \ \ ; \ \ \frac{\partial F}{\partial t} = cK - H \ . \tag{2.24}$$

The canonical transformation defined in this way is given implicitly, since F depends on both the old and the new coordinates. In order to determine it explicitly, one has to make the further hypothesis

$$\det\left(\frac{\partial^2 F}{\partial q \partial Q}\right) \neq 0 \ , \tag{2.25}$$

which allows (by the implicit function theorem) to invert either the first or the second of relations (2.24). Indeed, starting from the first of (2.24) one gets $Q = V(q, p, t)$ which, substituted in the second relation yields $P = -c^{-1}\partial F/\partial Q(q, V, t) = U(q, p, t)$. On the other hand, starting from the second relation one gets $q = v(Q, P, t)$ and substituting it in the first one yields $p = \partial F/\partial q(v, Q, t) = u(Q, P, t)$. Using the latter expressions and substituting them in the third of (2.24) yields the new Hamiltonian $K(Q, P, t) = c^{-1}[H(v, u, t) + \partial F/\partial t(v, Q, t)]$.

Very often one needs to generate canonical transformations by means of a function of q and P, for example. This is easily realized starting from (2.23) and defining the generating function $S(q, P, t) := F(q, Q, t) + cQ \cdot P$, satisfying

$$\mathrm{d}S = p \cdot \mathrm{d}q + cQ \cdot \mathrm{d}P + (cK - H)\mathrm{d}t \ , \tag{2.26}$$

which implies

$$\frac{\partial S}{\partial q} = p \ \ ; \ \ \frac{\partial S}{\partial P} = cQ \ \ ; \ \ \frac{\partial S}{\partial t} = cK - H \ . \tag{2.27}$$

The canonical transformation generated by S is explicitly determined under the condition

$$\det\left(\frac{\partial^2 S}{\partial q \partial P}\right) \neq 0 \ , \tag{2.28}$$

with reasonings similar to those made above for F.

Exercise 2.3. *Consider the Hamiltonian $H'(q,p,t)$, defined in (2.7) and corresponding to the a gauge shifted Lagrangian. Show that the time-dependent transformation $(q,p) \mapsto (Q,P)$ defined by $P = (p - \nabla F)/a$, $Q = aq$, is a canonical transformation, generated by $S(q,P,t) = aq \cdot P + F(q,t) + \psi(t)$. Show that the new Hamiltonian is*

$$K(Q,P,t) = H'(Q,P,t) + \frac{\partial F}{\partial t} + \dot{\psi} = aH(Q,P,t) \ .$$

Apply all this to the case of the particle in the e.m. field.

Exercise 2.4. *Given a time-dependent point transformation $q \mapsto Q = V(q,t)$, show that this is canonically complemented by*

$$P = \left(\frac{\partial V}{\partial q}\right)^{-T} (p - \nabla\varphi(q,t)) \ \ ,$$

where $\varphi(q,t)$ is an arbitrary function of its arguments. Show that the transformation is generated by $S(q,P,t) = V \cdot P + \varphi$; write down the new Hamiltonian.

Exercise 2.5. *Repeat the previous exercise complementing the point transformation on the momenta, namely $p \mapsto P = U(p,t)$. Hint: look for a generating function of the kind $F'(p,Q,t) = F(q,Q,t) + q \cdot p$, where F is determined by (2.23); as an alternative, invert first $P = U(p,t)$ and then look for $S(q,P,t)$.*

2.5 Hamilton-Jacobi equation

Suppose that one looks for a canonical change of variables $(q,p,H,t) \mapsto (Q,P,K,t)$ with valence $c = 1$ and such that $K = \varphi(t)$ depends (at most) on time only, i.e. independent of Q and P. This amounts to look for a canonical transformation such that the new Hamiltonian variables do not evolve in time. Notice that one can set $\varphi(t) \equiv 0$ without any loss of generality. Indeed, if S satisfies (2.27) with $K = \varphi(t)$, then $\tilde{S} := S - \int \varphi(t)\mathrm{d}t$ satisfies (2.27) with $K \equiv 0$. With this in mind, the first and the third of relations (2.27) yield the *Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0 \ , \tag{2.29}$$

a first order PDE in the unknown function $S(q,t)$. Notice that, among the possible solutions of equation (2.29), we are interested to the so-called *complete integrals*, namely those solutions depending on n parameters $P_1, \ldots, P_n$ and such that (2.28) holds.

A complete integral of the Hamilton-Jacobi equation generates a canonical transformation $(q,p,H,t) \mapsto (Q,P,0,t)$, since condition (2.28), together with the first two equations of (2.27), allows to get (q,p) in terms of (Q,P) and t and viceversa. More precisely, starting from $Q = \partial S/\partial P$, by (2.28) one can invert it and get $q = v(Q,P,t)$; the latter can then be inserted into the right hand side of $p = \partial S/\partial q$, which yields $p = u(Q,P,t)$. By construction, the functions of time $q(t)$ and $p(t)$ solve the H-equations with Hamiltonian H. On the other hand, one can start from $p = \partial S/\partial q$ and, again by (2.28), invert it to get $P = U(q, p, t)$; upon substitution of the latter into the right hand side of $Q = \partial S/\partial P$ one gets $Q = V(q, p, t)$. By construction, Q and P thus obtained are preserved by the flow of the original Hamiltonian. In particular, by hypothesis, the functions U_i are n independent first integrals of H:

$$\dot{P}_i = \{U_i, H\} + \frac{\partial U_i}{\partial t} = 0 \quad i = 1, \ldots, n \ . \tag{2.30}$$

Moreover, the functions U_i are new momenta and as a consequence $\{U_i, U_j\} = \{P_i, P_j\} = 0$.

Very often H is independent of time, and one looks for time-independent canonical transformations such that the new Hamiltonian K depends on the momenta P only, since in the latter case the Hamilton equations are immediately solved: $Q(t) = Q(0) + t(\partial K/\partial P)$, at constant P (such canonical transformations rectify the flow of the given Hamiltonian system). In this case, the following time-independent version of the Hamilton-Jacobi equation holds:

$$H\left(q, \frac{\partial S}{\partial q}(q, P)\right) = K(P) \ . \tag{2.31}$$

Here again, a complete integral of the above equation is required. Notice that the new Hamiltonian K is an unknown of the problem.

2.6 Integrability: Liouville theorem

A dynamical system is integrable if it possesses a number of first integrals (i.e. functions defined on the phase space not evolving in time along the flow of the system) which is high enough to geometrically constraint the motion, a priori, on a curve. For a generic system of the form $\dot{x} = u(x)$ in $\mathbb{R}^n$, integrability would require, a priori, $n-1$ first integrals (the intersection of the level sets of m first integrals has co-dimension m and dimension $n-m$). However, it turns out that the Hamiltonian structure reduces such a number to half the (even) dimension of the phase space.

Definition 2.3. *The system defined by the Hamiltonian $H(q, p, t)$, is said to be integrable in $\Gamma \subseteq \mathbb{R}^{2n}$, in the sense of Liouville, if it admits n independent first integrals $f_1(q, p, t), \ldots, f_n(q, p, t)$ in involution, i.e., for any $(q, p) \in \Gamma$ and $t \in \mathbb{R}$*

1. *$\partial f_j/\partial t + \{f_j, H\} = 0$ for any $j = 1, \ldots, n$;*

2. *$\sum_{j=1}^n c_j \nabla f_j(q, p, t) = 0 \Rightarrow c_1 = \cdots = c_n = 0$ (equivalently: the rectangular matrix of the gradients of the integrals has maximal rank n) for any (q, p, t);*

3. *$\{f_j, f_k\} = 0$ for any $j, k = 1, \ldots, n$.*

Notice that often H coincides with one of the first integrals. The introduction of the above definition is motivated by the following theorem.

Theorem 2.1 (Liouville). *Let the Hamiltonian system defined by H be Liouville-integrable in $\Gamma \subseteq \mathbb{R}^{2n}$, and let $a \in \mathbb{R}^n$ such that the level set*

$$M_a := \{(q,p) \in \Gamma \ : \ f_1(q,p,t) = a_1, \dots, f_n(q,p,t) = a_n\}$$

is non empty; let also M'_a denote a (nonempty) connected component of M_a. Then, a function $S(q,t;a)$ exists, such that $p \cdot \mathrm{d}q|_{M'_a} = d_q S(q,t;a)$ and S is a complete integral of the time-dependent Hamilton-Jacobi equation, i.e. the generating function of a canonical transformation $\mathscr{C} : (q,p,H,t) \mapsto (b,a,0,t)$, where $b := \partial S/\partial a$.

Remark 2.5. *If $H(q,p)$ does not depend explicitly on time, then in the above definition of integrable system all the f_j are independent of time as well, and condition* 1. *is replaced by $\{f_j, H\} = 0$. In such a case, the generating function $S(q;a)$ appearing in the Liouville theorem is a complete integral of the time-independent Hamilton-Jacobi equation $H(q, \nabla_a S) = K(a)$, thus generating a canonical transformation $\mathscr{C} : (q,p) \mapsto (b,a)$ such that $H(\mathscr{C}^{-1}(b,a)) = K(a)$.*

$\lhd$ PROOF. In order to prove the theorem 2.1, understand the meaning of the definition 2.3, we start by supposing that $H(q,p,t)$ admits n independent first integrals $f_1(q,p,t), \dots, f_n(q,p,t)$, but we do not suppose, for the moment, that such first integrals are in involution. Without any loss of generality, as a condition of independence of the first integrals one can assume

$$\det\left(\frac{\partial f}{\partial p}\right) = \det\left(\frac{\partial(f_1,\dots,f_n)}{\partial(p_1,\dots,p_n)}\right) \neq 0 \ ,$$

in such a way that the level set $M_a = \{(q,p) : f(q,p,t) = a\}$ of the first integrals can be represented, by means of the implicit function theorem, as

$$p_1 = u_1(q,t;a) \ ; \ \dots \ p_n = u_n(q,t;a) \ . \tag{2.32}$$

The above relations must hold at any time if they hold at $t = 0$. Differentiating the relation $p_i(t) = u_i(q(t),t;a)$ $(i = 1,\dots,n)$ with respect to time and using the Hamilton equations one gets

$$\frac{\partial u_i}{\partial t} + \sum_{j=1}^{n}\left(\frac{\partial u_i}{\partial q_j} - \frac{\partial u_j}{\partial q_i}\right)\frac{\partial H}{\partial p_j} = -\frac{\partial H}{\partial q_i} - \sum_{j=1}^{n}\frac{\partial u_j}{\partial q_i}\frac{\partial H}{\partial p_j}\Bigg|_{p=u(q,t;a)} \ . \tag{2.33}$$

Notice that, for the sake of convenience, the same sum of terms is artificially added on both sides of the equation. By introducing the quantities

$$\mathrm{rot}(u) := \left(\frac{\partial u}{\partial q}\right) - \left(\frac{\partial u}{\partial q}\right)^T \ , \tag{2.34}$$

$$v(q,t) := \frac{\partial H}{\partial p}\Bigg|_{p=u(q,t;a)} \ , \tag{2.35}$$

and

$$\widetilde{H}(q,t) := H(q,u(q,t;a),t) \ , \tag{2.36}$$

the equations (2.33) can be rewritten in compact, vector form as

$$\frac{\partial u}{\partial t} + \mathrm{rot}(u)v = -\nabla_q \widetilde{H} \ . \tag{2.37}$$

Notice the similarity of the latter equation with the (unitary density) Euler equation of hydrodynamics, namely

$$\frac{\partial \boldsymbol{u}}{\partial t} + \mathrm{rot}(\boldsymbol{u})\boldsymbol{u} = -\nabla\left(\frac{|\boldsymbol{u}|^2}{2} + p\right) \ , \tag{2.38}$$

where $\boldsymbol{u}$ is the velocity field, p is the pressure and $\mathrm{rot}(\boldsymbol{u})\boldsymbol{u} = \boldsymbol{\omega} \wedge \boldsymbol{u}$, $\boldsymbol{\omega} := \nabla \wedge \boldsymbol{u}$ being the vorticity of the fluid. The similarity of (2.37) and (2.38) is completely evident in the case of natural mechanical systems, whose Hamiltonian has the form

$$H(q,p,t) = \frac{p \cdot M^{-1}(q,t)p}{2} + V(q,t) \ ,$$

where $M^{-1}(q,t)$ is a $n \times n$ positive definite matrix. In such a case $v = M_{-1}u$ and equation (2.37) takes the rather simple form

$$\frac{\partial u}{\partial t} + \mathrm{rot}(u)M^{-1}u = -\nabla_q\left(\frac{u \cdot M^{-1}u}{2} + V\right) \ . \tag{2.39}$$

In particular, in those cases such that $G = I_n$ the latter equation is the Euler equation in space dimension n, with the potential energy V playing the role of pressure.

