ANALYTICAL MECHANICS

Lectures Notes 2019/2020

A. Ponno

December 19, 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). As sketched below, for a certain class of constrained problems, the reaction R can be computed as a function of X, $\dot{X}$ and t. However, such a computation turns out to be extremely involved in practice. As a matter of fact, it was such a difficulty 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 (bilateral) holonomic constraints the mechanical

[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.

system is supposed to move on a given L-dimensional manifold M (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). Let us denote by T_XM the L-dimensional tangent plane, or space, at each point $X\in M$. The fundamental hypothesis that allows to solve the problem for holonomic constraints is that of ideal constraints (the so-called D'Alambert principle), namely

$$R\cdot\xi=0\ \ \forall\xi\in T_XM\ . \tag{1.2}$$

Indeed, condition (1.2) means that the reaction R pointwise belongs to the normal space N_XM at $X\in M$, which is equivalent to the requirement

$$R=\sum_{r=1}^{N-L}c_r(X,t)\nabla_X\Phi_r(X,t)\ , \tag{1.3}$$

where the Φ_r are the $N-L$ components of Φ, whereas the c_r are unknown coefficients[2]. 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

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

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 (1.3)). Finally, one can reduce the dimension of the Newton equations from N to L. Such an approach obviously reveals rather involved in practice.

The alternative Lagrange approach works as follows. Under the usual hypotheses of the implicit function theorem the constraint manifold M is (locally) described 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 free, or generalized, or Lagrangian coordinates, and the L vectors $\partial X/\partial q_i$ are linearly independent and span the tangent space T_XM. Condition (1.2) is then locally expressed by

$$R\cdot\frac{\partial X}{\partial q_i}=0\ \ \forall i=1,\dots,L\ \ ,\ \forall t\ . \tag{1.4}$$

Upon scalar multiplication of the Newton equation $M\ddot X=F+R$ by $\partial X/\partial q_i$ the, due to condition (1.4), 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

$$M\ddot X\cdot\frac{\partial X}{\partial q_i}=Q_i\ \ ;\ \ Q_i:=F\cdot\frac{\partial X}{\partial q_i}\ . \tag{1.5}$$

[2]One has $\Phi_r(X,t)=0$ for any $r=1,\dots,N-L$. Let $s\mapsto\gamma(s)$ a smooth curve on M such that $\gamma(0)=X$; then $\xi:=\gamma'(0)\in T_XM$. The identity $\Phi_r(\gamma(s),t)=0$ holds for any s in some interval containing $s=0$, which implies

$$0=\frac{\mathrm{d}}{\mathrm{d}s}\Phi_r(\gamma(s),t)\Big|_{s=0}=\nabla_X\Phi_r(X,t)\cdot\xi\ .$$

The generalized force Q_i on the right hand side 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

$$\frac{\partial X}{\partial q_i} = \frac{\partial \dot{X}}{\partial \dot{q}_i} \quad ; \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial X}{\partial q_i} = \frac{\partial \dot{X}}{\partial q_i} \ . \tag{1.6}$$

Defining the kinetic energy $K := (\dot{X} \cdot M\dot{X})/2$ of the system, one easily shows that the equations (1.5) take on the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial K}{\partial \dot{q}} - \frac{\partial K}{\partial q} = Q \ , \tag{1.7}$$

to be meant by components. This is the most general form of Lagrange equations. In the case of conservative force F, i.e. if a function $U(X,t)$ exists such that $F = -\nabla_X U$, one easily checks that the Lagrange equations (1.7) read

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathscr{L}}{\partial \dot{q}} - \frac{\partial \mathscr{L}}{\partial q} = 0 \ , \tag{1.8}$$

where

$$\mathscr{L}(q,\dot{q},t) := K(q,\dot{q},t) - U(q,t) \tag{1.9}$$

is the Lagrange function, or Lagrangian of the system. In the definition (1.9) we have set, with abuse of notation

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

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.

1.2 General properties of Lagrange equations

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathscr{L}}{\partial \dot{q}} - \frac{\partial \mathscr{L}}{\partial q} = 0 \ , \tag{1.10}$$

with $\mathscr{L} = K - U$ 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.*

1. The L-equations are left invariant in form by any (possibly) time dependent (inverse) change of coordinates
$$q = q(Q,t) \;\; ; \;\; \dot{q} = \frac{\partial q}{\partial Q}\dot{Q} + \frac{\partial q}{\partial t} \; . \tag{1.11}$$
Such a property is directly checked by showing that the left hand side $\lambda(\mathscr{L})$ of the Lagrange equations (1.8) and its transformed Λ via the transformation (1.11) are related by $\Lambda = (\partial q/\partial Q)^T\lambda$.

2. For any constant $c \neq 0$ and any function $F(q,t)$ whose total derivative with respect to time is denote by $\dot{F} = \sum_j (\partial F/\partial q_j)\dot{q}_j + \partial F/\partial t$, the Lagrangians $\mathscr{L}$ and $\mathscr{L}' := c\mathscr{L} + \dot{F}$ are equivalent, in the sense that their associated L-equations are the same. Such a property is easily checked by noting that the L-equations are linear in $\mathscr{L}$ and that $\lambda(\dot{F}) \equiv 0$. The change of Lagrangian $\mathscr{L} \to \mathscr{L}'$ is referred to as a *gauge transformation*.

3. If $\partial\mathscr{L}/\partial t = 0$ then the function
$$\mathscr{H}(q,\dot{q}) := \frac{\partial \mathscr{L}}{\partial \dot{q}} \cdot \dot{q} - \mathscr{L}(q,\dot{q}) \tag{1.12}$$
is a constant of motion, or first integral, of the system (i.e. a function whose value is preserved along the solutions of the L-equations). This is directly checked by showing that $\dot{\mathscr{H}}|_{\lambda=0} = 0$. The function $\mathscr{H}$ is known as the Jacobi first integral which, in the conservative mechanical case, $\mathscr{L} = K - U$, is the total energy of the system: $\mathscr{H} = K + U$.

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

4. Consider a one-parameter group of coordinate transformations $q \mapsto Q = \Phi^s(q)$, $s \in \mathbb{R}$, such that $\Phi^0(q) = q$. The group operation $\circ$ here is the composition of transformations, namely $\Phi^s \circ \Phi^r = \Phi^{s+r}$ $\forall s, r \in \mathbb{R}$, which is easily checked to display the three fundamental group properties.[3]. The one-parameter (commutative) group $\{\Phi^s\}_{s\in\mathbb{R}}$ is said to be a symmetry group of (or admissible for) the Lagrangian system defined by $\mathscr{L}(q,\dot{q},t)$ if it leaves the Lagrangian invariant, namely
$$\mathscr{L}\left(\Phi^s(q), \frac{\mathrm{d}}{\mathrm{d}t}\Phi^s(q), t\right) = \mathscr{L}(q,\dot{q},t) \;\; \forall s \in \mathbb{R} \; . \tag{1.13}$$

[3] We recall that a group is a pair set-operation $(G,\circ)$, where $\circ : G \times G \to G$ is an associative application with a unit element e and inverse element g^{-1} of any element $g \in G$. More precisely:

(a) $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$ $\forall g_1, g_2, g_3 \in G$;

(b) $\exists e \in G$ such that $g \circ e = e \circ g = g$ $\forall g \in G$;

(c) $\forall g \in G$ $\exists g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.

A commutative or Abelian group is such that $g_1 \circ g_2 = g_2 \circ g_1$ $\forall g_1, g_2 \in G$.

Let $u(q) := \partial\Phi^s(q)/\partial s|_{s=0}$ be the generator vector field of the group. Then, the Nöther theorem holds:

Theorem 1.1 (Nöther)**.** *Let $\{\Phi^s\}$ be a one parameter symmetry group of the Lagrangian $\mathscr{L}$ (in the sense (1.13)). Then $J_u := u \cdot \partial\mathscr{L}/\partial\dot{q}$ is a first integral.*

The Theorem is easily proven by showing that

$$0 = \frac{\partial}{\partial s}\mathscr{L}\left(\Phi^s(q), \frac{\mathrm{d}}{\mathrm{d}t}\Phi^s(q), t\right)\Big|_{s=0}\Big|_{\lambda=0} = \frac{\partial\mathscr{L}}{\partial q}\cdot u + \frac{\partial\mathscr{L}}{\partial\dot{q}}\cdot\dot{u}\Big|_{\lambda=0} = \frac{\mathrm{d}J_u}{\mathrm{d}t} \ . \tag{1.14}$$

Observe that mechanical Lagrangians are quadratic forms 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 $\partial\mathscr{L}/\partial q_1 = 0$. In this case the symmetry group is $\Phi^s(q) = (q_1+s, q_2, \ldots, q_L)$, and the quantity $p_1 := \partial\mathscr{L}/\partial\dot{q}_1$ is constant, as also follows directly from the L-equations. We recall that if R over L coordinates are ignorable, then the dynamics of the problem is determined by the Routh Lagrangian depending on $L - R$ degrees of freedom[4]. One can prove that the case of one ignorable coordinate ($R = 1$) is the fundamental one: *if Φ^s is a one-parameter group of symmetry for $\mathscr{L}$, then there exists a change to new coordinates such that one of them, say the first one, is ignorable.* Such a Theorem is proven by exploiting the transformation that rectifies the vector field u generating the group Φ^s. The inverse of such a transformation is explicitly given by

$$q = \Phi^{Q_1}(\sigma(Q_2, \ldots, Q_L)) \ , \tag{1.15}$$

where $\sigma : \mathbb{R}^{L-1} \to \mathbb{R}^L$ defines an arbitrary surface of co-dimension 1 transversal to $u(\sigma)$ in the configuration space, whose equation is $Q_1 = 0$, i.e., by (1.15), $q = \sigma(Q_2, \ldots, Q_L)$. The transversality condition is expressed

$$\det\left(\frac{\partial q}{\partial Q}\right)\Big|_{Q_1=0} = \det\left(u(\sigma), \frac{\partial\sigma}{\partial Q_2}, \ldots, \frac{\partial\sigma}{\partial Q_L}\right) \neq 0 \ . \tag{1.16}$$