Remark 2.6. *Attention has to be paid to the fact that for the Euler equation (2.38) the pressure p is determined by the divergence-free condition $\nabla \cdot \boldsymbol{u} = 0$, while nothing similar holds, in general, for the equations (2.37) or (2.39).*

Now, by analogy with the case of fluids, we look for curl-free, i.e. irrotational solutions of the Euler-like equation (2.37) (we recall that in fluid dynamics, looking for a solution of the Euler equation (2.38) of the form $\boldsymbol{u} = \nabla\phi$ leads to the Bernoulli equation for the velocity potential ϕ, namely $\partial\phi/\partial t + |\nabla\phi|^2/2 + p = \mathit{constant}$). In simply connected domains (of the n-dimensional configuration space), one has

$$\mathrm{rot}(u) = 0 \ \text{ iff } \ u = \nabla S \ ,$$

where $S = S(q,t;a)$. Upon substitution of $u = \nabla S$ into equation (2.37) and lifting a gradient, one gets

$$\frac{\partial S}{\partial t} + H(q, \nabla_q S, t) = \varphi(t;a) \ . \tag{2.40}$$

One can set $\varphi(t;a) \equiv 0$ without any loss of generality, and the latter equation becomes the time-dependent Hamilton-Jacobi equation (if $\varphi \not\equiv 0$ then $\widetilde{S} := S - \int \varphi dt$ satisfies equation (2.40) with zero right hand side). Thus, The Hamilton-Jacobi equation is the analogue of the Bernoulli equation for the hydrodynamics of Hamiltonian systems. The interesting point is that the curl-free condition $\mathrm{rot}(u) = 0$ is equivalent to the condition of involution of the first integrals $f_1, \ldots, f_n$. Indeed, starting from the identity

$$f_i(q, u(q, t; a), t) \equiv a_i \ , \tag{2.41}$$

and taking its derivative with respect to q_j one gets

$$\frac{\partial f_i}{\partial q_s} + \sum_{r=1}^{n} \frac{\partial f_i}{\partial p_r}\frac{\partial u_r}{\partial q_s} = 0$$

for any $i = 1, \ldots, n$. Thus

$$\begin{aligned} \{f_i, f_j\} &= \sum_{s=1}^{n} \left(\frac{\partial f_i}{\partial q_s}\frac{\partial f_j}{\partial p_s} - \frac{\partial f_i}{\partial p_s}\frac{\partial f_j}{\partial q_s} \right) = \sum_{r,s=1}^{n} \left(\frac{\partial f_i}{\partial p_s}\frac{\partial f_j}{\partial p_r}\frac{\partial u_r}{\partial q_s} - \frac{\partial f_j}{\partial p_s}\frac{\partial f_i}{\partial p_r}\frac{\partial u_r}{\partial q_s} \right) = \\ &= \sum_{r,s=1}^{n} \frac{\partial f_j}{\partial p_r} \left(\frac{\partial u_r}{\partial q_s} - \frac{\partial u_s}{\partial q_r} \right) \frac{\partial f_i}{\partial p_s} = \left[\left(\frac{\partial f}{\partial p} \right) \mathrm{rot}(u) \left(\frac{\partial f}{\partial p} \right)^T \right]_{ji} \ , \end{aligned}$$

which implies $\mathrm{rot}(u) = 0$ iff $\{f_i, f_j\} = 0$ for any $i, j = 1, \ldots, n$ (the direct implication is obvious, the reverse one requires the independence condition $\det(\partial f/\partial p) \neq 0$). This is the key point: the condition of involution of the first integrals is equivalent to that of irrotational, i.e. gradient, velocity fields of the hydrodynamic equation (2.37). The velocity potential $S(q, t; a)$ satisfies the Hamilton-Jacobi equation and is actually a complete integral of the latter. In order to see this, one can start again from identity (2.41), setting there $u = \nabla S$ and taking the derivative with respect to a_j, getting the i, j component of the matrix identity

$$\left(\frac{\partial f}{\partial p} \right) \left(\frac{\partial^2 S}{\partial q \partial a} \right) = I_n \ ,$$

which, by the independence condition of the first integrals, yields $\det(\partial^2 S/\partial q \partial a) \neq 0$. We finally notice that if the first integrals and thus the velocity field u are known, then the potential S can be obtained by a simple integration, based on the identity $\mathrm{d}_q S = u \cdot \mathrm{d}q$, such as

$$S(q, t; a) - S(0, t; a) = \int_{0 \to q} u(q', t; a) \cdot \mathrm{d}q' = \int_0^1 u(\lambda q, t; a) \cdot q \mathrm{d}\lambda \ ,$$

where $S(0, t; a)$ may be set to zero. The function $S(q, t; a)$, satisfying the Hamilton-Jacobi equation, generates a canonical transformation $(q, p, H, t) \mapsto (b, a, 0, t)$ to a zero Hamiltonian, transformation defined by the implicit equations $p = \nabla_q S(q, t; a)$, $b := \nabla_a S(q, t; a)$. This concludes the proof of theorem 2.1. The restriction to the case where H, $f_1, \ldots, f_n$ are independent of time is left as an exercise. $\triangleright$

Example 2.4. *The Hamiltonian system of central motions is Liouville-integrable. Indeed, if* $H = \frac{|\boldsymbol{p}|^2}{2m} + V(|\boldsymbol{r}|)$ *is the Hamiltonian of the system, then it is easily proven that the angular momentum* $\boldsymbol{L} = \boldsymbol{r} \wedge \boldsymbol{p}$ *is a vector constant of motion (the Hamiltonian is invariant with respect* *to the "canonical rotations" $(\boldsymbol{r}, \boldsymbol{p}) \mapsto (\boldsymbol{r}', \boldsymbol{p}') = (R\boldsymbol{r}, R\boldsymbol{p})$, where R is any orthogonal matrix; the conservation of the angular momentum is a consequence of the Nöther theorem). The phase space of the system has dimension $2n = 6$, and three independent first integrals in involution are $f_1 := H$, $f_2 := |\boldsymbol{L}|^2$ and $f_3 := L_z$, for example (show that).*

Example 2.5. *The Hamiltonian of n noninteracting systems, $H = \sum_{j=1}^{n} h_j(q_j, p_j)$, is obviously Liouville integrable, with the choice $f_j := h_j(q_j, p_j)$, $j = 1, \dots, n$. As an example, consider the case of harmonic oscillators, where $h_j(q_j, p_j) = (p_j^2 + \omega_j^2 q_j^2)/2$.*

A fundamental result in the theory of integrable systems is the following theorem due to Arnol'd, whose proof is not reported.

Theorem 2.2 (Arnol'd)**.** *Let the Hamiltonian system defined by H be integrable in $\Gamma \subseteq \mathbb{R}^{2n}$ in the sense of Liouville, and let $a \in \mathbb{R}^n$ such that the level set*

$$M_a := \{(q, p) \in \Gamma \ : \ f_1(q, p) = a_1, \dots, f_n(q, p) = a_n\}$$

is non empty; let also M_a' denote a connected and <u>compact</u> component of M_a. Then M_a' is diffeomorphic to the n-dimensional torus $\mathbb{T}^n = \mathbb{T}^1 \times \cdots \times \mathbb{T}^1$ (n times), where $\mathbb{T}^1 = \mathbb{R}/(2\pi\mathbb{Z})$, the (group of) real numbers modulo 2π. Moreover, there exists a neighborhood U of M_a' in Γ that is canonically diffeomorphic to $\mathbb{T}^n \times B$, where $B \subset \mathbb{R}^n_+$, i.e. there exists a canonical transformation $\mathscr{C} : U \to \mathbb{T}^n \times B : (q, p) \mapsto (\varphi, I)$ to angle-action variables, such that $H(\mathscr{C}^{-1}(\varphi, I)) = E(I)$ and $f_j(\mathscr{C}^{-1}(q, p)) = \Phi_j(I)$ for any $j = 1, \dots, n$.

Thus, for Liouville-integrable Hamiltonian systems displaying compact families of level sets, canonical action-angle coordinates (φ, I) can be introduced, such that both the Hamiltonian and all the first integrals depend on the action variables I only. In terms of the variables (φ, I), the dynamics of the system becomes trivial: the Hamilton equations $\dot{\varphi} = \partial E/\partial I$, $\dot{I} = -\partial E/\partial \varphi = 0$ imply that $I(t) = I(0)$ and $\varphi(t) = \varphi(0) + \omega(I(0))t$, where

$$\omega(I) := \frac{\partial E(I)}{\partial I} \ . \tag{2.42}$$

The phase space of the system is thus locally foliated into invariant tori, on each of which the motion is a translation with a frequency vector (2.42) depending, in general, on the value of the action I_0 labeling the torus $\mathbb{T}^n$.

Chapter 3

General Hamiltonian systems

3.1 Poisson structures

The evolution in time of physical systems is described by differential equations of the form

$$\dot{x} = u(x) \ , \tag{3.1}$$

where the vector field $u(x)$ is defined on some *phase space* Γ, i.e. the space of all the possible states of the system.

Remark 3.1. *From a topological point of view, the phase space* Γ *of the system has to be a Banach (i.e. normed complete vector) space. This is due to the necessity to guarantee the existence and uniqueness of the solution* $x(t) = \Phi^t(x_0)$ *to the differential equation (3.1), for any initial condition* $x(0) = x_0 \in \Gamma$ *and any* $t \in I_0 \subseteq \mathbb{R}$*.*

Very often, Γ turns out to be a Hilbert space, i.e. a Euclidean space that is Banach with respect to the norm induced by the scalar product. This happens obviously if Γ has finite dimension or, for example, in the theory of the classical (linear and nonlinear) PDEs, such as the wave and the heat equations, and in quantum mechanics. In this case equation (3.1) can be written by components, namely $\dot{x}_i = u_i(x_1, x_2, \dots)$, $i \in \mathbb{N}$, where the x_i's are called the components, coordinates or Fourier coefficients of x in a given orthonormal basis $\{\varphi_i\}_{i\in\mathbb{N}}$ of the Hilbert space: $x = \sum_i x_i\varphi_i$, $x_i := \langle \varphi_i, x\rangle$ (where $\langle , \rangle$ denotes the scalar product of the Hilbert space). In the sequel, we will almost always suppose the phase space of the system to be a Hilbert space.

Hamiltonian systems are those particular dynamical systems whose phase space Γ is endowed with a Poisson structure, according to the following definition.

Definition 3.1 (Poisson bracket)**.** *Let $\mathscr{A}(\Gamma)$ be the algebra (i.e. vector space with a bilinear product) of real smooth (i.e. C^∞) functions defined on Γ. A function $\{\ ,\}:\mathscr{A}(\Gamma)\times\mathscr{A}(\Gamma)\to\mathscr{A}(\Gamma)$ is called a* Poisson bracket *on Γ if it satisfies the following properties:*

1. $\{F,G\}=-\{G,F\}\ \forall F,G\in\mathscr{A}(\Gamma)$ *(skew-symmetry);*
2. $\{\alpha F+\beta G,H\}=\alpha\{F,H\}+\beta\{G,H\}\ \forall\alpha,\beta\in\mathbb{R}$ *and* $\forall F,G,H\in\mathscr{A}(\Gamma)$ *(left linearity);*
3. $\{F,\{G,H\}\}+\{G,\{H,F\}\}+\{H,\{F,G\}\}=0\ \forall F,G,H\in\mathscr{A}(\Gamma)$ *(Jacobi identity);*
4. $\{FG,H\}=F\{G,H\}+\{F,H\}G\ \forall F,G,H\in\mathscr{A}(\Gamma)$ *(left Leibniz rule).*

The pair $(\mathscr{A}(\Gamma),\{\ ,\})$ is a Poisson algebra, i.e. a Lie algebra (a vector space with a skew-symmetric, left-linear and Jacobi product) satisfying the Leibniz rule. Observe that 1. and 2. in the above definition imply right-linearity, so that the Poisson bracket is actually bi-linear. Observe also that 1. and 4. imply the right Leibniz rule.

Definition 3.2 (Hamiltonian system)**.** *Given a Poisson algebra $(\mathscr{A}(\Gamma),\{\ ,\})$, a Hamiltonian system on Γ is a dynamical system described by a differential equation of the form (3.1) whose vector field has the form*

$$u_i(x)=[X_H(x)]_i:=\{x_i,H\}\ .$$

where $H\in\mathscr{A}(\Gamma)$ is called the Hamiltonian of the system.

In local coordinates, a skew-symmetric, bi-linear Leibniz bracket $\{\ ,\}$ on Γ is of the form

$$\{F,G\}=\nabla F\cdot J\nabla G:=\sum_{j,k}\left(\frac{\partial F}{\partial x_j}\right)J_{jk}(x)\left(\frac{\partial G}{\partial x_k}\right)\ ,\tag{3.2}$$

for all $F,G\in\mathscr{A}(\Gamma)$, where

$$J_{jk}(x):=\{x_j,x_k\}\ .\tag{3.3}$$

The following proposition holds:

Proposition 3.1. *The bracket (3.2) is Jacobi, i.e. is a Poisson bracket, iff the operator function $J(x)$ defining it satisfies the relation:*

$$\boxed{\sum_s\left(J_{is}\frac{\partial J_{jk}}{\partial x_s}+J_{js}\frac{\partial J_{ki}}{\partial x_s}+J_{ks}\frac{\partial J_{ij}}{\partial x_s}\right)=0}\tag{3.4}$$

for any $x\in\Gamma$ and any i,j,k.

$\triangleleft$ PROOF. The condition (3.4) is checked by a direct computation. One has

$$\begin{aligned}\{F,\{G,H\}\} &= \nabla F\cdot J\nabla(\nabla G\cdot J\nabla H)=\sum_{isjk}\frac{\partial F}{\partial x_i}J_{is}\frac{\partial}{\partial x_s}\left(\frac{\partial G}{\partial x_j}J_{jk}\frac{\partial H}{\partial x_k}\right)=\\ &= \sum_{isjk}\left[\frac{\partial F}{\partial x_i}J_{is}\frac{\partial^2 G}{\partial x_s\partial x_j}J_{jk}\frac{\partial H}{\partial x_k}+\frac{\partial F}{\partial x_i}J_{is}\frac{\partial^2 H}{\partial x_s\partial x_k}J_{jk}\frac{\partial G}{\partial x_j}+\frac{\partial F}{\partial x_i}\frac{\partial G}{\partial x_j}\frac{\partial H}{\partial x_k}J_{is}\frac{\partial J_{jk}}{\partial x_s}\right]=\\ &= \nabla F\cdot\left(J\frac{\partial^2 G}{\partial x^2}J\right)\nabla H-\nabla F\cdot\left(J\frac{\partial^2 H}{\partial x^2}J\right)\nabla G+\sum_{isjk}\frac{\partial F}{\partial x_i}\frac{\partial G}{\partial x_j}\frac{\partial H}{\partial x_k}J_{is}\frac{\partial J_{jk}}{\partial x_s}\ .\end{aligned}$$

Now, exploiting the skew-symmetry of J and the consequent symmetry of a matrix of the form $J(\text{hessian})J$, and suitably cycling over the functions F, G, H and over the indices i, j, k, one gets

$$\{F,\{G,H\}\}+\{G,\{H,F\}\}+\{H,\{F,G\}\} = \sum_{ijk} \frac{\partial F}{\partial x_i}\frac{\partial G}{\partial x_j}\frac{\partial H}{\partial x_k}\left[\sum_s \left(J_{is}\frac{\partial J_{jk}}{\partial x_s} + J_{js}\frac{\partial J_{ki}}{\partial x_s} + J_{ks}\frac{\partial J_{ij}}{\partial x_s}\right)\right] .$$

Such an expression is identically zero for all $F, G, H \in \mathscr{A}(\Gamma)$ iff (3.4) holds. Indeed, sufficiency (if) is obvious, whereas necessity (only if) is obtained by choosing $F = x_i$, $G = x_j$ and $F = x_k$ and varying i, j, k.$\triangleright$

As a consequence of the above proposition, the Hamiltonian vector fields are of the form (put $F = x_i$ and $G = H$ in (3.2))

$$\boxed{X_H(x) := \{x, H\} = J(x)\nabla_x H(x)} \ , \tag{3.5}$$

i.e. are proportional to the gradient of the Hamiltonian function through the function operator $J(x) := \{x, x\}$. The latter operator, when (3.4) is satisfied, takes the name of <u>Poisson tensor</u>, and generalizes the symplectic matrix of standard Hamiltonian mechanics (i.e. the one following from finite-dimensional Lagrangian mechanics).

Remark 3.2. *According to what just shown above,* ***any constant skew-symmetric tensor is a Poisson tensor*** *(explain why).*

Remark 3.3. *The standard formula expressing the derivative of a function $F(x)$ along the flow of $\dot{x} = X_H(x)$ in terms of the Poisson bracket of F and H holds, namely*

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \nabla F \cdot X_H = \nabla F \cdot J\nabla H = \{F, H\} \ . \tag{3.6}$$

Remark 3.4. *The deep meaning of the Jacobi identity (3. in Definition 3.1, or its equivalent (3.4)), is the following. Given two functions F and G, and their Poisson bracket $\{F, G\}$, the relation (3.6) implies $\dot{F} = \{F, H\}$, $\dot{G} = \{G, H\}$ and $\dot{\overline{\{F, G\}}} = \{\{F, G\}, H\}$. One then easily checks that the identity*

$$\dot{\overline{\{F, G\}}} = \{\dot{F}, G\} + \{F, \dot{G}\} \tag{3.7}$$

holds iff the Jacobi identity holds. In other words, the Jacobi identity is equivalent to the Leibniz rule on the Poisson bracket with respect to the time derivative.

A Poisson tensor $J(x)$ is singular at x if there exists a vector field $u(x) \not\equiv 0$ such that $J(x)u(x) = 0$, i.e. if $\ker J(x)$ is nontrivial. The functions $C(x)$ such that $\nabla C(x) \in \ker J(x)$ have a vanishing Poisson bracket with any other function F defined on Γ, since $\{F, C\} = \nabla F \cdot J\nabla C \equiv 0$ independently of F. Such special functions are called <u>Casimir invariants</u> associated to the given Poisson tensor J, and are constants of motion of *any* Hamiltonian system with vector field X_H associated to J, i.e. $\dot{C} = \{C, H\} = -\{H, C\} = -\nabla H \cdot J\nabla C \equiv 0$, which holds independently of H.

Example 3.1. *In the "mechanical case" $x = (q, p)$ and $J = J_{2n}$, i.e. the P-tensor is the standard symplectic matrix. This has been already discussed.*

Example 3.2. *The Euler equations, describing the evolution of the angular momentum $\boldsymbol{L}$ of a rigid body in a co-moving frame and in absence of external torque (moment of external forces), read $\dot{\boldsymbol{L}} = \boldsymbol{L} \wedge I^{-1}\boldsymbol{L}$, where I is the inertia tensor of the body (a 3×3 symmetric, positive definite matrix) and $\wedge$ denotes the standard vector product in $\mathbb{R}^3$. This is a Hamiltonian system with Hamiltonian function $H(\boldsymbol{L}) = \frac{1}{2}\boldsymbol{L} \cdot I^{-1}\boldsymbol{L}$ and Poisson tensor*

$$J(\boldsymbol{L}) := \left(\begin{array}{ccc} 0 & -L_3 & L_2 \\ L_3 & 0 & -L_1 \\ -L_2 & L_1 & 0 \end{array}\right) = \boldsymbol{L} \wedge \ . \tag{3.8}$$

In this way, the Euler equations have the standard form $\dot{\boldsymbol{L}} = J(\boldsymbol{L})\nabla_{\boldsymbol{L}} H(\boldsymbol{L})$, and the Poisson bracket of two functions $F(\boldsymbol{L})$ and $G(\boldsymbol{L})$ is $\{F, G\} = \boldsymbol{L} \cdot (\nabla G \wedge \nabla F)$ (one has of course to check that $J(\boldsymbol{L})$, as defined above, is actually a Poisson tensor, which is left as an exercise; see below). We notice that the Casimir invariants of $J(\boldsymbol{L})$ are all the functions $C(\boldsymbol{L}) := f\left(|\boldsymbol{L}|^2\right)$, since $\nabla C = f'\left(|\boldsymbol{L}|^2\right) 2\boldsymbol{L}$, so that $J(\boldsymbol{L})\nabla C = \boldsymbol{L} \wedge (2f'\boldsymbol{L}) = 0$.

Exercise 3.1. *Prove that the tensor (3.8) satisfies relation (3.4). Hint: observe that one can write $J_{ij}(\boldsymbol{L}) = \{L_i, L_j\} = -\sum_{k=1}^{3} \varepsilon_{ijk} L_k$, where ε_{ijk} is the Levi-Civita symbol with three indices $i, j, k = 1, 2, 3$, so defined: $\varepsilon_{ijk} = +1$ if (i, j, k) is an even permutation of $(1, 2, 3)$; $\varepsilon_{ijk} = -1$ if (i, j, k) is an odd permutation of $(1, 2, 3)$ and $\varepsilon_{ijk} = 0$ if any two indices are equal (recall that a permutation is even or odd when it is composed by an even or odd number of pair exchanges, respectively). The following relation turns out to be useful:*

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{kl}\delta_{jm} \ , \tag{3.9}$$

where δ_{ij} is the Kronecker delta, whose value is 1 if $i = j$ and zero otherwise.

Example 3.3. *Let us consider a single harmonic oscillator, with Hamiltonian $H = \frac{1}{2}\left(p^2 + \omega^2 q^2\right)$, and introduce the complex variables $z = (\omega q + \imath p)/\sqrt{2\omega}$ and $z^* = (\omega q - \imath p)/\sqrt{2\omega}$, where $\imath$ is the imaginary unit. In terms of such complex coordinates, known as complex Birkhoff coordinates, the Hamiltonian reads $H(z, z^*) = \omega |z|^2$, and the Hamilton equations become $\dot{z} = -\imath\omega z = -\imath\partial\widetilde{H}/\partial z^*$ and its complex conjugate $\dot{z}^* = \imath\omega z^* = \imath\partial\widetilde{H}/\partial z$. The new Poisson tensor is the second Pauli matrix[1], $\boldsymbol{\sigma}_2 = \left(\begin{array}{cc} 0 & -\imath \\ \imath & 0 \end{array}\right) = -\imath J_2$, that satisfies condition (3.4) because is constant (independent of z and z^*). Thus, with respect to the Birkhoff vector $\zeta = (z, z^*)^T$, the equations of motion of the Harmonic oscillator take on the new Hamiltonian form $\dot{\zeta} = \boldsymbol{\sigma}_2\nabla_\zeta H$.*

[1]The three Pauli matrices are

$$\boldsymbol{\sigma}_1 := \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right) \ ; \ \boldsymbol{\sigma}_2 := \left(\begin{array}{cc} 0 & -\imath \\ \imath & 0 \end{array}\right) \ ; \ \boldsymbol{\sigma}_3 := \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right) \ .$$

As further coordinates for the harmonic oscillator one can introduce the so-called action-angle variables. The latter are linked to the Birkhoff coordinates by the relations $z = \sqrt{I}e^{-\imath\varphi}$, $z^* = \sqrt{I}e^{\imath\varphi}$. *Observing that* $H = \omega|z|^2 = \omega I$, *so that* $\dot{I} = 0$, *from the equation of motion* $\dot{z} = -\imath\omega z$ *it follows that* $\dot{\varphi} = \omega$. *Thus the equations of motion in the action-angle variables read* $\dot{\varphi} = \partial H/\partial I$, $\dot{I} = -\partial H/\partial\varphi$, *with the standard symplectic matrix* J_2 *as Poisson tensor.*

3.1.1 Wave equation

The wave equation $u_{tt} = c^2 u_{xx}$ is an example of infinite dimensional Hamiltonian system. For what concerns the boundary conditions we consider the case of fixed ends: $u(t,0) = 0 = u(t,L)$, $u_t(t,0) = 0 = u_t(t,L)$. One can also write the wave equation in the equivalent first order form

$$\begin{cases} u_t = v \\ v_t = c^2 u_{xx} \end{cases} , \tag{3.10}$$

where fixed ends conditions hold now on v. The initial conditions for the latter system are $u_0(x) := u(0,x)$ and $v_0(x) := v(0,x)$. The quantity

$$H[u,v] = \int_0^L \frac{v^2 + c^2(u_x)^2}{2}\mathrm{d}x \tag{3.11}$$

is easily checked to be a constant of motion for system (3.10).

Now, the set of functions $\varphi_k(x) := \sqrt{2/L}\sin(\pi kx/L)$, $k = 1,2,\ldots$, constitutes an orthonormal basis in the Hilbert space $L_2([0,L])$ of square integrable functions on $[0,L]$ with fixed ends, since one easily checks that $\langle\varphi_i,\varphi_j\rangle := \int_0^L \varphi_i(x)\varphi_j(x)\mathrm{d}x = \delta_{ij}$. One can thus expand both $u(t,x)$ and $v(t,x)$ in Fourier series: $u(t,x) = \sum_{k\geq 1} q_k(t)\varphi_k(x)$, $v(t,x) = \sum_{k\geq 1} p_k(t)\varphi_k(x)$, with Fourier coefficients given by $q_k(t) = \langle\varphi_k,u\rangle$ and $p_k(t) = \langle\varphi_k,v\rangle$, respectively. Upon substitution of the latter Fourier expansions into the wave equation system (3.10) one easily gets $\dot{q}_k = p_k$, $\dot{p}_k = -\omega_k^2 q_k$, where $\omega_k := c\pi k/L$, $k \geq 1$. These are the Hamilton equations of a system of infinitely many harmonic oscillators, with Hamiltonian $K(q,p) = \sum_{k\geq 1}(p_k^2 + \omega_k^2 q_k^2)/2$, which is obviously integrable (in the Liouville sense). One easily finds that the substitution of the Fourier expansions of u and v into the function (3.11) yields $H = K$; to such a purpose, notice that $\int_0^L (u_x)^2\mathrm{d}x = uu_x|_0^L - \int_0^L uu_{xx}\mathrm{d}x$.

The solution of the Cauchy problem for the wave equation $u_{tt} = c^2 u_{xx}$ is thus immediately found:

$$u(t,x) = \sum_{k\geq 1}\left[q_k(0)\cos(\omega_k t) + \frac{p_k(0)}{\omega_k}\sin(\omega_k t)\right] e^{(k)}(x) , \tag{3.12}$$

where $q_k(0) = \langle\varphi_k,u_0\rangle$ and $p_k(0) = \langle\varphi_k,v_0\rangle$ are determined by the initial conditions.

Notice that if the Fourier coefficients of the string are initially zero for some given value of the index, then they are zero forever. In particular, if they are zero from some given value of the index on, then they are zero for any t: if $q_k(0) = 0 = p_k(0)$ for $k > M$ then $q_k(t) = 0 = p_k(t)$ for $k > M$. As a consequence, one can construct solutions of the string equation of any finite dimension M. The finite dimensional system thus obtained consists of M noninteracting harmonic oscillators. In particular, such a system is integrable in the Liouville sense: there exist M independent first integrals $H_1, \dots H_M$ in involution, i.e. such that $\{H_i, H_j\} = 0$ for any $i, j = 1, \dots, M$. In the specific case, the first integrals are the energies of the oscillators and the Hamiltonian is simply the sum of them. Such a property is valid for any M and is clearly preserved in the limit $M \to +\infty$ (the only attention to be paid deals with convergence problems).

3.1.2 Quantum mechanics

Quantum mechanics, in its simplest version, is build up as follows. One starts from the classical Hamiltonian $H = \sum_{j=1}^{N} \frac{|\boldsymbol{p}_j|^2}{2m_j} + U(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$ of a given system of N interacting particles, and therein substitutes all the momenta: $\boldsymbol{p}_j \to -\imath\hbar\nabla_j$, where $\hbar \simeq 10^{-27} erg \cdot sec$ is the normalized Planck constant[2] and $\nabla_j := \partial/\partial \boldsymbol{x}_j$. Such a procedure is referred to as the "canonical quantization" of the classical system. The operator thus obtained, namely

$$\hat{\mathsf{H}} := -\sum_{j=1}^{N} \frac{\hbar^2 \Delta_j}{2m_j} + U(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) \ , \tag{3.13}$$

where Δ_j is the Laplacian with respect to the coordinates of the j-th particle, is referred to as the quantum Hamiltonian, or Hamiltonian operator (or simply Hamiltonian, if the quantum mechanical context is given for granted) of the system. $\hat{\mathsf{H}}$ is naturally defined on a certain domain $D \subseteq \mathbb{R}^{3N}$ (depending on U) of the configuration space, and is easily checked to be Hermitian (or even self-adjoint) on the Hilbert space $L_2(D)$ of square integrable functions vanishing on ∂D, i.e., denoting by $\langle F, G\rangle := \int_D F^* G \mathrm{d}V$ the scalar product ($\mathrm{d}V := \mathrm{d}^3x_1 \dots \mathrm{d}^3x_N$), one has $\left\langle \hat{\mathsf{H}}F, G \right\rangle = \left\langle F, \hat{\mathsf{H}}G \right\rangle$. The latter condition is usually denoted by $\hat{\mathsf{H}} = \hat{\mathsf{H}}^\dagger$ that, in the sequel, will be always meant to denote self-adjointness.