Due to condition (1.13), when the transformation (1.15) is plugged into the Lagrangian L, the result is independent of Q_1 (it depends however on $\dot{Q}_1$, since Q_1 depends in turn on t). One can alternatively compute how $J_u = u\cdot\partial\mathscr{L}/\partial\dot{q}$ transforms under (1.16). Since

$$\frac{\partial\mathscr{L}}{\partial\dot{Q}_j} = \sum_j \frac{\partial\mathscr{L}}{\partial\dot{q}_i}\frac{\partial\dot{q}_i}{\partial\dot{Q}_j} = \sum_j \frac{\partial\mathscr{L}}{\partial\dot{q}_i}\frac{\partial q_i}{\partial Q_j}$$

then

$$J_u = u(q(Q))\cdot\left(\frac{\partial q}{\partial Q}\right)^{-T}\frac{\partial\mathscr{L}}{\partial\dot{Q}} = \left[\left(\frac{\partial q}{\partial Q}\right)^{-1} u(q(Q))\right]\cdot\frac{\partial\mathscr{L}}{\partial\dot{Q}} = \frac{\partial\mathscr{L}}{\partial\dot{Q}_1} \ ,$$

[4] The Routh reduced Lagrangian is $\mathscr{R} = \mathscr{L} - \sum_r p_r\dot{q}_r|_{\dot{q}_r=f_r(q,p)}$, where $p_r := \partial\mathscr{L}/\partial\dot{q}_r$, the sum $\sum_r$ ranges over the R ignorable coordinates, and the relations $\dot{q}_r = f_r(q,p)$ are obtained by inverting the R relations $p_r := \partial\mathscr{L}/\partial\dot{q}_r$. Such relations are finally used to reconstruct the complete motion after the L-equations associated to $\mathscr{R}$ are solved.

the last step being due to the fact that the vector field u is rectified by the transformation (1.15)[5]. One is then naturally led to ask whether in the presence of two one-parameter symmetry groups $\{\Phi^s\}$ and $\{\Psi^r\}$ it is still possible to find a transformation to new coordinates such that two of them are ignorable at the same time. It turns out that such a transformation exists iff $\Phi^s \circ \Psi^r = \Psi^r \circ \Phi^s$ for all $s, r \in \mathbb{R}$.

Exercise 1.2. *Consider the Lagrangian of a particle of mass m and charge q moving in a given electromagnetic field, namely*

$$\mathscr{L}(x, \dot{x}, t) = m\frac{|\dot{x}|^2}{2} + \frac{q}{c}A(x,t)\cdot\dot{x} - q\phi(x,t) \ , \tag{1.17}$$

being A and ϕ the vector and scalar potential, respectively, and c the velocity of light.

1. *Show that the Lagrange equation read*

$$m\ddot{x} = q\left(E + \frac{1}{c}\dot{x}\wedge B\right) \ , \tag{1.18}$$

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

$$E := -\nabla_x\phi - \frac{1}{c}\frac{\partial A}{\partial t} \quad ; \quad B := \nabla_x \wedge A \ . \tag{1.19}$$

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

$$A \mapsto A' = A + \nabla\chi \quad ; \quad \phi \mapsto \phi' = \phi - \frac{1}{c}\frac{\partial\chi}{\partial t} \ , \tag{1.20}$$

so that the Lagrange equation (1.18) is invariant as well.

3. *Show that under the gauge transformation (1.20) $\mathscr{L} \mapsto \mathscr{L}' = \mathscr{L} + \dot{F}$, 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*

$$\mathscr{H}(x, \dot{x}) = \frac{m|\dot{x}|^2}{2} + q\phi(x) \ , \tag{1.21}$$

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

[5] Consider the system of ODEs $\dot{q} = u(q)$. A transformation $q = g(Q)$ rectifies the vector field u if the system of ODEs in the new coordinates reads $\dot{Q} = \hat{e}_1 := (1, 0, \dots, 0)^T$, for example. Now, since $g'(Q)\dot{Q} = u(g(Q))$, then the rectification condition reads $(g'(Q))^{-1}u(g(Q)) = \hat{e}_1$. The latter relation can also be written $g'(Q)\hat{e}_1 = u(g(Q))$, or $\partial g/\partial Q_1 = u(g(Q_1, \dots, Q_L))$, which justifies the definition (1.15). Notice that, by definition, given a one parameter group $\{\Phi^s\}$, one can associate to it the generator $u = \mathrm{d}\Phi^s/\mathrm{d}s|_{s=0}$. Viceversa, given the system of ODEs $\dot{q} = u(q)$, its flow is the one parameter group generated by u. Indeed, one has

$$\frac{\mathrm{d}\Phi^s(Q)}{\mathrm{d}s} = \lim_{h\to 0}\frac{\Phi^{s+h}(Q) - \Phi^s(Q)}{h} = \lim_{h\to 0}\frac{\Phi^h(\Phi^s(Q)) - \Phi^s(Q)}{h} = \frac{\mathrm{d}\Phi^h(\Phi^s(Q))}{\mathrm{d}h}\Big|_{h=0} = u(\Phi^s(Q)) \ .$$

Exercise 1.3. *Consider the case of a particle in a constant uniform magnetic field $B = B_0 e_z$, with vector potential $A(x) = B_0(e_z \wedge x)/2$, and plane scalar potential ϕ that is a function of $x^2 + y^2$. Find the conservation laws due to translation and rotation symmetry.*

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* $A_{\mathscr{L}} : C_{a,b} \to \mathbb{R}$ is defined by

$$A_{\mathscr{L}}[q] := \int_{t_1}^{t_2} \mathscr{L}(q(t), \dot{q}(t), t) \; \mathrm{d}t \; . \tag{1.22}$$

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 $A_{\mathscr{L}}$ in $q \in C_{a,b}$ with increment $\delta q \in C_{0,0}$ is defined by

$$\delta A_{\mathscr{L}} := \frac{\mathrm{d}}{\mathrm{d}\epsilon} A_{\mathscr{L}}[q + \epsilon \delta q]\Big|_{\epsilon=0} = \lim_{\epsilon \to 0} \frac{A_{\mathscr{L}}[q + \epsilon \delta q] - A_{\mathscr{L}}[q]}{\epsilon} \; . \tag{1.23}$$

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 $\delta A_{\mathscr{L}}$ would be $\mathrm{d}A_{\mathscr{L}}[q]h := \mathrm{d}A_{\mathscr{L}}[q + \epsilon h]/\mathrm{d}\epsilon|_{\epsilon=0}$, 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 (1.23), a simple calculation shows that

$$\delta A_{\mathscr{L}} = \int_{t_1}^{t_2} \left(\frac{\partial \mathscr{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}} \right) \cdot \delta q \; \mathrm{d}t \; . \tag{1.24}$$

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

Proposition 1.1.

$$\delta A_{\mathscr{L}} = 0 \;\; \forall \delta q \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}} - \frac{\partial \mathscr{L}}{\partial q} = 0 \; . \tag{1.25}$$

$\lhd$ 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 $\delta A_{\mathscr{L}} \neq 0$. $\rhd$

In words, the critical points of the action $A_{\mathscr{L}}$ 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 1.4. *For the harmonic oscillator with unit frequency $\mathscr{L} = (\dot{q}^2 - q^2)/2$, 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 1.5. *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 $c\mathscr{L} + \dot{F}$ (for any constant $c \neq 0$ and any function $F(q, t)$) is $cA_{\mathscr{L}} + \Delta F$, 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 $A_{\mathscr{L}'}$ associated to the transformed Lagrangian $\mathscr{L}'(Q, \dot{Q}, t) := \mathscr{L}(q(Q, t), \dot{q}(Q, t), t)$ are the solutions of the L-equations in the new variables.

In the autonomous case, i.e. $\partial\mathscr{L}/\partial t = 0$, the flow of the L-equations preserves the Jacobi integral $\mathscr{H} = p \cdot \dot{q} - \mathscr{L}$, where $p := \partial\mathscr{L}/\partial\dot{q}$. 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

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

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.

Proposition 1.2.

$$\delta A_E = 0 \ \ \forall \delta q \quad \Longleftrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathscr{L}}{\partial \dot{q}} - \frac{\partial \mathscr{L}}{\partial q} = 0 \ . \tag{1.26}$$

1.4 Functional differentiation

Let us consider the space $C^{(n-1)}_{a_0 b_0,\ldots a_{n-1}b_{n-1}}$ of curves $[t_1, t_2] \ni t \mapsto q(t) \in \mathbb{R}^L$ such that the (vector) values of $q(t)$ and of its derivatives of any order up to $n-1$ included are fixed:

$$\frac{\mathrm{d}^j}{\mathrm{d}t^j} q(t)\Big|_{t=t_1} = a_j \quad , \quad \frac{\mathrm{d}^j}{\mathrm{d}t^j} q(t)\Big|_{t=t_2} = b_j \ , \ j = 0, \ldots, n-1 \ . \tag{1.27}$$

Obviously $C^0_{a_0 b_0} = C_{ab}$, the space previously considered in the formulation of the Hamilton principle, with $a = a_0$ and $b = b_0$.