The Schrödinger equation describing the quantum dynamics of the given system of N particles is

$$\imath\hbar \frac{\partial \Psi}{\partial t} = \hat{\mathsf{H}}\Psi \ , \tag{3.14}$$

where the complex function $\Psi(t, \boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$ is the so-called "wave function" of the system, satisfying the boundary condition $\Psi|_{\partial D} = 0$. Notice that the unitary evolution (3.14) implies that $\int_D |\Psi|^2 \mathrm{d}V$ is independent of time; in other words, $\Psi|_{t=0} \in L_2(D)$ implies $\Psi \in L_2(D)$ for any $t \in \mathbb{R}$. The physical meaning of Ψ is that $\int_A |\Psi|^2 \mathrm{d}V$ is the probability to find the particles in the region $A \subseteq D$ of the configuration space at time t.

We now suppose that the spectrum of the Hamiltonian operator (3.13) is discrete, and denote by $\{\Phi_j\}_j$ and $\{E_j\}_j$ the set of orthonormal eigenfunctions and of the corresponding eigenvalues of $\hat{\mathsf{H}}$, respectively, so that $\hat{\mathsf{H}}\Phi_j = E_j\Phi_j$, and $\langle \Phi_k, \Phi_j\rangle := \int_D \Phi_k^* \Phi_j \mathrm{d}V = \delta_{kj}$. One can then expand the wave-function Ψ on the basis of the eigenfunctions of $\hat{\mathsf{H}}$: $\Psi = \sum_j c_j \Phi_j$.

[2] The Planck constant is $h \simeq 6.626\ 10^{-27} erg \cdot sec$; the normalized constant, which actually enters the theory, is $\hbar := h/(2\pi)$.

Substituting the latter expression into the Schrödinger equation (3.14) and taking the scalar product with Φ_k one gets

$$\imath\dot{c}_k = \omega_k c_k \quad ; \quad \omega_k := \frac{E_k}{\hbar} \ , \tag{3.15}$$

valid for any "index" k. The latter equation is that of the harmonic oscillator in complex Birkhoff coordinates, so that the Hamiltonian of the system is

$$K(c, c^*) = \sum_j E_j |c_j|^2 \ . \tag{3.16}$$

Equation (3.15) reads thus $\imath\dot{c}_k = \hbar^{-1}\partial K/\partial c_k^*$ for any k. Making use of the expansion $\Psi = \sum_j c_j \Phi_j$ of the wave function one easily checks that

$$K(c, c^*) = \left\langle \Psi, \hat{\mathsf{H}}\Psi \right\rangle \ , \tag{3.17}$$

known as the "quantum expectation of the total energy". One can easily check that the conservation of $K = \left\langle \Psi, \hat{\mathsf{H}}\Psi \right\rangle$ follows directly from the Schrödinger equation with the condition that $\hat{\mathsf{H}}$ is independent of time t. Another conserved quantity is clearly

$$\int_D |\Psi|^2 \mathrm{d}V = \langle \Psi, \Psi \rangle = \sum_j |c_j|^2 \ (= 1) \ . \tag{3.18}$$

The solution of the Cauchy problem for the Schrödinger equation (3.14) is

$$\Psi(t, \boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \sum_k c_k(0) e^{-\imath(E_k/\hbar)t} \Phi_k(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) \ , \tag{3.19}$$

where $c_k(0) = \langle \Phi_k, \Psi \rangle |_{t=0}$. The solution (3.19) can be obtained also by introducing the formal solution of the Schrödinger equation in the operator form, namely

$$\Psi(t) = e^{-\imath \frac{\hat{\mathsf{H}}}{\hbar} t} \Psi(0) := \hat{\mathsf{U}}^t \Psi(0) \ , \tag{3.20}$$

where the unitary evolution operator $\hat{\mathsf{U}}^t := e^{-\imath(\hat{\mathsf{H}}/\hbar)t}$ has been introduced. Expanding the wave function $\Psi(0)$ on the basis of the eigenfunctions Φ_j of $\hat{\mathsf{H}}$ and observing that $\hat{\mathsf{U}}^t\Phi_j = e^{-\imath(E_j/\hbar)t}\Phi_j$, one gets (3.19).

Thus when the spectrum of $\hat{\mathsf{H}}$ is purely discrete, as just assumed above, the quantum mechanical problem of any system of interacting particles is essentially equivalent to that of a vibrating string: the dynamics of the system is decomposed into its "normal modes", i.e. a collection of noninteracting harmonic oscillators, and is thus integrable in the classical Liouville sense. In this respect, no relaxation phenomenon can take place if the system is left free to evolve, isolated from external perturbations. Things are instead quite different in the presence of a continuous spectral component of $\hat{\mathsf{H}}$, as can be seen through simple examples.

3.2 Change of variables

Let us consider a Hamiltonian dynamical system

$$\dot{x} = J(x)\nabla_x H(x) \ , \quad (x \in \Gamma) \tag{3.21}$$

and perform a change of variables $f : x \mapsto y = f(x)$, with inverse $g := f^{-1} : y \mapsto x = g(y)$. The gradients with respect to x and y are connected by the chain rule, namely

$$\frac{\partial}{\partial x_i} = \sum_j \left(\frac{\partial y_j}{\partial x_i}\right) \frac{\partial}{\partial y_j} \ \text{ i.e. } \ \nabla_x = \left(\frac{\partial f}{\partial x}\right)^T \nabla_y \ .$$

The Jacobians of f and g are instead linked by the identity

$$\left(\frac{\partial g}{\partial y}\right)^{-1} = \frac{\partial f}{\partial x}(g(y))$$

that in turn follows differentiating the identity $f(g(y)) = y$. Using such relations one immediately shows that in the new variables the Hamilton equation (3.21) reads

$$\dot{y} = J^{\#}(y)\nabla_y \widetilde{H}(y) \ , \quad (y \in f(\Gamma)) \tag{3.22}$$

where

$$\widetilde{H}(y) := H(g(y)) \ , \tag{3.23}$$

and

$$J^{\#}(y) := \left(\frac{\partial f}{\partial x}\right) J(x) \left(\frac{\partial f}{\partial x}\right)^T \Bigg|_{x=g(y)} = \left(\frac{\partial g}{\partial y}\right)^{-1} J(g(y)) \left(\frac{\partial g}{\partial y}\right)^{-T} \ , \tag{3.24}$$

where the superscript $-T$ on the right hand side of the latter definition means inverse and transpose[3].

The first question that naturally poses is whether and when equation (3.22) is still Hamiltonian, i.e. whether and when the tensor $J^{\#}(y)$, defined in (3.24), is a Poisson tensor. The equivalent (and deeper) question is how a given Poisson algebra transforms under change of variables. In this respect, the following proposition holds.

Proposition 3.2. *Poisson brackets are characterized by coordinate-independent properties.*

$\triangleleft$ PROOF. Given a Poisson bracket $\{F, G\}(x) = \nabla F(x) \cdot J(x)\nabla G(x)$ we want to show that it transforms into another Poisson bracket under **any** change of variables $f : x \mapsto y$; as above,

[3]For any invertible operator L one has $LL^{-1} = \mathbb{I}$, so that $(LL^{-1})^T = (L^{-1})^T L^T = \mathbb{I}$ and $(L^T)^{-1} = (L^{-1})^T$. Thus the notation L^{-T} is well defined, independent of the order of the inversion and of the transposition.

when convenient, $g := f^{-1}$. According to (3.23), a tilde denotes composition with the inverse, namely $\widetilde{F}(y) := F(g(y))$. By means of (3.24) one finds

$$\begin{aligned}
\widetilde{\{F,G\}}(y) &= \{F,G\}(g(y)) = \left[\left(\frac{\partial f}{\partial x}\right)^T \nabla_y F\right] \cdot J(x) \left(\frac{\partial f}{\partial x}\right)^T \nabla_y G\Bigg|_{x=g(y)} = \\
&= \nabla_y \widetilde{F}(y) \cdot \left[\left(\frac{\partial g}{\partial y}\right)^{-1} J(g(y)) \left(\frac{\partial g}{\partial y}\right)^{-T}\right] \nabla_y \widetilde{G}(y) = \\
&= \nabla_y \widetilde{F}(y) \cdot J^{\#}(y) \nabla_y \widetilde{G}(y) := \{\widetilde{F}, \widetilde{G}\}_{\#}(y) \ . \qquad (3.25)
\end{aligned}$$

Equivalently, with notation independent of coordinates, one has

$$\widetilde{\{F,G\}} = \{\widetilde{F}, \widetilde{G}\}_{\#} \ \Leftrightarrow \ \{F,G\} \circ g = \{F \circ g, G \circ g\}_{\#} \ . \qquad (3.26)$$

We must show that the bracket $\{\widetilde{F}, \widetilde{G}\}_{\#}$, formally defined in the last step of (3.25), is an actual Poisson bracket on the algebra of the transformed (i.e. "tilded") functions. To this end, we observe that skew-symmetry, bi-linearity and the Leibniz property follow directly from those of $\{\ ,\}$, through the identity (3.26) (show it). In fact, repeated use of the latter relation also implies the validity of the Jacobi identity:

$$\begin{aligned}
0 &\equiv \{F, \widetilde{\{G,H\}}\} + \{G, \widetilde{\{H,F\}}\} + \{H, \widetilde{\{F,G\}}\} = \\
&= \{\widetilde{F}, \widetilde{\{G,H\}}\ \}_{\#} + \{\widetilde{G}, \widetilde{\{H,F\}}\}_{\#} + \{\widetilde{H}, \widetilde{\{F,G\}}\}_{\#} = \\
&= \{\widetilde{F}, \{\widetilde{G}, \widetilde{H}\}_{\#}\}_{\#} + \{\widetilde{G}, \{\widetilde{H}, \widetilde{F}\}_{\#}\}_{\#} + \{\widetilde{H}, \{\widetilde{F}, \widetilde{G}\}_{\#}\}_{\#} \ . \qquad (3.27)
\end{aligned}$$

Thus, the change of variables f transforms Poisson brackets into Poisson Brackets.▷

Since $\{\ ,\}_{\#}$ is a Poisson bracket, it follows that (3.24) is a Poisson tensor and that, as a consequence, the transformed system (3.22) is Hamiltonian. The conclusion is that **a given dynamical system is Hamiltonian independently of the coordinates chosen**.

Notice that, from (3.26), with the choice $F = f_i$ and $G = f_j$, it follows that

$$\{y_i, y_j\}_{\#} = \{f_i, f_j\} \circ g \qquad (3.28)$$

holds for any change of variables $f : x \mapsto y$. It must be stressed that in the bracket on the left hand side of (3.28) y is the independent variable, whereas in the bracket on the right hand side the independent variable is x (and the composition with g yields back a function of y).

Example 3.4. *Let us reconsider the example of the complex Birkhoff coordinates for the harmonic oscillator. Let us set $x = (q,p)^T$ and $y = (z, z^*)^T$, so that the change of variables $y = f(x)$ is defined by the formulas $z = (\omega q + \imath p)/\sqrt{2\omega}$, $z^* = (\omega q - \imath p)/\sqrt{2\omega}$. Formula (3.24)* *for the transformed Poisson tensor reads*

$$
\begin{aligned}
J^{\#} &= \frac{\partial(z,z^*)}{\partial(q,p)} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \frac{\partial(z,z^*)}{\partial(q,p)}^T = \\
&= \begin{pmatrix} \sqrt{\omega/2} & \imath/\sqrt{2\omega} \\ \sqrt{\omega/2} & -\imath/\sqrt{2\omega} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{\omega/2} & \sqrt{\omega/2} \\ \imath/\sqrt{2\omega} & -\imath/\sqrt{2\omega} \end{pmatrix} = \\
&= \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix} = \boldsymbol{\sigma}_2 \ ,
\end{aligned}
\tag{3.29}
$$

which shows how the standard, symplectic Poisson structure is mapped into the Birkhoff one. The "new" Poisson bracket in Birkhoff coordinates reads

$$
\begin{aligned}
\{\widetilde{F}, \widetilde{G}\}_{\#} &= \begin{pmatrix} \partial\widetilde{F}/\partial z \\ \partial\widetilde{F}/\partial z^* \end{pmatrix} \cdot \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix} \begin{pmatrix} \partial\widetilde{G}/\partial z \\ \partial\widetilde{G}/\partial z^* \end{pmatrix} \\
&= -\imath \left(\frac{\partial\widetilde{F}}{\partial z}\frac{\partial\widetilde{G}}{\partial z^*} - \frac{\partial\widetilde{F}}{\partial z^*}\frac{\partial\widetilde{G}}{\partial z} \right) \ .
\end{aligned}
$$

The Poisson tensor in Birkhoff coordinates is determined by the relations $\{z, z^*\} = -\imath$, $\{z, z\} = \{z^*, z^*\} = 0$.

Remark 3.5. *To a Hamiltonian sum of harmonic oscillator Hamiltonians, i.e.* $H = \sum_j \omega_j |z_j|$, *there corresponds the Poisson bracket*

$$
\{F, G\} = -\imath \sum_j \left(\frac{\partial F}{\partial z_j}\frac{\partial G}{\partial z_j^*} - \frac{\partial F}{\partial z_j^*}\frac{\partial G}{\partial z_j} \right) \ ,
$$

with a Poisson tensor specified by the relations $\{z_j, z_k^*\} = -\imath\delta_{jk}$, $\{z_j, z_k\} = \{z_j^*, z_k^*\} = 0$.

This is the case, for example, of quantum mechanics (where $z_j = c_j$), but for a prefactor $\hbar^{-1}$. One easily checks that with the Poisson bracket

$$
\{F, G\} = -\frac{\imath}{\hbar} \sum_j \left(\frac{\partial F}{\partial c_j}\frac{\partial G}{\partial c_j^*} - \frac{\partial F}{\partial c_j^*}\frac{\partial G}{\partial c_j} \right) \ , \tag{3.30}
$$

and the Hamiltonian $K = \sum_j E_j |c_j|^2$, the Hamilton equations $\dot{c}_k = \{c_k, K\}$ are exactly the Schrödinger ones (3.15).

Exercise 3.2. *Suppose* $F(c, c^*) = \left\langle \Psi, \hat{\mathsf{F}}\Psi \right\rangle$ *and* $G(c, c^*) = \left\langle \Psi, \hat{\mathsf{G}}\Psi \right\rangle$, *where* $\Psi = \sum_k c_k \Phi_k$. *Show that*

$$
\{F, G\} = \left\langle \Psi, \frac{1}{\imath\hbar}[\hat{\mathsf{F}}, \hat{\mathsf{G}}]\Psi \right\rangle \ , \tag{3.31}
$$

where $[\hat{\mathsf{F}}, \hat{\mathsf{G}}] := \hat{\mathsf{F}}\hat{\mathsf{G}} - \hat{\mathsf{G}}\hat{\mathsf{F}}$ *is the commutator of the self-adjoint operators* $\hat{\mathsf{F}}$ *and* $\hat{\mathsf{G}}$.

Example 3.5. *Let us consider the case of a constant Poisson tensor $J := A$ in finite dimension n. The real $n \times n$ matrix A is only supposed to be skew-symmetric: $A^T = -A$. We consider the Hamiltonian system*

$$\dot{x} = A\nabla_x H(x) \tag{3.32}$$

and perform a linear change of variables $x \mapsto y = Lx$. In this case one has $\partial f/\partial x = L$, so that formula (3.24) yields the transformed Poisson tensor $J^{\#} = LAL^T$. The Linear change of variables is built up in such a way that $J^{\#}$ turns out to be as close as possible to a standard symplectic matrix. To this purpose, the main properties of the matrix A are listed below.

1. *A is a normal matrix, i.e. commutes with its Hermitian conjugate*[4], *so that it is unitarily diagonalizable: $\exists U$ such that $U^\dagger U = \mathbb{I}$ and $U^\dagger AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

2. *The spectrum of A consists of pure imaginary numbers $\lambda_1, \ldots, \lambda_n \in \imath\mathbb{R}$; if λ is an eigenvalue with eigenvector $\boldsymbol{x} =\in \mathbb{C}^n$, also $\lambda^* = -\lambda$ is an eigenvalue, with eigenvector $\boldsymbol{x}^*$. Thus there are $2k = n - r$ non zero eigenvalues, where $r := \dim(\ker A)$.*

3. *Consider the eigenvalue problem*

$$A(\boldsymbol{a} + \imath\boldsymbol{b}) = \imath\mu(\boldsymbol{a} + \imath\boldsymbol{b}) \ , \ \Leftrightarrow \ \begin{cases} A\boldsymbol{a} = -\mu\boldsymbol{b} \\ A\boldsymbol{b} = \mu\boldsymbol{a} \end{cases} \tag{3.33}$$

where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ and $\mu > 0$ without loss of generality. Then there exist a solution consisting of k positive numbers $\mu_1, \ldots, \mu_k$ and k pairs of real vectors $(\boldsymbol{a}_j, \boldsymbol{b}_j)$, $j = 1, \ldots, k$, such that $\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 0$, $\boldsymbol{a}_i \cdot \boldsymbol{a}_j = \boldsymbol{b}_i \cdot \boldsymbol{b}_j = \delta_{ij}$ for any $i, j = 1, \ldots, n$.

4. *The orthonormal vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ are also orthogonal to $\ker A$. Thus, given any orthonormal basis $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ of $\ker A$, the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ constitute an orthonormal basis of $\mathbb{R}^n$.*

The matrix L that defines the linear change of variables is now given by

$$L^T := \left(\frac{\boldsymbol{a}_1}{\sqrt{\mu_1}} \middle| \cdots \middle| \frac{\boldsymbol{a}_k}{\sqrt{\mu_k}} \middle| \frac{\boldsymbol{b}_1}{\sqrt{\mu_1}} \middle| \cdots \middle| \frac{\boldsymbol{b}_k}{\sqrt{\mu_k}} \middle| \boldsymbol{w}_1 \middle| \cdots \middle| \boldsymbol{w}_r \right) \ , \tag{3.34}$$

defined by its column vectors. One then easily checks that

$$J^{\#} = LAL^T = \begin{pmatrix} \mathbb{O}_k & \mathbb{I}_k & \mathbb{O}_{k,r} \\ -\mathbb{I}_k & \mathbb{O}_k & \mathbb{O}_{k,r} \\ \mathbb{O}_{r,k} & \mathbb{O}_{r,k} & \mathbb{O}_{r,r} \end{pmatrix} \ . \tag{3.35}$$

The Hamilton equations $\dot{y} = J^{\#}\nabla_y \widetilde{H}(y)$ take on a very simple form. Indeed, setting $y := (q, p, \xi)$, where $q, p \in \mathbb{R}^k$ and $\xi \in \mathbb{R}^r$, one gets

$$\dot{q} = \frac{\partial \widetilde{H}}{\partial p} \ ; \ \dot{p} = -\frac{\partial \widetilde{H}}{\partial q} \ ; \ \dot{\xi} = 0 \ . \tag{3.36}$$

[4] Recall that the Hermitian conjugate of a matrix, or of a linear operator, M is the transpose of its complex conjugate, or vice versa, namely $M^\dagger := (M^*)^T = (M^T)^*$.

Thus the Hamiltonian $\widetilde{H}(q,p;\xi)$ depends on the r parameters $\xi_1,\dots,\xi_r$. Such parameters are nothing but the Casimir invariants of the system, and their expression in terms of the old variables x is obtained by computing the last r components of the vector $y = Lx$.

3.2.1 Canonical transformations

Given the Hamiltonian dynamical system (3.21), among all the possible changes of variables concerning it, a privileged role is played by those leaving the Poisson tensor and, as a consequence, the Hamilton equations invariant in form, i.e. mapping the equation (3.21) into the equation $\dot{y} = J(y)\nabla_y\widetilde{H}(y)$, with the same Poisson tensor and a transformed Hamiltonian $\widetilde{H} = H \circ g$. Such particular changes of variables are the so-called *canonical transformations* of the given Poisson structure and are characterized by the following equivalent conditions:

$$J^{\#}(y) = J(y) \ ; \tag{3.37}$$
$$\{y_i, y_j\} = \{f_i, f_j\} \circ g \ ; \tag{3.38}$$
$$\{F \circ g, G \circ g\} = \{F, G\} \circ g \ . \tag{3.39}$$

Exercise 3.3. *Prove the equivalence of conditions (3.37), (3.38) and (3.39).*

The set of all the canonical transformations of a given Poisson structure has a natural group structure with respect to composition (they are actually a subgroup of all the change of variables).

Example 3.6. *Let us consider the standard Hamiltonian case, where the reference Poisson tensor is the standard symplectic matrix J_{2n} defined in (2.18). If $x = (q,p)$ and $y = (Q,P) = f(q,p)$, taking into account formula (3.24) with $J = J_{2n}$, condition (3.37) reads*

$$J_{2n} = \left(\frac{\partial f}{\partial x}\right) J_{2n} \left(\frac{\partial f}{\partial x}\right)^T \ ,$$

namely

$$\begin{pmatrix} \mathbb{O}_n & \mathbb{I}_n \\ -\mathbb{I}_n & \mathbb{O}_n \end{pmatrix} = \left(\frac{\partial(Q,P)}{\partial(q,p)}\right) \begin{pmatrix} \mathbb{O}_n & \mathbb{I}_n \\ -\mathbb{I}_n & \mathbb{O}_n \end{pmatrix} \left(\frac{\partial(Q,P)}{\partial(q,p)}\right)^T \ ,$$

which is equivalent to the relations

$$\left[\frac{\partial Q}{\partial q}\left(\frac{\partial Q}{\partial p}\right)^T - \frac{\partial Q}{\partial p}\left(\frac{\partial Q}{\partial q}\right)^T\right]_{ij} = \sum_{s=1}^n \left(\frac{\partial Q_i}{\partial q_s}\frac{\partial Q_j}{\partial p_s} - \frac{\partial Q_i}{\partial p_s}\frac{\partial Q_j}{\partial q_s}\right) := \{Q_i, Q_j\}_{q,p} = 0 \ ;$$
$$\left[\frac{\partial P}{\partial q}\left(\frac{\partial P}{\partial p}\right)^T - \frac{\partial P}{\partial p}\left(\frac{\partial P}{\partial q}\right)^T\right]_{ij} = \sum_{s=1}^n \left(\frac{\partial P_i}{\partial q_s}\frac{\partial P_j}{\partial p_s} - \frac{\partial P_i}{\partial p_s}\frac{\partial P_j}{\partial q_s}\right) := \{Q_i, Q_j\}_{q,p} = 0 \ ;$$
$$\left[\frac{\partial Q}{\partial q}\left(\frac{\partial P}{\partial p}\right)^T - \frac{\partial Q}{\partial p}\left(\frac{\partial P}{\partial q}\right)^T\right]_{ij} = \sum_{s=1}^n \left(\frac{\partial Q_i}{\partial q_s}\frac{\partial P_j}{\partial p_s} - \frac{\partial Q_i}{\partial p_s}\frac{\partial P_j}{\partial q_s}\right) := \{Q_i, P_j\}_{q,p} = \delta_{ij} \ ,$$

for all $i, j = 1, \dots, n$. Such relations are the necessary and sufficient conditions for a change of variables to be canonical in standard Hamiltonian mechanics. The same relations can be obtained by applying the equivalent conditions (3.38) (check it).

Example 3.7. *Consider again the harmonic oscillator of example 3.3. The transformation $(q, p) \mapsto (\varphi, I)$ defined by $q = \sqrt{2I/\omega}\cos\varphi$, $p = -\sqrt{2\omega I}\sin\varphi$, is canonical, since $\{\varphi, I\}_{q,p} = 1$. The new Hamiltonian reads $\widetilde{H}(\varphi, I) = \omega I$, and the corresponding equations read $\dot{\varphi} = \omega$, $\dot{I} = 0$.*

Example 3.8. *Consider the Euler equations $\dot{\boldsymbol{L}} = \boldsymbol{L} \wedge I^{-1}\boldsymbol{L} = J(\boldsymbol{L})\nabla_L H(\boldsymbol{L})$ for the free rigid body, where $J(\boldsymbol{L}) = \boldsymbol{L}\wedge$ and $H = \frac{1}{2}\boldsymbol{L} \cdot I^{-1}\boldsymbol{L}$. The linear change of variable $\boldsymbol{L}' = R\boldsymbol{L}$, where R is any orthogonal matrix ($RR^T = \mathbb{I}_3$), is canonical. Indeed, formula (3.24), with $\partial f/\partial x = \partial \boldsymbol{L}'/\partial \boldsymbol{L} = R$, gives*

$$J^{\#}(\boldsymbol{L}')\boldsymbol{\xi} = RJ(\boldsymbol{L})R^T\big|_{\boldsymbol{L}=R^T\boldsymbol{L}'}\boldsymbol{\xi} = R[(R^T\boldsymbol{L}') \wedge (R^T\boldsymbol{\xi})] = \boldsymbol{L}' \wedge \boldsymbol{\xi} = J(\boldsymbol{L}')\boldsymbol{\xi} \ ,$$

which, being valid for any vector $\boldsymbol{\xi}$, implies $J^{\#}(\boldsymbol{L}') = J(\boldsymbol{L}')$. The transformed Hamiltonian is

$$\widetilde{H}(\boldsymbol{L}') = H(R^T\boldsymbol{L}') = \frac{1}{2}\boldsymbol{L}' \cdot (RI^{-1}R^T)\boldsymbol{L}' := \frac{1}{2}\boldsymbol{L}' \cdot \tilde{I}^{-1}\boldsymbol{L}' \ ,$$

which has the same functional form of H, but for the transformed inertia tensor $\tilde{I} := RIR^T$; observe that $\tilde{I}^{-1} = (RIR^T)^{-1} = RI^{-1}R^T$.

Sometimes, the requirement of canonicity in the sense stated above turns out to be too restrictive. For example, the simple re-scaling $(q, p, H, t) \mapsto (Q, P, K, T) = (aq, bp, cH, dt)$, depending on four real parameters a, b, c, d, preserves the form of the Hamilton equations, namely $\mathrm{d}Q/\mathrm{d}T = \partial K/\partial P$, $\mathrm{d}P/\mathrm{d}T = -\partial K/\partial Q$, under the unique condition $ab = cd$. On the other hand, canonicity in strict sense would require $ab = 1$. In this case, the extra factor ab gained by the transformed Poisson tensor is re-absorbed by a rescaling of Hamiltonian and time.

One is thus naturally led to call canonical transformations those changes of variables (time and Hamiltonian included) that preserve the final form of Hamilton equations, with the same Poisson tensor. In particular, a change of phase space coordinates $x \mapsto y = f(x)$ such that $J^{\#}(y) = cJ(y)$, can be completed to a canonical transformation by re-absorbing the constants c through the time rescaling $T = ct$. Indeed, the change of variables

$$(x, J, H, t) \mapsto (y, J^{\#}, \widetilde{H}, T)$$

maps the original Hamilton equations $\mathrm{d}x/\mathrm{d}t = J\nabla_x H$ into $\mathrm{d}y/\mathrm{d}T = J(y)\nabla_y \widetilde{H}(y)$, and is thus canonical, in the extended sense of preserving the given (Poisson) structure of the Hamilton equations.

Example 3.9. *Consider the Euler equations for the free rigid body. Since $|\boldsymbol{L}|$ is a constant of motion (Casimir invariant), one can reasonably consider only unit vectors. Let us set $\ell := |\boldsymbol{L}|$ and $\boldsymbol{u} := \boldsymbol{L}/\ell$. One easily checks that the change of phase space variable $\boldsymbol{L} \mapsto \boldsymbol{u}$ gives a transformed Poisson tensor $J^{\#} = J/\ell$ and a transformed Hamiltonian $\widetilde{H} = \ell^2 H$. It thus follows that the change of variables*

$$(\boldsymbol{L}, J, H, t) \mapsto (\boldsymbol{u}, J/\ell, \ell^2 H, \ell t) \ ,$$

is canonical. The Euler equations on the unit sphere read $\mathrm{d}\boldsymbol{u}/\mathrm{d}T = \boldsymbol{u} \wedge I^{-1}\boldsymbol{u}$.