Let us consider a functional $F_{\mathscr{F}} : C^{(n-1)}_{a_0 b_0,\ldots a_{n-1}b_{n-1}} \to \mathbb{R}$

$$F_f[q] := \int_{t_1}^{t_2} f\left(q(t), q^{(1)}(t), \ldots, q^{(n)}(t), t\right) \mathrm{d}t \tag{1.28}$$

defined through a function $f : \mathbb{R}^{(n+1)L} \times \mathbb{R} \to \mathbb{R}$. Here the shorthand notation adopted is

$$q^{(j)}(t) := \frac{\mathrm{d}^j}{\mathrm{d}t^j} q(t) \quad , \ j = 0, \ldots, n \ ,$$

where $q^{(0)} = q$ is understood. The action functional $A_{\mathscr{L}}$ defined by a given Lagrangian $\mathscr{L}$ is a particular example of such a functional in the case $n = 1$ (a Lagrangian is assumed to depend on q and $\dot{q}$ and not on $\ddot{q}$ and higher order derivatives). Let δq be any increment curve in the space $C^{(n-1)}_{00,\ldots 00}$, with all the derivatives equal to zero up to order $n-1$ included. Then, by repeated use of integration by parts one easily gets

$$\delta F_f = \int_{t_1}^{t_2} \frac{\delta F_f}{\delta q} \cdot \delta q \ \mathrm{d}t \ , \tag{1.29}$$

where

$$\frac{\delta F_f}{\delta q} := \frac{\partial f}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial f}{\partial \dot{q}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial f}{\partial \ddot{q}} - \cdots = \sum_{k=0}^{n} (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}t^k}\frac{\partial f}{\partial q^{(k)}} \tag{1.30}$$

is the so-called *functional gradient of F_f with respect to q*; the ith component of such a quantity, denoted by $\delta F_f / \delta q_i$ is called functional derivative of F_f with respect to q_i. The notation is lent from differential calculus of functions of real variables, where

$$\mathrm{d}F = \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(x + \epsilon h)\Big|_{\epsilon=0} = \nabla F(x) \cdot h \ ,$$

where we notice that the gradient is the object multiplying the increment h with respect to the Euclidean scalar product "$\cdot$". The similarity to such a case becomes even more evident if the space $L_2([t_1, t_2]; \mathbb{R}^L)$ is introduced, namely the space of curves in $\mathbb{R}^L$ defined in the interval $[t_1, t_2]$, such that $\int_{t_1}^{t_2} |x(t)|^2 \mathrm{d}t < +\infty$ and endowed with the scalar product

$$\langle x, y \rangle_{L_2([t_1,t_2];\mathbb{R}^L)} := \int_{t_1}^{t_2} x(t) \cdot y(t) \ \mathrm{d}t \ , \tag{1.31}$$

where $x, y : [t_1, t_2] \to \mathbb{R}^L$ and the $\cdot$ denotes the ordinary Euclidean scalar product. One easily checks that the fundamental properties of the scalar product hold for (1.31): it is a positive, symmetric and bilinear form. The Euclidean space L_2 just defined turns out to be complete with respect to the norm $\|x\| := \sqrt{\langle x, x \rangle}$ induced by the scalar product, so that it is a Hilbert space. One can now rewrite (1.30) as follows

$$\delta F_f = \left\langle \frac{\delta F_f}{\delta q}, \delta q \right\rangle_{L_2([t_1,t_2];\mathbb{R}^L)} , \tag{1.32}$$

which completely resembles the case of functions of real variables. For such a reason the functional gradient is also called L_2-gradient, which explicitly makes reference to the relevant scalar product in the problem: $\delta F_f / \delta q = \nabla_{L_2} F_f$.

The same formula (1.30) for the functional gradient holds in the case of a functional $F_f : C_T^{per} \to \mathbb{R}$ defined on the space C_T^{per} of T-periodic curves in $\mathbb{R}^L$, namely curves $q : \mathbb{R}/T\mathbb{Z} \to \mathbb{R}^L$. In such a case one can choose the increments $\delta q \in C_T^{per}$ and perform the integral $\int f \mathrm{d}t$ on any period of length T.

More in general, one can consider functionals depending on functions of more than one independent variable. For example, let us consider the functional

$$F_f[u] = \int_D f(u, u_x, u_y, x, y)\ \mathrm{d}x\mathrm{d}y\ , \tag{1.33}$$

where $u : D \to \mathbb{R}$ is a function of two variables. Consider the case where $D = [0, p] \times [0, q]$ is a rectangle with sides of length p and q and either $u|_{\partial D} = 0$, or u is p periodic in the x direction and q periodic in the y direction, which can be also written as $D = \mathbb{R}/(p\mathbb{Z}) \times \mathbb{R}/(q\mathbb{Z})$. In both cases the increments δu satisfy the same boundary conditions of the functions u. One then easily proves that the functional derivative of F_f is given by the formula

$$\frac{\delta F_f}{\delta u} = \frac{\partial f}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial u_x} - \frac{\mathrm{d}}{\mathrm{d}y}\frac{\partial f}{\partial u_y}\ . \tag{1.34}$$

1.5 Beaded string and its continuum limit I: Lagrangian formulation

Consider a system of $N-1$ identical moving mass points, of mass m, in the x, y plane, the jth mass moving along the vertical line at the fixed abscissa $x_j = ja$, $j = 1, \dots, N-1$. The jth mass is connected to its nearest neighbors $j-1$ and $j+1$ by an ideal spring of elastic constant γ, and it is connected to the the point $(x_j, 0)$ by an ideal spring of elastic constant k. The first γ-spring on the left is connected to the the origin$(0, 0)$, whereas the last γ-spring on the right is connected to the the point $(L, 0)$, where $L = Na$. Such a system represents a kind of beaded string with masses concentrated in $N-1$ points.

1.5.1 Solution of the discrete problem

We are interested in solving explicitly the equations of motion of the system first, and then in performing the continuum limit $a \to 0$, $N \to \infty$ such that $L = Na$ is kept fixed, everything being done within the Lagrangian formalism. Such a study is divided into the following steps.

1. Show that the Lagrangian of the system is
$$\mathscr{L}_N = \sum_{j=1}^{N-1} \frac{m\dot{y}_j^2}{2} - \sum_{j=0}^{N-1} \left[k\frac{y_j^2}{2} + \gamma \frac{(y_{j+1} - y_j)^2}{2} \right] - \gamma \frac{Na^2}{2} \tag{1.35}$$
with fixed-ends boundary conditions $y_0(t) = 0 = y_N(t)$ for any t.

2. Show that the L-equations are
$$m\ddot{y}_j = -ky_j + \gamma(y_{j+1} + y_{j-1} - 2y_j) \ , \quad j = 1, \dots, N-1 \ , \tag{1.36}$$
always taking in mind that $y_0 = 0 = y_N$.

3. Solve the equations of motion (1.36) by separation of variables, looking for solutions of the form
$$y_j(t) = c(t)z^j \ , \tag{1.37}$$
where c is an unknown function of time and z is a parameter. Start to look for space periodic solutions of period $2N$, namely $y_{j+2N} = y_{2N}$ for any $j \in \mathbb{Z}$. Show that there exist $2N$ independent complex solutions of the form (1.37), namely
$$y_j^{(s)}(t) = c_s(t)e^{\imath \frac{2\pi sj}{2N}} \ , \quad s = 0, 1, \dots, 2N-1 \ , \tag{1.38}$$
where c_s is the solution of
$$\ddot{c} = -\omega_s^2 c \ , \tag{1.39}$$
and
$$\omega_s^2 = \frac{k}{m} + 4\frac{\gamma}{m} \sin\left(\frac{\pi s}{2N}\right) \ , \quad s = 0, 1, \dots, 2N-1 \ , \tag{1.40}$$
is the dispersion relation of the system (frequency vs. mode number).

4. Show that a $2N$-periodic, odd initial condition satisfying $y_{j+2N} = y_j$ and $y_{-j} = y_j$ for any $j \in \mathbb{Z}$ at $t = 0$, and the same for $\dot{y}_j$, evolves according to equation (1.36) preserving such a property for any t. In particular, show that this implies that the problem with fixed ends is a sub-case of the periodic one (consider what happens from $-N$ to N).

5. By exploiting the linearity of the equations (1.36), show that the solutions (1.38) can be written in the real form. Then, imposing $y_0^{(s)} = 0 = y_N^{(s)}$, show that the $N-1$ independent solutions of the equations (1.36) with fixed ends are
$$y_j^{(s)}(t) = q_s(t)\varphi_s(j) \ , \quad \varphi_s(j) := \sqrt{\frac{2}{N}} \sin\left(\frac{\pi js}{N}\right) \ , \tag{1.41}$$
where here and in (1.40) $s = 1, \dots, N-1$, whereas $q_s(t)$ is a real solution of the harmonic oscillator equation (1.39).

6. Show that the functions $\varphi_s(j)$ constitute an orthonormal set, namely they satisfy

$$\langle \varphi_r, \varphi_s \rangle := \sum_{j=1}^{N-1} \varphi_r(j)\varphi_s(j) = \delta_{r,s} \ , \tag{1.42}$$

where $\delta_{r,s}$ denotes the Kronecker delta whose value is 1 if $r = s$ and zero otherwise (Hint: make use of the identity

$$\sum_{n=0}^{N-1} \cos(\pi J n/N) = \begin{cases} N & \text{if } J = 2sN \ , \ s \in \mathbb{Z} \\ 1 & \text{if } J = (2s+1) \ , \ s \in \mathbb{Z} \end{cases}$$

which in turn can be proven by means of the geometric sum).

7. Show that the solution of the problem with initial conditions $y_j(0) = \xi_j$, $\dot{y}_j(0) = \eta_j$, $j = 1, \ldots, N-1$ is given by

$$y_j(t) = \sum_{s=1}^{N-1} \left[\langle \varphi_s, \xi \rangle \cos(\omega_s t) + \frac{\langle \varphi_s, \eta \rangle}{\omega_s} \sin(\omega_s t) \right] \varphi_s(j) \ . \tag{1.43}$$

The method of solution just exposed is known as *reduction to normal modes* of the given Lagrangian problem. The sth normal mode of the beaded string is the solution $y_j^{(s)}(t) = q_s(t)\varphi_s(j)$, characterized by a space profile $\varphi_s(j)$ oscillating at the frequency ω_s. The general solution of the problem is a linear combination of all the normal modes of the system.