3.2.2 Canonicity of Hamiltonian flows

A very convenient way of performing canonical transformations is to do it through Hamiltonian flows. To such a purpose, let us consider a Hamiltonian $G(x)$ and its associated Hamilton equations $\dot{x} = X_G(x)$. Let Φ_G^s denote the flow of G, so that $\Phi_G^s(\xi)$ is the solution of the Hamilton equations at time s, corresponding to the initial condition ξ at $s = 0$. We also denote by

$$L_G := \{\ , G\} = (J\nabla G)\cdot\nabla = X_G\cdot\nabla \tag{3.40}$$

the Lie derivative along the Hamiltonian vector field X_G; notice that $L_G F = \{F, G\}$.

Lemma 3.1. *For any function F one has*

$$F\circ\Phi_G^s = e^{sL_G}F\ . \tag{3.41}$$

◁ PROOF. Set $\widetilde{F}(s) := F\circ\Phi_G^s$ (observe that $\widetilde{G}(s) = G\circ\Phi_G^s = G$), and notice that $\widetilde{F}(0) = F$. Then

$$\frac{\mathrm{d}}{\mathrm{d}s}\widetilde{F}(0) = \{F, G\} = L_G F\ .$$

Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\widetilde{F}(s) &= \lim_{h\to 0}\frac{\widetilde{F}(s+h)-\widetilde{F}(s)}{h} = \lim_{h\to 0}\frac{\widetilde{F}(s)\circ\Phi_G^h-\widetilde{F}(s)}{h} = \\ &= \frac{\mathrm{d}}{\mathrm{d}h}\widetilde{F}(s)\circ\Phi_G^h\Big|_{h=0} = \{\widetilde{F}(s), G\} = L_G\widetilde{F}(s)\ ,\end{aligned}$$

whose solution is

$$\widetilde{F}(s) = e^{sL_G}\widetilde{F}(0) = e^{sL_G}F\ .\ \ ▷$$

Proposition 3.3. *If G is independent of s, the change of variables $x\mapsto y = \Phi_G^{-s}(x)$ defined by its flow at time $-s$ constitutes a one-parameter group of canonical transformations.*

◁ PROOF. The group properties follow from those of the flow. For what concerns canonicity, we prove the validity of condition (3.39), with $f := \Phi_G^{-s}$ and $g := f^{-1} = \Phi_G^s$, namely we prove that

$$\{F\circ\Phi_G^s, H\circ\Phi_G^s\} = \{F, H\}\circ\Phi_G^s\ \ \forall s\ . \tag{3.42}$$

The equivalent statement is that the difference

$$D(s) := \{F\circ\Phi_G^s, H\circ\Phi_G^s\} - \{F, H\}\circ\Phi_G^s = \{e^{sL_G}F, e^{sL_G}H\} - e^{sL_G}\{F, H\}$$

identically vanishes. Observe that $D(0) = 0$; moreover $G = \widetilde{G}$. One finds

$$\begin{aligned}\frac{\mathrm{d}D}{\mathrm{d}s} &= \{L_G\widetilde{F}, \widetilde{H}\} + \{\widetilde{F}, L_G\widetilde{H}\} - L_G\widetilde{\{F, H\}} = \{\{\widetilde{F}, \widetilde{G}\}, \widetilde{H}\} + \{\widetilde{F}, \{\widetilde{H}, \widetilde{G}\}\} - \{\widetilde{\{F, H\}}, \widetilde{G}\} = \\ &= \{\{\widetilde{F}, \widetilde{G}\}, \widetilde{H}\} + \{\{\widetilde{G}, \widetilde{H}\}, \widetilde{F}\} - \{\widetilde{\{F, G\}}, \widetilde{H}\} = \{\{\widetilde{F}, \widetilde{H}\}, \widetilde{G}\} - \{\widetilde{\{F, H\}}, \widetilde{G}\} = \\ &= \{\{\widetilde{F}, \widetilde{H}\} - \widetilde{\{F, H\}}, \widetilde{G}\} = \{D(s), G\} = L_G D(s)\ .\end{aligned}$$

The unique solution of this differential equation, namely $D'(s) = L_G D(s)$, with initial datum $D(0) = 0$, is $D(s) \equiv 0$. $\rhd$

An interesting application of the above formalism is the following Hamiltonian version of the Nöther theorem, linking symmetries to first integrals.

Proposition 3.4. *If the Hamiltonian H is invariant with respect to the Hamiltonian flow of the Hamiltonian K, i.e. $H \circ \Phi_K^s = H$, then K is a first integral of H.*

$\lhd$ PROOF. One has

$$0 = \frac{\mathrm{d}H}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s} H \circ \Phi_K^s = \{H, K\} \circ \Phi_K^s \ ,$$

for any $s \in \mathbb{R}$. In particular, for $s = 0$ one gets $\{H, K\} = 0$.$\rhd$

In order to apply the previous proposition, one usually has to

1. find a one parameter group of symmetry for H, namely a transformation of coordinates $x \mapsto y = f^s(x)$, depending on a real parameter s, such that $f_1^s \circ f^{s_2} = f^{s_1+s_2}$ for any pair $s_1, s_2 \in \mathbb{R}$ and $f^0(x) = x$ for any x, and such that $H \circ f^s = H$;

2. check whether f^s is a Hamiltonian flow, namely whether there exists a Hamiltonian K such that $f^s = \Phi_K^s$.

As a matter of fact, there is no recipe for point 1., whereas for point 2. one has to check whether the *generator* of the group, namely the vector field $\partial f^s(x)/\partial s|_{s=0}$, is a Hamiltonian vector field. In practice one writes the equation

$$\left. \frac{\partial f^s(x)}{\partial s} \right|_{s=0} = J(x) \nabla K(x)$$

and looks for a solution $K(x)$; if such a solution exists then K is a first integral of H. In the equation above $J(x)$ is the Poisson tensor fixed for H. We recall that the generator of the group $u := \partial f^s/\partial s|_{s=0}$ is the vector field whose flow is f^s, i.e. the vector field of the differential equation whose solution at any time s with initial datum x is just $y(s) := f^s(x)$. This is easily checked as follows:

$$\frac{\mathrm{d}y}{\mathrm{d}s} = \lim_{h \to 0} \frac{y(s+h) - y(s)}{h} = \lim_{h \to 0} \frac{f^h(y) - y}{h} = \left. \frac{\partial f^h(y)}{\partial h} \right|_{h=0} = u(y) \ .$$

Exercise 3.4. *Consider the single particle Hamiltonian*

$$H = \frac{|\boldsymbol{p}|^2}{2m} + V(\boldsymbol{q}) \ ,$$

where $\boldsymbol{q}, \boldsymbol{p} \in \mathbb{R}^3$. i) Determine the conditions under which H is invariant under the one-parameter group of space translations $\boldsymbol{q} \mapsto \boldsymbol{q} + s\boldsymbol{u}$, $\boldsymbol{p} \mapsto \boldsymbol{p}$, where $\boldsymbol{u}$ is a given unit vector. Write such a transformation as a Hamiltonian flow and determine the corresponding first integral. ii) Repeat the same analysis with the one-parameter group of rotations $\boldsymbol{q} \mapsto R(s)\boldsymbol{q}$, $\boldsymbol{p} \mapsto R(s)\boldsymbol{p}$, where $R(s) = e^{sA}$, A being a given skew-symmetric matrix. Finally, show that for one parameter subgroups of rotation matrices the exponential form $R(s) = e^{sA}$ is generic.

Exercise 3.5. *Repeat the previous exercise for the corresponding quantum problem and show that the conserved quantities corresponding to the two symmetries are the canonically quantized classical ones. Hint:* $x \mapsto x'(s)$ *implies* $\psi(x) \mapsto \psi' := \psi(x'(s))$.

Chapter 4

Perturbation theory

The basic elements of Hamiltonian perturbation theory are reported, with particular attention to the construction of the normal form to some given order.

4.1 Normal form Hamiltonian

Let us consider a Hamiltonian system of the form

$$H_\lambda = h + \lambda P_1 + \lambda^2 P_2 + \cdots + \lambda^n P_n + R_{n+1} \ , \tag{4.1}$$

where λ is a small parameter ($|\lambda| \ll 1$), $h, P_j = O(1)$ ($j = 1, \dots, n$) are given functions and $R_{n+1} = O(\lambda^{n+1})$. Now, if the Hamiltonian $h = H_0$ is *integrable*, for example in the Liouville sense, though the notion will be better specified below. The Hamiltonian (4.1) is said to be *quasi-integrable* or *close to integrable*, and $P(\lambda) = H_\lambda - h$ is called the *perturbation*. It has to be stressed that in many applications the perturbation P does not appear as explicitly ordered in terms of a small parameter λ, but could be split into a leading part, say P_1, plus a remainder, say P_2, the latter being possibly split in turn into a leading part, say P_3, and so on. Moreover, the splitting of P can be different in different regions of the phase space. In such cases the parameter λ can be artificially inserted in the theory as a *tracer* of the ordering, and set to one at the end of the calculations (some authors, in this case, refer to λ as a "bookkeeping" parameter). Another remark concerns the closeness of the perturbed system, defined by H_λ, to the unperturbed one defined by h. What really matters is not, or not so much, the ratio $|P|/|h|$ of the perturbation to the unperturbed Hamiltonian, but the ratio of the respective vector fields, namely $\|X_P\|/\|X_h\|$, in some norm. One understands this by thinking that a constant perturbation does not affect the dynamics independently of its size.

The central idea of Hamiltonian perturbation theory, which goes back to Lagrange and Poincaré and has then be developed by Birkhoff, Bogoliubov and Kolmogorov, consists in looking for a change of variables that removes, completely or partially, the perturbation $P(\lambda)$ from H_λ, up to some pre-fixed order. As will be shown below, the complete removal of the perturbation, already at first order (i.e. the complete removal of P_1) is not possible, in general. What is actually meant by "partial removal of the perturbation up to some pre-fixed order" is clarified by the following definition.

Definition 4.1 (Normal form). *A Hamiltonian K_λ of the form*

$$K_\lambda = h + \sum_{j=1}^{n} \lambda^j S_j + R_{n+1} \ , \tag{4.2}$$

where $\{S_j, h\} = 0$ for any $j = 1, \dots, n$, and $R_{n+1} = O(\lambda^{n+1})$, is said to be in ***normal form to order*** *n* ***with respect to*** *h.*

The Hamiltonian (4.2) is of the form (4.1), but the perturbation terms S_j are first integrals of h, which includes the case of $S_j = 0$ for some j, i.e. of absence of the perturbation term of the order j at hand. From a technical point of view, the aim of Hamiltonian perturbation theory becomes that of finding a suitable change of variables that maps the quasi-integrable Hamiltonian (4.1) into a normal form of the type (4.2). To that purpose, suppose that

1. the Hamiltonian $H_\lambda(x)$ has the form (4.1) for any $x \in D \subset \Gamma$;

2. the unperturbed Hamiltonian $H_0 = h$ is integrable in D, which means that its flow $\Phi_h^t(\xi)$, i.e. the solution of the Hamilton equations $\dot{x} = X_h(x) = (J\nabla H)(x)$ is known for any $\xi \in D$;

3. the flow $\Phi_h^t(\xi)$ is bounded in D, uniformly in time: there exists a constant C such that $\|\Phi_h^t(\xi)\| \leq C$ for any $\xi \in D$, in some suitable norm $\| \ \|$.

Then one looks for a smooth, λ-dependent, λ-close to the identity, canonical change of variables $\mathscr{C}_\lambda$ such that $\tilde{H}_\lambda(y) := H_\lambda\left(\mathscr{C}_\lambda^{-1}(y)\right)$ is in normal form to order n with respect to h, i.e. has the form (4.2):

$$\mathscr{C}_\lambda : x \mapsto y = \mathscr{C}_\lambda(x) = x + O(\lambda) \ \ ; \ \ \tilde{H} = H_\lambda \circ \mathscr{C}_\lambda^{-1} = K_\lambda \ . \tag{4.3}$$

The role of the third hypothesis (boundedness) on Φ_h^t made above will appear below. The λ-closeness to the identity of the change of variables $\mathscr{C}_\lambda$ is necessary in order to match the unperturbed problem as $\lambda \to 0$. Finally, the canonicity of $\mathscr{C}_\lambda$ allows one to deal with the transformation of the Hamiltonian function only, without minding about the consequent deformation of the Poisson structure. However, it has to be stressed that this is a choice, made for the sake of simplicity, and represents an actual restriction in the framework of perturbation theory.

In order to go on with the above program, one has to choose how to build up the canonical transformation $\mathscr{C}_\lambda$. This is very conveniently made by composing the Hamiltonian flows of suitable generating Hamiltonians at λ-dependent times. More precisely, one looks for a canonical transformation of the form

$$\mathscr{C}_\lambda = \Phi_{G_n}^{-\lambda^n} \circ \Phi_{G_{n-1}}^{-\lambda^{n-1}} \circ \cdots \circ \Phi_{G_2}^{-\lambda^2} \circ \Phi_{G_1}^{-\lambda} \ , \tag{4.4}$$

with inverse

$$\mathscr{C}_\lambda^{-1} = \Phi_{G_1}^{\lambda} \circ \Phi_{G_2}^{\lambda^2} \circ \cdots \circ \Phi_{G_{n-1}}^{\lambda^{n-1}} \circ \Phi_{G_n}^{\lambda^n} \ , \tag{4.5}$$

where $\Phi_{G_j}^{\pm\lambda^j}$ is the flow of the Hamiltonian G_j at time $\pm\lambda^j$; notice that the choice of the minus sign in defining the direct transformation (4.4) is made just for formal convenience: one always needs the inverse transformation to compute the transformed Hamiltonian. The n Hamiltonians $G_1, \ldots, G_n$ are called the generating Hamiltonians of the canonical transformation $\mathscr{C}_\lambda$: the latter transformation is completely specified when the n generating functions are completely specified. As a matter of fact, the generating Hamiltonian $G_1, \ldots, G_n$ are the unknowns of the perturbative construction described above: their form is determined order by order. It will turn out that actually, in a very precise sense, there are infinitely many possible sets of n generating functions allowing to set H_λ in normal form. In other words, **the normal form of a given quasi-integrable Hamiltonian is not unique**.

In the sequel, the following notation will be made use of. Given the Hamiltonian h, for any real function F on Γ its time-average $\langle F\rangle_h$ along the flow of h and its deviation from the average $\delta_h F$ are defined by

$$\langle F\rangle_h := \lim_{t\to+\infty} \frac{1}{t}\int_0^t \left(F\circ\Phi_h^s\right)\, \mathrm{d}s \ ; \tag{4.6}$$

$$\delta_h F := F - \langle F\rangle_h \ . \tag{4.7}$$

We will also need the following technical lemmas.

Lemma 4.1. *The time-average $\langle\ \rangle_h$ is invariant with respect to the flow of h, i.e. for any function F one has*

$$\langle F\rangle_h \circ \Phi_h^r = \langle F\rangle_h \ \ \forall r \quad \Leftrightarrow \quad L_h \langle F\rangle_h = 0 \ . \tag{4.8}$$

$\lhd$ PROOF. Two equivalent proofs of the statement are given. The first one starts by writing down explicitly $\langle F\rangle_h \circ \Phi_h^r$ and making use of the group property of the flow, namely $\Phi_h^s \circ \Phi_h^r = \Phi_h^{s+r}$, which yields

$$\langle F\rangle_h \circ \Phi_h^r = \lim_{t\to+\infty}\frac{1}{t}\int_0^t F\circ\Phi_h^{s+r}\mathrm{d}s = \lim_{t\to+\infty}\frac{1}{t}\int_r^{t+r} F\circ\Phi_h^u \mathrm{d}u \ .$$

Now, by splitting $\int_r^{t+r}\mathrm{d}u = \int_r^0 \mathrm{d}u + \int_0^t \mathrm{d}u + \int_t^{t+r}\mathrm{d}u$, observing that the first and the third integral are on bounded intervals, and making use of the boundedness hypothesis on Φ_h^t, the thesis of the Lemma follows. The second proof starts from the chain of identities

$$\frac{\mathrm{d}}{\mathrm{d}s} F\circ\Phi_h^s = L_h F\circ\Phi_h^s = \{F, h\}\circ\Phi_h^s = \{F\circ\Phi_h^s, h\} \ ,$$

where use has been made of Lemma 3.1 first, while the last step follows by the canonicity of the flow, Proposition 3.3, and by the obvious invariance of h with respect to Φ_h^s. Now, integrating the above identity (left and rightmost members) from 0 to t and dividing by t, one gets

$$\frac{F\circ\Phi_h^t - F}{t} = \frac{1}{t}\int_0^t \{F\circ\Phi_h^s, h\}\mathrm{d}s = \left\{\frac{1}{t}\int_0^t F\circ\Phi_h^s \mathrm{d}s, h\right\} \ ,$$

where the second equality follows by the bi-linearity of the Poisson bracket (think of computing the integral as the limit of Riemann sums). The thesis of the lemma, equation (4.8), right form, follows in the limit as $t\to+\infty$, by observing that the left hand side of above identity vanishes in the limit (recall the boundedness of Φ_h^t). $\rhd$

Lemma 4.2. *For any function F, the solution of the equation*

$$L_h G = \delta_h F \quad \Leftrightarrow \quad \{G, h\} = F - \langle F \rangle_h \tag{4.9}$$

is given by

$$G = \mathscr{G} + L_h^{-1} \delta_h F := \mathscr{G} + \lim_{t \to +\infty} \frac{1}{t} \int_0^t (s - t) \left(\delta_h F \circ \Phi_h^s \right) \mathrm{d}s \ , \tag{4.10}$$

where $\mathscr{G}$ is an arbitrary element of $\ker L_h$*, i.e. any function satisfying $\{\mathscr{G}, h\} = 0$.*

◁ PROOF. Consider the left form of equation (4.9), compose both sides with Φ_h^s, multiply by $(s - t)$, integrate withe respect to s from 0 to t and divide by t. This yields

$$\frac{1}{t} \int_0^t (s - t) L_h G \circ \Phi_h^s \mathrm{d}s = \frac{1}{t} \int_0^t (s - t) \delta_h F \circ \Phi_h^s \mathrm{d}s \ .$$

Now make use of Lemma 3.1 on the left hand side of the latter equation, which, upon integrating by parts, implies

$$G = \frac{1}{t} \int_0^t e^{s L_h} G \ \mathrm{d}s + \frac{1}{t} \int_0^t (s - t) \delta_h F \circ \Phi_h^s \mathrm{d}s \ .$$

The thesis of the Lemma follows by taking the limit of the latter equation for $t \to +\infty$, and by observing that $\mathscr{G} := \langle G \rangle_h$, by the Lemma 4.1, is an arbitrary element of $\ker L_h$. ▷

The following Theorem holds.

Theorem 4.1 (Averaging principle)**.** *Consider a quasi integrable Hamiltonian H_λ of the form (4.1) and satisfying the hypotheses 1., 2. and 3. made above. Then*

I) For <u>any</u> choice of the generating Hamiltonians $G_1, \dots, G_n$ defining the canonical transformation (4.4)-(4.5), one has

$$\tilde{H}_\lambda = H_\lambda \circ \mathscr{C}_\lambda^{-1} = h + \sum_{j=1}^{n} \lambda^j \mathscr{P}_j + \mathscr{R}_{n+1} \ , \tag{4.11}$$

where, for $j = 1, \dots, n$ and starting with $F_1 = 0$,

$$\mathscr{P}_j = -L_h G_j + P_j + F_j[h, P_1, \dots, P_{j-1}, G_1, \dots, G_{j-1}] \ ; \tag{4.12}$$

$$\mathscr{R}_{n+1} = \sum_{j \geq n+1} \lambda^j (P_j + F_j[h, P_1, \dots, P_{j-1}, G_1, \dots, G_n]) \ . \tag{4.13}$$

II) The perturbation at order $j = 1, \dots, n$ of the normal form is given by

$$\mathscr{P}_j := S_j = \langle P_j + F_j \rangle_h \ . \tag{4.14}$$

III) The n normalizing, generating Hamiltonians $G_1, \dots, G_n$ are given by

$$G_j = \mathscr{G}_j + L_h^{-1} \delta_h (P_j + F_j) \quad , \ \mathscr{G}_j \in \ker L_h \ . \tag{4.15}$$

(the operators $\langle\ \rangle_h$, δ_h and $L_h^{-1}\delta_h$ are those defined above in (4.6), (4.7) and (4.10)).

$\triangleleft$ PROOF. Let us define $L_j := L_{G_j} = \{\ , G_j\}$ and recall the obvious property $L_j h = -L_h G_j$, to be repeatedly used in the sequel. Making use of the definition (4.5) of $\mathscr{C}_\lambda^{-1}$ and by repeated use of Lemma 3.1, one gets

$$\tilde{H}_\lambda = H_\lambda \circ \mathscr{C}_\lambda^{-1} = e^{\lambda^n L_n} \cdots e^{\lambda^2 L_2} e^{\lambda L_1} \left(h + \lambda P_1 + \lambda^2 P_2 + \cdots + \lambda^n P_n + R_{n+1}\right) \ . \tag{4.16}$$

Let us first write down explicitly the above transformation in the case $n = 2$. By expanding the two exponentials one finds

$$\begin{aligned} \tilde{H}_\lambda &= \left(1 + \lambda^2 L_2 + \dots\right)\left(1 + \lambda L_1 + \lambda^2 L_1^2/2 + \lambda^3 L_1^3/6 + \dots\right)\left(h + \lambda P_1 + \lambda^2 P_2 + \lambda^3 P_3 + \dots\right) = \\ &= h + \lambda\left(L_1 h + P_1\right) + \lambda^2\left(L_2 h + P_2 + L_1 P_1 + L_1^2 h/2\right) + \mathscr{R}_3 \ , \end{aligned}$$

which is of the form (4.11)-(4.13) with

$$\begin{aligned} &\mathscr{P}_1 = -L_h G_1 + P_1 \quad ; \quad F_1 = 0 \ ; \\ &\mathscr{P}_2 = -L_h G_2 + P_2 + F_2 \quad ; \quad F_2 = L_1 P_1 + L_1^2 h/2 \ ; \\ &\mathscr{R}_3 = \lambda^3\left(P_3 + F_3\right) + O(\lambda^4) \quad ; \quad F_3 = L_1 P_2 + L_2 P_1 + L_1^2 P_1/2 + L_1^3 h/6 + L_2 L_1 h \ . \end{aligned}$$

Moreover, any F_j, for $j \geq 4$, depends on h, all the perturbation terms up to order $j-1$, and on the two generating functions G_1, G_2.

Thus, statement I) of the theorem holds for $n = 1, 2$. Let us now suppose that statement holds up to order $m-1$, so that $\tilde{H}_\lambda$ looks in the form (4.11)-(4.13) with $n = m-1$; call the latter $\tilde{H}_\lambda^{(m-1)}$. Then one lets the exponential operator $e^{\lambda^m L_m} = 1 + \lambda^m L_m + O(\lambda^{2m})$ act on the left of such an expression, getting $\tilde{H}_\lambda^{(m)} = e^{\lambda^m L_m} \tilde{H}_\lambda^{(m-1)}$, namely

$$\begin{aligned} \tilde{H}_\lambda^{(m)} &= e^{\lambda^m L_m}\left[h + \sum_{j=1}^{m-1} \lambda^j \mathscr{P}_j + \lambda^m\left(P_m + F_m\right) + \sum_{j \geq m+1} \lambda^j (P_j + F_j)\right] = \\ &= h + \sum_{j=1}^{m-1} \lambda^j \mathscr{P}_j + \lambda^m\left(-L_h G_m + P_m + F_m\right) + \sum_{j \geq m+1} \lambda^j (P_j + F_j') = \\ &= h + \sum_{j=1}^{m} \lambda^j \mathscr{P}_j + \mathscr{R}'_{m+1} \ , \end{aligned}$$

of the form (4.11)-(4.12). By induction I) holds for any $j \geq 1$ and thus up to any pre-fixed n.

In order to prove statements II) and III), let us consider equation (4.12) at a generic order j, $1 \leq j \leq n$. When the generating set of Hamiltonians $G_1, \dots G_n$ is specified, the latter gives the form of $\mathscr{P}_j$. On the other hand, in order to get the transformed Hamiltonian (4.11) in normal form, we have to impose the condition $\mathscr{P}_j := S_j \in \ker L_h$ and solve the equation

$$\boxed{S_j = -L_h G_j + P_j + F_j} \tag{4.17}$$

with respect to two unknowns, namely S_j and G_j (supposing we have solved for $G_1, \dots, G_{j-1}$ at previous orders, which specifies F_j).

Remark 4.1. *Observe that* $\mathscr{P}_j[G_1, \ldots, G_j]$ *has the form (4.12) for any choice of the generating Hamiltonians* $G_1, \ldots, G_j$. *We then look for those special sequences* $\bar{G}_1, \ldots, \bar{G}_j$ *such that* $S_j := \mathscr{P}[\bar{G}_1, \ldots, \bar{G}_j] \in \ker L_h$.

Equation (4.17) is the so-called **homological equation** of Hamiltonian perturbation theory to order j. We first solve it for S_j, and then for G_j. Let us compose both sides of the homological equation (4.17) with Φ_h^s, and take into account that $S_j \circ \Phi_h^s = S_j$, and $L_h G_j \circ \Phi_h^s = \mathrm{d}(G_j \circ \Phi_h^s)/\mathrm{d}s$. Then, integrating with respect to s from 0 to t and dividing by t one gets

$$S_j = -\frac{G_j \circ \Phi_h^t - G_j}{t} + \frac{1}{t}\int_0^t (P_j + F_j) \circ \Phi_h^s \mathrm{d}s \ .$$

Statement II) of the Theorem, equation (4.14), follows in the limit as $t \to +\infty$, given the boundedness of the unperturbed flow. Inserting the expression $S_j = \langle P_j + F_j \rangle_h$ into the homological equation (4.17), one is left with the equation

$$L_h G_j = P_j + F_j - \langle P_j + F_j \rangle_h = \delta_h (P_j + F_j) \ .$$

By Lemma 4.2, its solution is

$$G_j = \mathscr{G}_j + L_h^{-1} \delta_h (P_j + F_j) \ ,$$

namely statement III) of the Theorem. ▷

It can be useful to give the explicit formulas of the normal form Hamiltonian to second order ($n = 2$). After some minor manipulation, one finds $\tilde{H}_\lambda = h + \lambda S_1 + \lambda^2 S_2 + \mathscr{R}_3$, with

$$S_1 = \langle P_1 \rangle_h \ ; \tag{4.18}$$

$$\delta_h G_1 = L_h^{-1} \delta_h P_1 \ ; \tag{4.19}$$

$$S_2 = \langle P_2 \rangle_h + \frac{1}{2} \langle \{\delta_h P_1, \delta_h G_1\} \rangle + \{\langle P_1 \rangle_h , \mathscr{G}\} \ , \tag{4.20}$$

where $\mathscr{G}$ is an arbitrary element of $\ker L_h$.

Exercise 4.1. *Deduce formulas (4.18), (4.19) and (4.20).*

Exercise 4.2. *Show that in the case of* τ*-periodic unperturbed flow, i.e.* $\Phi_h^\tau(x) = \Phi_h^0(x) = x$ *for any* x*, formula (4.6) for the time average of a function and formula (4.10) for the operator* $L_h^{-1}\delta_h$ *simplify to*

$$\langle F \rangle_h = \frac{1}{\tau} \int_0^\tau (F \circ \Phi_h^s) \, \mathrm{d}s \ ; \tag{4.21}$$

$$L_h^{-1} \delta_h F = \frac{1}{\tau} \int_0^\tau s \, (\delta_h F \circ \Phi_h^s) \, \mathrm{d}s \ . \tag{4.22}$$

4.1.1 Quasi-periodic unperturbed flows

In large part of the applications of Hamiltonian perturbation theory, one actually deals with quasi-periodic, or multi-periodic, unperturbed flows, characterized by a certain number d of frequencies $\omega_1, \dots, \omega_d$, including the possibility $d \to +\infty$. This means that the dynamics of the unperturbed system takes place on a d-dimensional torus parametrized by d angles, each of them advancing linearly in time. For example, for mechanical systems with n degrees of freedom, this is the case when the unperturbed system is Liouville-Arnol'd integrable, so that action-angle variables (I, φ) exist such that $h = E(I)$ and E is a function of $d \leq n$ action variables only; the unperturbed flow in this case reads $\Phi_h^t(\varphi, I) = (\varphi + \omega(I)t, I)$, $\omega(I) := \partial E/\partial I$.