1.5.2 Continuum limit

We are interested in studying the continuum limit of the beaded string with a fixed length L, as $N \to \infty$ and, as a consequence, $a = L/N \to 0$. In order to do this, we start by supposing that there exists a smooth interpolating function $u(t, x)$ defined on the interval $[0, L]$, vanishing at its ends and such that

$$y_j(t) = u(t, ja) \ , \quad j = 0, \ldots, N \ . \tag{1.44}$$

The following three hypotheses are made:

1. the linear mass density $\rho := \lim_{a\to 0} Nm/L = \lim_{a\to 0} m/a$ exists;

2. the tension $\tau := \lim_{a\to 0} \gamma a$ exists;

3. the local square frequency $\omega^2 := (\lim_{a\to 0} k/a)/\rho$ exists.

Thus the mass of the particles and the elastic constants of the springs connecting them must depend in a precise way on the step $a = L/N$.

1. Show that as $a \to 0$, the limit of the dispersion relation (1.40) exists and is given by

$$\lim_{a\to 0} \omega_s^2 = \omega^2 + \left(\frac{\pi c s}{L}\right)^2 \ , \quad s = 1, 2 \ldots \ . \tag{1.45}$$

2. Show that if y_j, as given by (1.44), satisfies the L-equations (1.36), then u satisfies the one-dimensional Klein-Gordon (1DKG) equation

$$u_{tt} = c^2 u_{xx} - \omega^2 u \ , \quad c^2 := \tau/\rho \ , \tag{1.46}$$

whereas, in the same limit (i.e. $a \to 0$ or $N \to \infty$)

$$\mathscr{L}_N \to \mathscr{L}_\infty[u, u_t] := \int_0^L \left[\rho \frac{(u_t)^2}{2} - \rho\omega^2 \frac{u^2}{2} - \tau \frac{(u_x)^2}{2} \right] \mathrm{d}x \ . \tag{1.47}$$

As a consequence, the action functional

$$A_N := \int_{t_1}^{t_2} \mathscr{L}_N \mathrm{d}t \to \int_{t_1} t_2 \mathscr{L}_\infty =: A_\infty[u] \ . \tag{1.48}$$

3. Show that the differential of any Lagrangian functional of the form

$$\mathscr{L}[u, u_t] = \int_0^L \mathcal{L}(u, u_t, u_x) \ \mathrm{d}x \ , \tag{1.49}$$

and defined on functions that vanish at $x = 0$ and $x = L$, is given by

$$\delta\mathscr{L}[u, u_t] = \int_0^L \left[\left(\frac{\partial \mathcal{L}}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial \mathcal{L}}{\partial u_x} \right) + \frac{\partial \mathcal{L}}{\partial u_t} \right] \ \mathrm{d}x \ . \tag{1.50}$$

Show that, as a consequence, the Hamilton variational principle

$$\delta A = 0 \ \forall \ \delta u \ \Leftrightarrow \ \frac{\delta A}{\delta u} = \frac{\partial \mathcal{L}}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial \mathcal{L}}{\partial u_x} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial u_t} = 0 \tag{1.51}$$

holds for the action functional $A = \int \mathscr{L}[u, u_t] \mathrm{d}t$.

4. Write down explicitly the Lagrange equation on the right hand side of (1.51) for the Lagrangian $\mathscr{L}_\infty$ defined in (1.47), and show that it coincides with the 1D KG equation (1.46).

5. Suppose that the initial conditions ξ, η of the beaded string be given by two C^∞ functions u_0 and v_0 that vanish at $x = 0$ and $x = L$, as follows:

$$\xi_j = u_0(ja) \ \ ; \ \ \eta_j = v_0(ja) \ , \quad j = 0, \ldots, N \ .$$

Write the corresponding solution (1.43) and find the interpolating function $u(t, x)$ supposed to exist at the beginning of our treatment of the continuum limit. Find the limit of such a solution and check that it solves the 1D KG equation (Hint: study first the limit of $\langle \varphi_s, \varphi_r \rangle = \delta_{sr}$).

6. Solve directly the 1D KG equation (1.46) by separation of variables and linear superposition of the normal mode solutions.

Chapter 2

Hamiltonian formalism

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

$\triangleleft$ 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 $\dot{p} = \partial L/\partial q$. $\triangleright$

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 (1.12) 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 (1.17) 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 (1.21).

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 (1.17) 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 The action as a function

The Hamilton principle means that the value of the action (2.9) does not change close to a critical curve $t \mapsto (q, p)(t)$ solution of the Hamilton equations with boundary conditions $q(t_1) = a$ and $q(t_2) = b$. More precisely, if a curve is ϵ-close to the critical one, then the variation of A_H is $O(\epsilon^2)$. Of course the specific value of the action at the critical point depends on the boundary conditions. One can thus consider the action (2.9) as a function $S(q, t)$ of the arrival time $t_2 := t$ and of the arrival coordinate $q(t) := q$ along a critical path, for example. One thus defines

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

the integral on the right being computed on a critical curve $t \mapsto (q,p)(t)$ solution of the Hamilton equations with boundary conditions $q(t_1) = a$, thought of as fixed, and $q(t) = q$. Taking the differential (or the total time-derivative) 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 $p(t) := p$ 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

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

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 Hamiltonian systems

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 $\dot q = \partial H/\partial p$, $\dot p = -\partial H/\partial q$

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial q}\cdot\dot q + \frac{\partial H}{\partial p}\cdot\dot p + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}\ , \tag{2.14}$$

   so that H is a first integral, or conserved quantity, 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 (i.e. explicitly time-dependent) 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 $\dot\xi = \partial K/\partial\eta = 1$).

4. By introducing the quantities

$$x := \begin{pmatrix} q \\ p \end{pmatrix} \quad ; \quad \nabla_x H(x,t) := \begin{pmatrix} \partial H/\partial q \\ \partial H/\partial p \end{pmatrix} ,$$

and

$$J_{2n} := \begin{pmatrix} \mathbb{O}_n & \mathbb{I}_n \\ -\mathbb{I}_n & \mathbb{O}_n \end{pmatrix} , \tag{2.15}$$

the Hamilton equations take on the simpler and more compact form

$$\boxed{\dot{x} = J_{2n} \nabla_x H(x,t)} . \tag{2.16}$$

The matrix (2.15) is called *standard symplectic matrix*, and its fundamental role, properties and meaning are going to be discussed below. The lighter notation $J\nabla H$ will be also used in the sequel, when no confusion is possible.

5. The x-space, or (q,p)-space of a Hamiltonian system is referred to as *phase space* and usually denoted by $\Gamma \subseteq \mathbb{R}^{2n}$. Hamiltonian systems are those dynamical systems defined by a vector ordinary differential equation (ODE) of first order $\dot{x} = u(x,t)$ whose vector fiels u is the *symplectic gradient* of a given function $H : \Gamma \times \mathbb{R} \to \mathbb{R}$, namely a vector field of the form

$$u(x,t) = J\nabla H(x,t) := X_H(x,t) . \tag{2.17}$$

The algebra[1] of real smooth functions F defined on $\Gamma \times \mathbb{R}$ is the so-called *algebra of observables* (i.e. the space of all possible Hamiltonians). To each function F in the algebra is associated its vector field $X_F = J\nabla F$, i.e. its symplectic gradient.

6. The solution of the Hamilton equations $\dot{x} = J\nabla H$ at time t with initial condition ξ at time $t = t_0$ is called the flow of the given Hamiltonian system, or the H-flow in short, and denoted by $\Phi_H^{t,t_0}(\xi)$. The H-flow has the following properties:

$$\Phi_H^{t_0,t_0} = \mathrm{id}_\Gamma \quad ; \quad \Phi^{t_2,t_0} = \Phi_H^{t_2,t_1} \circ \Phi_H^{t_1,t_0}$$

where id_Γ denotes the identity on Γ and $t_0 \le t_1 \le t_2$ are arbitrary times. Of course, in the autonomous case ($\partial H/\partial t = 0$) $\Phi_H^{t,t_0} = \Phi_H^{t-t_0}$. Such properties are general and hold for flows of any ODE. Notice that the H-flow is a one parameter group, with the property $\Phi_H^t \circ \Phi_H^s = \Phi_H^{t+s}$ only in the autonomous case and only if the solution is global for any initial condition, i.e. if $\Phi_H^t(\xi)$ exists for any $t \in \mathbb{R}$ and $\xi \in \Gamma$ (in such a case the flow is also said to be complete). In the autonomous case the Hamiltonian H is a constant of motion and the dynamics takes place on the surface $\Sigma_E := \{x \in \Gamma : H(x) = E\} := H^{-1}(E)$. A sufficient condition for global existence of the solution is the compactness of Σ_E.

Exercise 2.3. *Check that in the autonomous case $X_H = J\nabla H$ is tangent to Σ_E (Hint: ∇H is orthogonal to Σ_E).*

In the sequel no distinction is made between global and local H-flows, unless strictly necessary.

[1]An algebra is a vector, or linear space closed with respect to a product that is distributive with respect to the sum.

2.3.1 Symplectic structure

Let us consider in some detail the algebraic and geometric structure of H-systems $\dot{x} = X_H(x,t) = J\nabla H(x,t)$, where $J = J_{2n}$ is the standard symplectic matrix defined in (2.15). One immediately checks that $J^T = J^{-1} = -J$, and $J^2 = -\mathbb{I}_{2n}$, the latter implying $|\det J| = 1$; actually $\det J = 1$, as can be easily proven. Indeed, consider the eigenvalue problem $J\xi = \lambda\xi$; $\lambda \neq 0$ since $\det J \neq 0$. Setting $\xi = (u,v)^T$ 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 is real, the eigenvalues come in complex conjugate pairs, and the product of them is one.