In the considered case of quasi-periodic unperturbed motions, in solving the homological equation at any given order, one has to compute the time average of a function of the form

$$(P+F)\circ\Phi_h^s = \sum_{k\in\mathbb{Z}^d} C_k e^{i(k\cdot\omega)s} \ , \tag{4.23}$$

where $\omega := (\omega_1, \dots, \omega_d)$. Its time average is

$$\langle P+F\rangle_h = \sum_{k\in\mathbb{Z}^d} C_k \lim_{t\to+\infty} \frac{1}{t}\int_0^t e^{i(k\cdot\omega)s}\mathrm{d}s = \sum_{\substack{k\in\mathbb{Z}^d:\\ k\cdot\omega=0}} C_k \ . \tag{4.24}$$

Thus, a coefficients C_k survives the averaging iff $k\cdot\omega = k_1\omega_+ \cdots + k_d\omega_d = 0$. If the frequencies $\omega_1, \dots, \omega_d$ are rationally independent, i.e. if $k\cdot\omega = 0$ implies $k = 0$, then $\langle P+F\rangle_h = C_0$. Observe that C_k is always a given function of the point x in the phase space, whereas the frequency vector ω may depend or not on x, which one refers to as the anisochronous and isochronous case, respectively. Now, once one has computed the average (4.24), subtracting the latter from (4.23) gives the deviation

$$\delta_h(P+F)\circ\Phi_h^s = \sum_{\substack{k\in\mathbb{Z}^d:\\ k\cdot\omega\neq0}} C_k e^{i(k\cdot\omega)s} \ . \tag{4.25}$$

Then, the particular solution $\delta_h G = G - \mathscr{G}$ of the homological equation $L_h G = \delta_h(P+F)$ for the generating Hamiltonian G is given by

$$\begin{aligned} \delta_h G &= L_h^{-1}\delta_h(P+F) = \sum_{\substack{k\in\mathbb{Z}^d:\\ k\cdot\omega\neq0}} C_k \lim_{t\to+\infty}\frac{1}{t}\int_0^t (s-t)e^{i(k\cdot\omega)s}\mathrm{d}s = \\ &= \sum_{\substack{k\in\mathbb{Z}^d:\\ k\cdot\omega\neq0}} C_k \left[\frac{1}{ik\cdot\omega} - \lim_{t\to+\infty}\frac{e^{i(k\cdot\omega)t}-1}{t(ik\cdot\omega)^2}\right] = \sum_{\substack{k\in\mathbb{Z}^d:\\ k\cdot\omega\neq0}} \frac{C_k}{ik\cdot\omega} \ . \end{aligned} \tag{4.26}$$

In the case of τ-periodic unperturbed flow, there is only one frequency $\omega_1 := \omega = 2\pi/\tau$ $(d = 1)$. In such a case formulas (4.24) and (4.26) simplify to

$$\langle P+F\rangle_h = C_0 \quad ; \quad \delta_h G = \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{C_k}{ik\omega} \ . \tag{4.27}$$

Exercise 4.3. *Consider an anharmonic oscillator, described by the Hamiltonian*

$$H_\lambda(q,p) = \frac{p^2+\omega^2 q^2}{2} + \lambda\alpha\frac{q^3}{3} + \lambda^2\beta\frac{q^4}{4} \ .$$

Express the latter in terms of the complex Birkhoff variable $z = (\omega q + ip)/\sqrt{2\omega}$ and compute its normal with respect to $h = H_0$, to second order.

Exercise 4.4. *Consider the Henón-Heiles family of Hamiltonian*

$$H_\lambda(q_1,q_2,p_1,p_2) = \frac{p_1^2+\omega_1^2 q_1^2}{2} + \frac{p_2^2+\omega_2^2 q_2}{2} + \lambda C(q_1,q_2) \ ,$$

where C is a homogeneous polynomial of degree 3. Write it in terms of the complex Birkhoff variables $z_j = (\omega_j q_j + ip_j)/\sqrt{2\omega_j}$, $j = 1,2$, and compute its normal form (with respect to $h = H_0$) to first order, discussing its dependence on the frequency ratio ω_2/ω_1.

Exercise 4.5. *Consider the system of three rotators defined by the Hamiltonian*

$$H(\varphi,I) = \frac{I_1^2+I_2^2+I_3^2}{2} + \lambda\left[\cos(\varphi_2-\varphi_1)+\cos(\varphi_3-\varphi_2)+\cos(\varphi_1-\varphi_3)\right] \ ,$$

where $\varphi = (\varphi,\varphi_2,\varphi_3) \in (\mathbb{R}/(2\pi\mathbb{Z}))^3$ and $I = (I_1,I_2,I_3) \in \mathbb{R}^3_+$ are canonically conjugated angle-action variables. Compute the normal form of H_λ with respect to $h = H_0$ to first order, discussing in detail its dependence on the point of the action space $\mathbb{R}^3_+$.

Exercise 4.6. *Consider the nonlinear, non homogeneous Klein-Gordon equation*

$$u_{tt} = -m^2 u + \lambda\left[u_{xx} - gu^3 - V(x)u\right] \ , \tag{4.28}$$

where m and g are real parameters, $V(x)$ is a given function and λ is a small parameter. Here the domain of x can be an interval with periodic boundary conditions or the whole real line.

Check that the equation (4.28) can be written in Hamiltonian form, namely $u_t = \delta H_\lambda/\delta\pi$, $\pi_t = -\delta H_\lambda/\delta u$, where $H_\lambda[u,\pi] = h + \lambda P$, with

$$h[u,\pi] = \int \frac{\pi^2+m^2u^2}{2}\,\mathrm{d}x \ ;$$

$$P[u,\pi] = \int \left[\frac{(u_x)^2}{2} + g\frac{u^4}{4} + V(x)\frac{u^2}{2}\right]\,\mathrm{d}x \ .$$

Perform the change of variables $(u,\pi) \mapsto (\psi,\psi^*)$, where $\psi := (mu + \imath\pi)/\sqrt{2m}$ and $\imath$ denotes the imaginary unit; write the transformed Poisson tensor and the transformed Hamiltonian.

Compute the normal form Hamiltonian to first order with respect to h. Prove that the Hamilton equations corresponding to the normal form Hamiltonian are the Gross-Pitaevskii equation

$$\imath\varphi_t = m\varphi + \frac{\lambda}{m}\left[-\varphi_{xx} + V(x)\varphi + \frac{3g}{2m}|\varphi|^2\varphi\right] \tag{4.29}$$

and its complex conjugate.

4.1.2 Application to quantum mechanics

The formalism of Hamiltonian perturbation theory just developed can be applied to quantum mechanics once one identifies the relevant quantum objects such as Hamiltonian flows and canonical transformations. To such a purpose, we start by observing that unitary transformations of the wave function Ψ, the unknown of the Schrödinger equation, are canonical transformations of the latter equation. Indeed, given *any* unitary operator $\hat{\mathsf{U}}$ independent of time, and defining $\Psi' := \hat{\mathsf{U}}^\dagger\Psi$, $\hat{\mathsf{H}}' = \hat{\mathsf{U}}^\dagger\hat{\mathsf{H}}\hat{\mathsf{U}}$, one has

$$\imath\hbar\Psi_t = \hat{\mathsf{H}}\Psi \iff \imath\hbar\Psi'_t = \hat{\mathsf{H}}'\Psi' \ , \tag{4.30}$$

the equation on the right being identical in form to that on the left. Thus, in the perturbative context where $\hat{\mathsf{H}} = \hat{\mathsf{h}} + \lambda\hat{\mathsf{V}} + \cdots$, one can try to remove the leading order perturbation $\hat{\mathsf{V}}$ (and then the higher order contributions as well) by looking for a particular unitary operator $\hat{\mathsf{U}}_\lambda$ that conjugates $\hat{\mathsf{H}}$ to its normal form, to leading order. In particular, by analogy with the classical case one looks for a unitary operator $\hat{\mathsf{U}}_\lambda$ that is the Schrödinger flow at time λ of some unknown Hamiltonian (Hermitian) operator $\hat{\mathsf{G}}$, namely

$$\hat{\mathsf{U}}_\lambda = e^{-\imath\lambda\hat{\mathsf{G}}/\hbar} \tag{4.31}$$

(we hare recall that the solution of the Schrödinger equation $\imath\hbar\Psi_t = \hat{\mathsf{H}}\Psi$ is formally given by $\Psi(t) = e^{-\imath t\hat{\mathsf{H}}/\hbar}\Psi(0)$). To any Hermitian operator $\hat{\mathsf{G}}$ one can associate the operator

$$\hat{\mathsf{L}}_{\hat{\mathsf{G}}} := -\frac{\imath}{\hbar}\left[\ ,\hat{\mathsf{G}}\right] \ , \tag{4.32}$$

i.e. the quantum Lie derivative associated to $\hat{\mathsf{G}}$. One easily proves the following

Lemma 4.3. *For any pair of Hermitian operators $\hat{\mathsf{F}}$ and $\hat{\mathsf{G}}$ independent of λ, one has*

$$e^{+\imath\lambda\hat{\mathsf{G}}/\hbar}\hat{\mathsf{F}}e^{-\imath\lambda\hat{\mathsf{G}}/\hbar} = e^{\lambda\hat{\mathsf{L}}_{\hat{\mathsf{G}}}}\hat{\mathsf{F}} \ . \tag{4.33}$$

$\lhd$ PROOF. Define $\hat{\mathsf{F}}(\lambda)$ the left hand side of (4.33) and take its derivative with respect to λ, getting

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\hat{\mathsf{F}}(\lambda) = \frac{\imath}{\hbar}\left(\hat{\mathsf{G}}\hat{\mathsf{F}}(\lambda) - \hat{\mathsf{F}}(\lambda)\hat{\mathsf{G}}\right) = \hat{\mathsf{L}}_{\hat{\mathsf{G}}}\hat{\mathsf{F}}(\lambda) \ .$$

The latter differential equation can be formally integrated with the initial condition $\hat{\mathsf{F}}(0) = \hat{\mathsf{F}}$, to yield (4.33).$\rhd$

In particular, from the latter Lemma it follows that $\hat{\mathsf{F}}$ is invariant with respect to the flow of $\hat{\mathsf{G}}$ iff $[\hat{\mathsf{F}},\hat{\mathsf{G}}] = 0$ (prove it). Now, supposing that $\hat{\mathsf{H}}_\lambda = \hat{\mathsf{h}} + \lambda\hat{\mathsf{V}}$ and transforming it by $\hat{\mathsf{U}}_\lambda = e^{-\imath\lambda\hat{\mathsf{G}}/\hbar}$, one gets

$$\hat{\mathsf{H}}'_\lambda := \hat{\mathsf{U}}^\dagger_\lambda\hat{\mathsf{H}}\hat{\mathsf{U}}_\lambda = e^{\lambda\hat{\mathsf{L}}_{\hat{\mathsf{G}}}}\left(\hat{\mathsf{h}} + \lambda\hat{\mathsf{V}}\right) = \hat{\mathsf{h}} + \lambda\left(\hat{\mathsf{V}} + \hat{\mathsf{L}}_{\hat{\mathsf{G}}}\hat{\mathsf{h}}\right) + O(\lambda^2) \ .$$

One now requires that the latter expression be in normal form with respect to $\hat{\mathsf{h}}$ to first order, i.e. that $\hat{\mathsf{H}}'_\lambda = \hat{\mathsf{h}} + \lambda\hat{\mathsf{S}} + O(\lambda^2)$, with $\hat{\mathsf{S}}$ invariant with respect to the flow of $\hat{\mathsf{h}}$: $\hat{\mathsf{L}}_{\hat{\mathsf{h}}}\hat{\mathsf{S}} = 0$. Thus, taking into account that $\hat{\mathsf{L}}_{\hat{\mathsf{G}}}\hat{\mathsf{h}} = -\hat{\mathsf{L}}_{\hat{\mathsf{h}}}\hat{\mathsf{G}}$, one writes down the (first order) quantum homological equation

$$\hat{\mathsf{S}} = \hat{\mathsf{V}} - \hat{\mathsf{L}}_{\hat{\mathsf{h}}}\hat{\mathsf{G}} \tag{4.34}$$

and solves it for $\hat{\mathsf{S}}$ and $\hat{\mathsf{G}}$ exactly as done in the classical case. The result is

$$\hat{\mathsf{S}} = \left\langle \hat{\mathsf{V}} \right\rangle_{\hat{\mathsf{h}}} = \lim_{t\to+\infty} \frac{1}{t}\int_0^t e^{\frac{\imath}{\hbar}s\hat{\mathsf{h}}}\hat{\mathsf{V}}e^{-\frac{\imath}{\hbar}s\hat{\mathsf{h}}}\mathrm{d}s \ ; \tag{4.35}$$

$$\hat{\mathsf{G}} = \hat{\mathscr{G}} + \lim_{t\to+\infty} \frac{1}{t}\int_0^t (s-t)e^{\frac{\imath}{\hbar}s\hat{\mathsf{h}}}\left(\hat{\mathsf{V}} - \left\langle \hat{\mathsf{V}} \right\rangle_{\hat{\mathsf{h}}}\right)e^{-\frac{\imath}{\hbar}s\hat{\mathsf{h}}}\mathrm{d}s \ . \tag{4.36}$$

Exercise 4.7. *Prove formulas (4.35) and (4.36).*

It follows that the averaging principle, Theorem 4.1, holds in quantum mechanics to all orders, with a formulation that, up to the replacement of the Poisson bracket with the commutator divided by $\imath\hbar$, is completely analogous to the classical one. In particular, to first order, the perturbation in normal form is the time average of the perturbation along the flow of the unperturbed system.

Exercise 4.8. *Consider the classical anharmonic oscillator defined by the Hamiltonian*

$$H(x,p) = \frac{p^2}{2m} + \frac{kx^2}{2} + \mu\frac{x^4}{4} \ . \tag{4.37}$$

Quantize the problem, introduce the quantum analogue of the Birkhoff variables and compute the normal form to first order with respect to the quadratic Hamiltonian. Compare the result with the corresponding classical one.