The matrix J is called standard symplectic matrix for the following reason. Consider the skew-symmetric bilinear form on $\mathbb{R}^{2n}$

$$\omega(x,y) := x \cdot Jy \ , \tag{2.18}$$

called the *symplectic product* (which is not a scalar product since $\omega(x,x) = 0$ for any x). Now look for the linear transformations $x \mapsto Sx$, $y \mapsto Sy$ that leave such a form invariant, namely $\omega(Sx,Sy) = \omega(x,y)$ for any x,y. 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. Indeed 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 $S^T JS = J$ 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 $S^T \in Sp(2n,\mathbb{R})$, i.e. also $SJS^T = J$ holds. Obviously, $J \in Sp(2n,\mathbb{R})$. From the relation $SJS^T = J$ it follows that $|\det S| = 1$. Actually $\det S = \det J = +1$. Indeed,

$$\begin{aligned} p_S(\lambda) &:= \det(S - \lambda\mathbb{I}) = \det(-JS^{-T}J - \lambda\mathbb{I}) = \det(-J(S^{-T} - \lambda\mathbb{I})J) = \\ &= \det(S^{-T}(\mathbb{I} - \lambda S^T)) = \frac{\lambda^{2n}}{\det S}\det(S^T - \lambda^{-1}\mathbb{I}) = \\ &= \frac{\lambda^{2n}}{\det S}\, p_S(1/\lambda) \ . \end{aligned} \tag{2.19}$$

From such an identity, and from the fact that p_S is a polynomial with real coefficients, it follows that the eigenvalues of S occur in quartets: $\lambda, 1/\lambda, \bar{\lambda}, 1/\bar{\lambda}$ (an overbar denoting complex conjugation). Moreover, one easily checks that the algebraic multiplicity of the four members of a quartet is the same. Then, the product of the eigenvalues is one.

The dimension $\dim Sp(2n,\mathbb{R})$ of the symplectic group is defined as the number of independent parameters that uniquely determine any of its elements. Indeed, due to the fundamental relation $SJS^T = J$, the $4n^2$ real elements of the matrix S are not independent of each other. It turns out that $\dim Sp(2n,\mathbb{R}) = n(2n+1)$. In order to show this, let us write the generic symplectic matrix S in block form, namely

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ ,$$

where a, b, c, d are $n \times n$ blocks. Then, the fundamental relation $SJS^T = J$ is easily shown to be equivalent to the following three conditions

$$\begin{cases} ab^T = (ab^T)^T \\ cd^T = (cd^T)^T \\ ad^T - bc^T = \mathbb{I}_n \end{cases} .$$

The first two relations require a matrix to be symmetric, so that they give rise $2n(n-1)/2 = n^2 - n$ equations; the last one gives rise to n^2 equations. The number of independent entries of S is given by $4n^2 - (n^2 - n) - n^2 = 2n^2 + n = n(2n+1)$.

As will be shown below, it is interesting to characterize the tangent space of the group $Sp(2n, \mathbb{R})$ at the identity $\mathbb{I}_{2n}$, namely $T_{\mathbb{I}}Sp(2n, \mathbb{R}) := \mathfrak{sp}(2n, \mathbb{R})$. In order to do this, let us consider a curve $t \mapsto S(t)$ in the symplectic group passing through the identity, namely $S(t) \in Sp(2n, \mathbb{R})$ as t varies in a real interval containing $t = 0$, and $S(0) = \mathbb{I}_{2n}$. Call $M := \dot{S}(0)$ the tangent vector to such a curve in the symplectic group. Taking the time derivative of $S(t)JS^T(t) = J$ with respect to t and setting $t = 0$ one gets

$$MJ + JM^T = 0 . \tag{2.20}$$

The relation above defines the vector space $\mathfrak{sp} = T_{\mathbb{I}}Sp$ (that this is a vector space follows from the fact that (2.20) is linear and homogeneous in M). One easily proves that $M \in \mathfrak{sp}$ implies $M^T \in \mathfrak{sp}$. Moreover, one checks that if $M, M' \in \mathfrak{sp}$ then the commutator

$$[M, M'] := MM' - M'M \in \mathfrak{sp} . \tag{2.21}$$

Such a properties implies that the vector space $\mathfrak{sp}$ with the product $[\ ,]: \mathfrak{sp} \times \mathfrak{sp} \to \mathfrak{sp}$ is a Lie algebra[2]. Matrices satisfying (2.20) are also called Hamiltonian matrices. An important characterization of $\mathfrak{sp}$ is the following: $M \in \mathfrak{sp}$ iff $M = JB$ with $B = B^T$. The proof is immediate: observe that J is nonsingular and plug $M = JB$ into the left hand side of (2.20), thus getting $MJ + JM^T = J(B - B^T)J$. One can observe that the dimension of the tangent vector space $\mathfrak{sp}(2n, \mathbb{R})$ coincides with the dimension of the symplectic group (Hint: compute the number of independent parameters needed to specify a symmetric $2n \times 2n$ matrix). As a final remark, it is observed that

$$M \in \mathfrak{sp} \ \Leftrightarrow \ S(t) = e^{tM} \in Sp . \tag{2.22}$$

This is easily proven by setting $W(t) := e^{tM} J e^{tM^T}$; by differentiating with respect to t one gets the equation $\dot{W} = MW + WM^T$, and notice that $W(0) = J$ is a solution of such equation. The conclusion $W(t) = W(0) = J$ is obtained by uniqueness. $\{e^{tM}\}_{t \in \mathbb{R}}$ is a one parameter subgroup of Sp.

[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 matrix vector space $\mathfrak{sp}$ and $[\ ,]$ is the ordinary commutator which is easily checked to satisfy the Jacobi identity, namely $[[M, M'], M''] + [[M', M''], M] + [[M'', M], M'] \equiv 0$ for any triple M, M', M''.

What has been just shown is an example of a general result. Actually the symplectic group $Sp(2n, \mathbb{R})$ is a Lie group of dimension $N = n(2n+1)$. A Lie group G of dimension N is a group with a structure of N-dimensional differentiable manifold compatible with the group structure. In practice this means that there exists a smooth map $\mathbb{R}^N \ni \mu \mapsto g(\mu) \in G$ (the compatibility of the manifold structure with the group one is already understood by requiring the closure: $g(\mu) \circ g(\nu) = g(\omega)$ means that there exists a function $\varphi : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ such that $\omega = \varphi(\mu, \nu)$. Other properties of φ are easily inferred by the group properties). A general theorem states that the tangent pace $\mathfrak{g} := T_e G$ to G at its identity e has a natural structure of Lie algebra (and is referred to as *the* Lie algebra of the given Lie group). For matrix groups, the skew symmetric bilinear product characterizing the Lie algebra $\mathfrak{g}$ is always the commutator. Moreover, in general, $m \in \mathfrak{g}$ iff $e^{tm} \in G$ for any t. The following are the most important examples of matrix Lie groups and their Lie algebras.

- $O(n)$, the orthogonal group, which is the invariance group of the real, Euclidean, scalar product $x \cdot y$ in $\mathbb{R}^n$. The fundamental relation is $R^T R = \mathbb{I}$. The dimension of the group is $n(n-1)/2$. The component connected to the identity of the group, whose matrices have determinant $\det R = +1$, is the special orthogonal group $SO(n)$, or the group of proper rotations in $\mathbb{R}^n$, whose Lie algebra $\mathfrak{so}(n)$ is the algebra of skew symmetric matrices $A^T + A = 0$. One parameters subgroups of $SO(n)$ have the form $R(t) = e^{tA}$.

- $U(n)$, the unitary group, which is the invariance group of the complex, Euclidean, scalar product $x \cdot y$ in $\mathbb{C}^n$. The fundamental relation is $U^\dagger U = \mathbb{I}$, where $U^\dagger = \bar{U}^T$ denotes complex conjugation and transposition. The dimension of the group is n^2. The component connected to the identity of the group, whose matrices have determinant $\det U = +1$, is the special unitary group $SU(n)$, or the group of proper rotations in $\mathbb{C}^n$, whose Lie algebra $\mathfrak{su}(n)$ is the algebra of anti Hermitian matrices $A^\dagger + A = 0$. Notice that the dimension of $SU(n)$ and its algebra is $n^2 - 1$: the determinant of the unitary group has unit modulus in $\mathbb{C}$. A anti Hermitian matrix is always of the form $A = \imath H$, where $H^\dagger = H$ is Hermitian and $\imath$ is the imaginary unit. One parameters subgroups of $SU(n)$ have the form $R(t) = e^{\imath t H}$.

The symplectic group naturally arises when one consider a H-system $\dot{x} = J\nabla_x H$ and performs a linear change of coordinates $y(t) = Sx(t)$. The gradient transforms according to

$$\frac{\partial}{\partial x_j} = \sum_k \frac{\partial y_k}{\partial x_j} \frac{\partial}{\partial y_k} \ \Leftrightarrow\ \nabla_x = \left(\frac{\partial y}{\partial x}\right)^T \nabla_y = S^T \nabla_y \ ,$$

which implies

$$\dot{y} = SJS^T \ \nabla_y \tilde{H} \ , \quad \tilde{H}(y) := H(S^{-1} y) \ .$$

The transformed system (in the y-variables) is Hamiltonian if $\partial y/\partial x = S$ is symplectic: $SJS^T = J$. A linear transofrmation of coordinates with symplectic Jacobian is an example of *canonical transformation* (whose definition, properties and characterizations will be given in the sequel). The symplectic algebra naturally arises in considering H-systems linearized around a non degenerate critical point x_0 of H: $\nabla H(x_0) = 0$, $\partial^2 H(x_0)$ nonsingular ($\partial^2 H = \partial^2 H/\partial x^2$ denotes the Hessian matrix of H). Setting $x = x_0 + \xi$ and linearizing one gets

$$\dot{\xi} = M\xi := J\partial^2 H(x_0)\xi \ . \tag{2.23}$$

Then $M = J\partial^2 H(x_0)$, the Jacobian of the linear Hamiltonian vector field, is a Hamiltonian matrix, i.e. an element of the symplectic algebra. The solution of equation (2.23) is $\xi(t) = e^{tM}\xi_0$, and the Jacobian e^{tM} of such a Hamiltonian flow with respect to the initial condition is symplectic.

2.3.2 Poisson bracket

Let us 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 a H-flow, namely

$$\begin{aligned} \dot{F} &:= \frac{\mathrm{d}}{\mathrm{d}t} F(q(t), p(t), t) = \frac{\partial F}{\partial q} \cdot \dot{q} + \frac{\partial F}{\partial p} \cdot \dot{p} + \frac{\partial F}{\partial t} = \\ &= \frac{\partial F}{\partial q} \cdot \frac{\partial H}{\partial p} - \frac{\partial H}{\partial q} \cdot \frac{\partial F}{\partial p} + \frac{\partial F}{\partial t} := \{F, H\} + \frac{\partial F}{\partial t} \ . \end{aligned} \tag{2.24}$$

Such a formula holds even if the flow is not globally defined. In the last step of (2.24) the *Poisson (P) bracket* $\{F, G\}$ of two functions F and G defined on Γ (possibly depending on time explicitly) has been defined:

$$\{F, G\} = \frac{\partial F}{\partial q} \cdot \frac{\partial G}{\partial p} - \frac{\partial G}{\partial q} \cdot \frac{\partial F}{\partial p} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i} \frac{\partial F}{\partial p_i} \right) \ . \tag{2.25}$$

One easily checks that in the notation introduced above the P-bracket reads

$$\{F, G\} = \omega(\nabla F, \nabla G) = \omega(X_F, X_G) \ , \tag{2.26}$$

where $\omega(\cdot, \cdot)$ is the symplectic product defined in (2.18) (check it explicitly). It is also observed that the P-bracket is invariant with respect to linear symplectic transformations $y = Sx$, with $S \in Sp$. Indeed

$$\{F, G\}_x = \omega(\nabla_x F, \nabla_x G) = \omega(S^T \nabla_y \tilde{F}, S^T \nabla_y \tilde{G}) = \omega(\nabla_y \tilde{F}, \nabla_y \tilde{G}) = \{\tilde{F}, \tilde{G}\}_y \ ,$$

where $\tilde{F}(y) = F(S^{-1}y)$ and $\tilde{G}(y) = G(S^{-1}y)$.

The P-bracket is a function defined on $\Gamma \times \mathbb{R}$, 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),

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

Remark 2.3. *The algebra of observables, i.e. functions defined on $\Gamma \times \mathbb{R}$, endowed with the operation $\{\ ,\}$ has the structure of a Lie algebra (properties 1., 2., and 3.) of Leibniz type (property 4.). A Lie-Leibniz algebra is called a Poisson (P) algebra.*

In terms of the P-bracket, one can rewrite the H-equations as follows:

$$\dot{x} = \{x, H\}\ , \tag{2.27}$$

to be meant by components, which is an alternative (and equivalent) definition of Hamiltonian system.

2.3.3 Hamiltonian flows and vector fields

It has been shown above that the Jacobian of a linearized Hamiltonian vector field is a Hamiltonian matrix (an element of the symplectic algebra), whereas the Jacobian of the H-flow of a linearized system is a symplectic matrix. Actually, such two properties characterize Hamiltonian vector fields and flows, in general.

Proposition 2.3. *Consider an ODE*

$$\dot{x} = u(x, t) \tag{2.28}$$

in $\mathbb{R}^{2n}$, with solution $x(t) = \Phi^{t,t_0}(\xi)$ corresponding to the initial condition $x(t_0) = \xi$.

1. *The vector field of (2.28) is Hamiltonian iff its Jacobian $\partial u/\partial x$ is Hamiltonian.*

2. *The flow of (2.28) is Hamiltonian iff its Jacobian $\partial x/\partial \xi$ is symplectic.*

$\triangleleft$ PROOF. 1. If u is Hamiltonian then $u = J\nabla_x H$, so that $\partial u/\partial x = J\partial_x^2 H$ which is of the Hamiltonian form. On the other hand, if $\partial u/\partial x \in \mathfrak{sp}$ then it satisfies $(\partial u/\partial x)J + J(\partial u/\partial x)^T = 0$. Set $u := Jv$, which defines v. Then $\partial u/\partial x = J(\partial v/\partial x)$ and

$$0 = \frac{\partial u}{\partial x} J + J \left(\frac{\partial u}{\partial x}\right)^T = J \left[\frac{\partial v}{\partial x} - \left(\frac{\partial v}{\partial x}\right)^T\right] J := J\left(\mathrm{rot}(v)\right) J\ .$$

Thus $\mathrm{rot}(v) = 0$, which implies $v = \nabla H$ and $u = J\nabla H$.

2. Define $\mathcal{J}(\xi, t) := \partial x/\partial \xi$. One easily checks that $\dot{\mathcal{J}} = (\partial u/\partial x)\mathcal{J}$. Define also $W(\xi, t) := \mathcal{J} J \mathcal{J}^T$. By taking the time derivative of W one finds

$$\dot{W} = \frac{\partial u}{\partial x}\, W + W \left(\frac{\partial u}{\partial x}\right)^T\ ;\ W(t_0) = J\ . \tag{2.29}$$

Thus, if $\mathcal{J}$ is symplectic then $W(t) = J$ for any t and $\dot{W} = 0$, so that (2.29) implies that $\partial u/\partial x$ is Hamiltonian, $u = J\nabla H$ and the flow of (2.28) is a Hamiltonian flow. On the other hand, if the flow of (2.28) is Hamiltonian, then $u = J\nabla H$ and $\partial u/\partial x$ is a Hamiltonian matrix. Then $W(t) = W(t_0) = J$ is a solution of the linear system of ODEs (2.29). By uniqueness this is the solution of (2.29), which implies $\mathcal{J}$ symplectic. $\triangleright$

2.3.4 Liouville equation and statistical mechanics

As an application of the above properties, let us consider again the ODE (2.28) in $\Gamma \subseteq \mathbb{R}^N$, where N is not necessarily even. Let us imagine to choose, on the phase space Γ, a probability measure $\rho(x,t)\mathrm{d}V(x)$, where $\rho(x,t)$ is the probability density at (x,t) and $\mathrm{d}V(x) = \mathrm{d}^N x$ is the volume element. Such a measure defines the probability to find the system in a set $\Omega_0 \subseteq \Gamma$ at a given time t_0, say, namely

$$\mathrm{Prob}(\xi \in \Omega_0 | t_0) := \int_{\Omega_0} \rho(\xi, t_0)\mathrm{d}V(\xi) \ . \tag{2.30}$$

Of course $\mathrm{Prob}(x \in \Gamma | t) = \int_\Gamma \rho \mathrm{d}V = 1$. Let Ω_t be the evolution at time $t > t_0$ of the set Ω_0 along the flow of (2.28), namely $\Omega_t := \{x(t) \in \Gamma : \dot{x} = u;\ x(t_0) = \xi \in \Omega_0\}$. The following hypothesis allows to get a partial differential equation satisfied by ρ:

$$\mathrm{Prob}(x(t) \in \Omega_t | t) = \mathrm{Prob}(\xi \in \Omega_0 | t_0) \ , \tag{2.31}$$

which can be interpreted as a law of mass conservation in absence of sources or sinks. By taking the time derivative of (2.31) and using the definition (2.30), one gets

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t} \rho(x(t),t)\mathrm{d}V(x(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_0} \rho(x(t),t) \det\left(\frac{\partial x(t)}{\partial \xi}\right) \mathrm{d}V(\xi) = \\ &= \int_{\Omega_0} \left[\frac{\partial \rho}{\partial t} + u \cdot \frac{\partial \rho}{\partial x} + \frac{\dot{D}}{D}\rho\right] D(\xi,t)\mathrm{d}V(\xi) \ , \end{aligned} \tag{2.32}$$

where $D(\xi,t) := \det\left(\frac{\partial x(t)}{\partial \xi}\right)$. By a direct computation one easily finds

$$\frac{\dot{D}}{D} = \mathrm{tr}\left(\frac{\partial u}{\partial x}\right) = (\nabla_x \cdot u) \ . \tag{2.33}$$

Exercise 2.4. *Prove formula (2.33). Hint: apply the definition of derivative.*

Equations (2.32) and (2.33) together imply that ρ has to satisfy the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla_x \cdot (\rho u) = 0 \ . \tag{2.34}$$

Such an equation describes the time evolution of the probability density once the initial condition $\rho(\xi, t_0)$ is specified.

Now let us restrict to consider Hamiltonian systems ($N = 2n$). In such a case $u = X_H := J\nabla_x H(x,t)$ and

$$\nabla_x \cdot X_H \equiv 0 = \sum_{i=1}^n \left(\frac{\partial}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i}\frac{\partial H}{\partial q_i}\right) \equiv 0 \ ,$$

which in turn implies, by (2.33), $D(\xi,t) = D(\xi,t_0) = 1$. Such a statement is usually referred to as the Liouville theorem on conservation of the phase space volume along Hamiltonian flows:

$$\mathrm{d}V(x(t)) = D(\xi,t)\mathrm{d}V(\xi) = \mathrm{d}V(\xi) \ .$$

An alternative way to get such a conclusion, which does not make use of the general equation (2.33), consists in using the fact that $\partial x(t)/\partial\xi$ is symplectic and its determinant is one. As a further remark, we notice that in the Hamiltonian case

$$\nabla_x \cdot (\rho X_H) = \nabla_x \rho \cdot X_H = \nabla_x \rho \cdot J \nabla_x H = \{\rho, H\} \ ,$$

so that the continuity equation simplifies to

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0 \ . \tag{2.35}$$

The latter equation is known as the *Liouville equation*, describing the time evolution of the probability density in the phase space of a Hamiltonian system.

Statistical mechanics is the theory that reconstructs equilibrium thermodynamics of Hamiltonian systems starting from a the microscopic properties. In a nutshell, it consists in choosing particular stationary measures, i.e. probability densities independent of times, such that one is able to consistently formulate the first and second principle of thermodynamics for reversible transformations. The condition $\partial\rho/\partial t = 0$ in the Liouville equation implies $\{\rho, H\} = 0$ (where $\partial H/\partial t = 0$ is also understood). If H does not display non trivial first integrals, one is led to the condition $\rho = f(H)$. The (essentially unique) right choice to build up thermodynamics is the so-called Gibbs density, namely $f(H) = e^{-\beta H}/Z$, where $\beta = 1/T$ is the inverse temperature and $Z = \int_\Gamma e^{-\beta H} \mathrm{d}V$ is a normalization constant. The fastest way to get the Gibbs distribution is to observe that if one artificially puts together two noninteracting systems, whose Hamiltonian is the sum of the two, namely $H = H_1 + H_2$, the they must be statistically independent of each other:

$$f(H_1 + H_2) = f(H_1) f(H_2) \ ,$$

which immediately leads to the exponential function.

2.4 Canonical transformations

A natural 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) or (2.16)? It has been shown above that linear transformations of coordinates defined by a symplectic Jacobian (constant) matrix map any Hamiltonian system into a Hamiltonian system, preserving the Hamiltonian and the Poisson bracket. In order to characterize the transformations of coordinates that map any Hamiltonian system into a Hamiltonian system, it is first convenient to consider how a general system transforms under a general change of variables.

Let us consider the system

$$\dot{x} = u(x, t) \ ; \quad x \in \mathbb{R}^N \ , \tag{2.36}$$

where N is any dimension (not necessarily even) and $u(x, t)$ is a given vector field. Consider then the time-dependent change of variables

$$x \mapsto y = f(x, t) \ \ ; \ \ y \mapsto x = g(y, t) \ . \tag{2.37}$$

First of all, one easily checks that, by differentiation, the identities $x = g(f(x,t),t)$ and $y = f(g(y,t),t)$ imply

$$\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} = \mathbb{I}_N = \frac{\partial g}{\partial y}\frac{\partial f}{\partial x} \quad ; \quad \frac{\partial f}{\partial x}\frac{\partial g}{\partial t} + \frac{\partial f}{\partial t} = 0 = \frac{\partial g}{\partial y}\frac{\partial f}{\partial t} + \frac{\partial g}{\partial t} \ . \tag{2.38}$$

Then, taking the time derivative of one of the two relations $y = f(x,t)$ or $x = g(y,t)$, and taking into account (2.36), one finds that the new variable $y(t)$ evolves according to

$$\dot{y} = v(y,t) \ , \tag{2.39}$$

where the transformed vector field v is given by

$$v(y,t) := \left[\frac{\partial f}{\partial x}u(x,t) + \frac{\partial f}{\partial t}\right]\Big|_{x=g(y,t)} = \left(\frac{\partial g}{\partial y}\right)^{-1}\left[u(g(y,t),t) - \frac{\partial g}{\partial t}\right] \ . \tag{2.40}$$

The vector field v (u, respectively) is said to be the conjugate of u (v) by the transformation f (g). This is completely general. Now, restricting to the Hamiltonian case ($N = 2n$ now), and taking into account that

$$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial y_j}{\partial x_i}\frac{\partial}{\partial y_j} \ \Leftrightarrow \ \nabla_x = \left(\frac{\partial f}{\partial x}\right)^T \nabla_y \ ,$$

one finds that the vector field v conjugate to the Hamiltonian vector field $u = X_H := J\nabla_x H(x,t)$ by f is

$$v(y,t) = \frac{\partial f}{\partial x} J \left(\frac{\partial f}{\partial x}\right)^T \Big|_{x=g(y,t)} \nabla_y \tilde{H}(y,t) + \frac{\partial f}{\partial t}\Big|_{x=g(y,t)} = \tag{2.41}$$

$$= \left(\frac{\partial g}{\partial y}\right)^{-1} J \left(\frac{\partial g}{\partial y}\right)^{-T} \nabla_y \tilde{H}(y,t) - \left(\frac{\partial g}{\partial y}\right)^{-1}\frac{\partial g}{\partial t} \ , \tag{2.42}$$

where the transformed Hamiltonian

$$\tilde{H}(y,t) := H(g(y,t),t) \tag{2.43}$$

has been defined. The following definition is now naturally suggested by the above formulas.

Definition 2.2. *The change of variables, or transformation f (defined by the formulas (2.37)) is said to be canonical if it conjugates any Hamiltonian vector field $u = X_H$ to a Hamiltonian vector field $v = X_K$.*

An equivalent definition formulated in terms of coordinates and momenta is the following

Definition 2.3. *A change of variables, or transformation*

$$(q,p) \mapsto (Q,P) = (V(q,p,t), U(Q,P,t)) \quad ; \quad (Q,P) \mapsto (q,p) = (v(Q,P,t), u(Q,P,t)) \tag{2.44}$$

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

Observe how extracting the information written down above in terms of the latter notation would be rather cumbersome (though perfectly possible). According to the definition 2.2 of canonical transformation, one is led to find the necessary and sufficient conditions under which the vector field v defined in (2.41)-(2.42) turns out to be of the form $J\nabla_y K(y,t)$, specifying K. More precisely, we want to characterize the transformations f such that, for any Hamiltonian $H(x,t)$ there exists a Hamiltonian $K(y,t)$ satisfying

$$J\nabla_y K(f(x,t),t)\Big|_{x=g(y,t)} = \frac{\partial f}{\partial x} J \left(\frac{\partial f}{\partial x}\right)^T \Big|_{x=g(y,t)} \nabla_y \tilde{H}(y,t) + \frac{\partial f}{\partial t}\Big|_{x=g(y,t)} , \tag{2.45}$$

or, equivalently

$$J\nabla_y K(y,t) = \left(\frac{\partial g}{\partial y}\right)^{-1} J \left(\frac{\partial g}{\partial y}\right)^{-T} \nabla_y \tilde{H}(y,t) - \left(\frac{\partial g}{\partial y}\right)^{-1} \frac{\partial g}{\partial t} , \tag{2.46}$$

with $\tilde{H}$ as defined in (2.43). The following theorem characterizes completely the canonical transformations.

Theorem 2.1 (Characterization of canonical transformations)**.** *The following statements are equivalent.*

1. *f is canonical;*
2. *$\mathcal{J} = \partial f/\partial x$ is symplectic; then $K = \tilde{H} + K_0$, where either f depends explicitly on time and is the flow associated to K_0, or K_0 is a constant.*
3. *f preserves the P-bracket of any two functions, namely $\{F,G\}_x = \{\tilde{F},\tilde{G}\}_y$, where $\tilde{F}(y,t) = F(g(y,t),t)$ and $\tilde{G}(y,t) = F(g(y,t),t)$;*
4. *$\{f_i,f_j\}_x = \{y_i,y_j\}_y = J_{ij}$ for any $i,j = 1,\dots,2n$;*
5. *for any $H(x,t)$ there exists a Hamiltonian $K(y,t)$ such that the difference of the differential 1-forms $\pi_H(x,t) := \frac{1}{2}(Jx)\cdot \mathrm{d}x - H(x,t)\mathrm{d}t$ and $\pi_K(y,t)|_{y=f(x,t)}$ is exact.*

$\triangleleft$ PROOF. 1. $\Rightarrow$ 2. If f is canonical then (2.45) holds for any H; in particular, it holds for $H = 0$. If f depends explicitly on time, then this implies

$$\frac{\partial f}{\partial t} = J\nabla_y K_0 , \tag{2.47}$$

where K_0 denotes the Hamiltonian corresponding to $H = 0$. Then $f(x,t)$ is the flow of the Hamiltonian K_0 and its Jacobian $\mathcal{J}$, due to Proposition 2.3, is symplectic. By lifting a gradient in (2.45) one gets $K = \tilde{H} + K_0$. If instead f does not explicitly depend on time, then, upon multiplying (2.45) by $-J$, one gets $M\nabla_y\tilde{H} = \nabla_y K$, where $M := -J\mathcal{J}J\mathcal{J}^T$. By properly choosing the form of the Hamiltonian $\tilde{H}$ one easily proves that $M = c\mathbb{I}$, where $c \neq 0$ is a real constant (i.e. independent of y; it cannot depend of t because f does not by hypothesis). This implies $\mathcal{J}J\mathcal{J}^T = cJ$. Now, one excludes negative values of c by requiring that the class of transformations is connected to the identity (where $\mathcal{J} = \mathbb{I}$ and $c = 1$), after which one can set $c = 1$ by the rescaling $f \to f/\sqrt{c}$. Then $\mathcal{J}$ is symplectic and $K = \tilde{H}$ (up to a constant K_0).

2. $\Rightarrow$ 1. This is trivial by (2.45).

2. $\Rightarrow$ 3. The Poisson bracket transforms through f as follows:

$$\{F, G\}_x = \nabla_x F \cdot J \nabla_x G = \nabla_y \tilde{F} \cdot (\mathcal{J} J \mathcal{J}^T) \nabla_y \tilde{G} \ . \tag{2.48}$$

The latter expression coincides with $\{\tilde{F}, \tilde{G}\}_y$ if $\mathcal{J}$ is symplectic (no matter whether f depends or not explicitly on time).

3. $\Rightarrow$ 2. Viceversa, if $\{F, G\}_x = \{\tilde{F}, \tilde{G}\}_y$ for any pair of functions F and G, then (2.48) implies

$$\nabla_y \tilde{F} \cdot (\mathcal{J} J \mathcal{J}^T - J) \nabla_y \tilde{G} = 0 \ .$$

By the arbitrariness of F and G this implies that $\mathcal{J}$ is symplectic.

3. $\Rightarrow$ 4. Take $F = f_i$ and $G = f_j$. Then $\{f_i, f_j\}_x = \{y_i, y_j\}_y = J_{ij}$.

4. $\Rightarrow$ 3. Notice that $\{f_i, f_j\} = J_{ij}$ is the ij element of the matrix relation $\mathcal{J} J \mathcal{J}^T = J$.

5. $\Leftrightarrow$ 2. Since

$$\pi_K(f(x,t), t) = \frac{1}{2}(Jf) \cdot \mathrm{d}f - K\mathrm{d}t = \frac{1}{2}(\mathcal{J}^T J f) \cdot \mathrm{d}x + \left(\frac{1}{2}(Jf) \cdot \frac{\partial f}{\partial t} - K \right) \mathrm{d}t \ ,$$

The difference of 1-forms

$$\Delta\pi := \pi_H(x,t) - \pi_K(f,t) = \alpha \cdot \mathrm{d}x + \beta \mathrm{d}t \tag{2.49}$$

turns out to be defined by the quantities

$$\alpha := \frac{1}{2}\left(Jx - \mathcal{J}^T J f \right) \quad ; \quad \beta := K - H - \frac{1}{2}(Jf) \cdot \frac{\partial f}{\partial t} \ . \tag{2.50}$$

At this stage H is any given Hamiltonian and K is arbitrary. One easily checks that

$$\frac{\partial \alpha}{\partial x} - \left(\frac{\partial \alpha}{\partial x} \right)^T = J - \mathcal{J}^T J \mathcal{J} \ ; \tag{2.51}$$

$$\frac{\partial \alpha}{\partial t} - \frac{\partial \beta}{\partial x} = -\frac{\partial (K - H)}{\partial x} - \mathcal{J}^T J \frac{\partial f}{\partial t} \ . \tag{2.52}$$

From the latter relations one sees that $\Delta\pi$ is exact (i.e. the left hand sides of (2.51) and (2.52) vanish) iff $\mathcal{J}^T$, and thus $\mathcal{J}$, is symplectic, and

$$\frac{\partial f}{\partial t} = J \mathcal{J}^{-T} \nabla_x (K - H) = J \nabla_y (K - \tilde{H}) \ ,$$

which concludes the proof. $\triangleright$

It is convenient to rewrite some of the canonicity conditions in terms of the canonical variables $x = (q, p)$ and $y = (Q, P)$. Observe that, from what has been shown above, the notation $f : (q, p, H) \mapsto (Q, P, K)$ is more appropriate to specify a canonical transformation.

First of all let us rewrite the Poincaré-Cartan 1-form as follows

$$\begin{aligned}\pi_H(q,p,t) &:= \frac{1}{2}(Jx)\cdot\mathrm{d}x - H(x,t)\mathrm{d}t = \frac{1}{2}(p\cdot\mathrm{d}q - q\cdot\mathrm{d}p) - H(q,p,t)\mathrm{d}t = \\ &= p\cdot\mathrm{d}q - H(q,p,t)\mathrm{d}t - \frac{1}{2}\mathrm{d}(p\cdot q) := \hat{\pi}_H(q,p,t) + \mathrm{d}\varphi(q,p)\ ,\end{aligned}$$

where $\hat{\pi}_H := p\cdot\mathrm{d}q - H\mathrm{d}t$ and $\varphi(q,p) = -(p\cdot q)/2$. One sees that the 1-forms π_H and $\hat{\pi}_H$ differ by an exact differential. The 1-form $\hat{\pi}_H$ is suggested by the structure of the Hamiltonian action. The condition 5. of exactness of $\Delta\pi := \pi_H - \pi_K$ is equivalent to the condition of exactness of the equivalent 1-form $\Delta\hat{\pi} := \hat{\pi}_H - \hat{\pi}_K = \mathrm{d}F$, namely

$$\mathrm{d}F(q,Q,t) = p\cdot\mathrm{d}q - P\cdot\mathrm{d}Q + (K-H)\mathrm{d}t\ , \tag{2.53}$$

The latter relation implies

$$\frac{\partial F}{\partial q} = p\ \ ;\ \ \frac{\partial F}{\partial Q} = -P\ \ ;\ \ \frac{\partial F}{\partial t} = K - H\ . \tag{2.54}$$

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

$$\det\left(\frac{\partial^2 F}{\partial q\partial Q}\right) \neq 0\ , \tag{2.55}$$

which allows (by the implicit function theorem) to invert either the first or the second of relations (2.54). Indeed, starting from the first of (2.54) one gets $Q = V(q,p,t)$ which, substituted in the second relation yields $P = -\partial F/\partial Q(q,V,t) = U(q,p,t)$. 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 (2.54) yields the new Hamiltonian $K(Q,P,t) = [H(v,u,t) + \partial F/\partial t(v,Q,t)]$.

Very often one needs to generate canonical transformations by means of a generating function of q and P, for example. This is easily realized starting from (2.53) and defining the generating function $S(q,P,t) := F(q,Q,t) + Q\cdot P$, satisfying

$$\mathrm{d}S = p\cdot\mathrm{d}q + Q\cdot\mathrm{d}P + (K-H)\mathrm{d}t\ , \tag{2.56}$$

which implies

$$\frac{\partial S}{\partial q} = p\ \ ;\ \ \frac{\partial S}{\partial P} = Q\ \ ;\ \ \frac{\partial S}{\partial t} = K - H\ . \tag{2.57}$$

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

$$\det\left(\frac{\partial^2 S}{\partial q\partial P}\right) \neq 0\ , \tag{2.58}$$

with reasonings similar to those made above for F.

Exercise 2.5. *Consider the Hamiltonian $H'(q,p,t)$, defined in (2.7) and corresponding to the a gauge shifted Lagrangian with $a=1$. Show that the time-dependent transformation $(q,p)\mapsto(Q,P)$ defined by $P=(p-\nabla F)$, $Q=q$, is a canonical transformation, generated by $S(q,P,t)=q\cdot P+F(q,t)+\psi(t)$. Show that the new Hamiltonian is*

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

Observe that H and $H+\dot{\psi}(t)$ are equivalent. Apply all this to the case of the particle in the e.m. field described by the Hamiltonian (2.6).

Exercise 2.6. *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 2.7. *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 (2.53); as an alternative, invert first $P=U(p,t)$ and then look for $S(q,P,t)$.*

Let us rewrite now the condition of symplectic Jacobian $\mathcal{J}J\mathcal{J}^T=J$, where $\mathcal{J}=\partial f/\partial x$, in canonical variables. If $x=(q,p)$ and $y=(Q,P)=f(q,p)$, one has

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

which is equivalent to the relations

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

for all $i,j=1,\dots,n$. Such relations are the necessary and sufficient conditions for a change of variables to be canonical in Hamiltonian mechanics. Condition 3. of the theorem above state that a transformation is canonical iff it preserves the Poisson bracket of any two observables, namely $\{F, G\}_{q,p} = \{\tilde{F}, \tilde{G}\}_{Q,P}$, whereas condition 4. states that this is the case iff the transformation preserves the fundamental brackets, namely the brackets of any two canonical variables. Indeed the relations

$$\{Q_i, Q_j\}_{q,p} = \{Q_i, Q_j\}_{Q,P} = 0 \ ;$$
$$\{P_i, P_j\}_{q,p} = \{P_i, P_j\}_{Q,P} = 0 \ ;$$
$$\{Q_i, P_j\}_{q,p} = \{Q_i, P_j\}_{Q,P} = \delta_{i,j} \ ,$$

are equivalent to those written above (here the computation is made on the right hand side of each relation and is trivial, recalling that the Q's and P's are independent of each other).

2.5 Canonical rescaling

Let us see which rescalings of the canonical variables, Hamiltonian and time leave the Hamilton equations invariant in form. This means setting

$$Q_i = \alpha_i q_i \ , \ P_i = \beta_i p_i \ (i = 1, \dots, n) \ ; \ K = aH \ , \ T = bt \ , \tag{2.59}$$

$\alpha_1, \dots, \alpha_n$, $\beta_1, \dots, \beta_n$, a and b being $2n + 2$ real parameters. Assuming now that $\mathrm{d}q_i/\mathrm{d}t = \partial H/\partial p_i$ and $\mathrm{d}p_i/\mathrm{d}t = -\partial H/\partial q_i$, one gets,

$$\frac{\mathrm{d}Q_i}{\mathrm{d}T} = \left(\frac{\alpha_i\beta_i}{ab}\right)\frac{\partial K}{\partial P_i} \quad , \quad \frac{\mathrm{d}P_i}{\mathrm{d}T} = -\left(\frac{\alpha_i\beta_i}{ab}\right)\frac{\partial K}{\partial Q_i} \ . \tag{2.60}$$

One thus sees that the rescaling (2.59) leaves the Hamilton equations invariant in form, and is thus *canonical in extended sense*, if the condition

$$\alpha_i\beta_i = ab \ , \ \forall i = 1, \dots, n \tag{2.61}$$

holds. Notice that if one restricts to the case $ab = 1$, which includes that of no time-rescaling a priori, i.e. $b = 1$, and no rescaling of the Hamiltonian a priori, then condition (2.61) reduces to $\alpha_i\beta_i = 1$ for any i. One can easily prove that the rescaling in this case is canonical in strict sense, namely the Jacobian of the transformation (2.59) is symplectic. Indeed, one has

$$\mathcal{J} := \frac{\partial(Q, P)}{\partial(q, p)} = \mathrm{diag}(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n) \ ,$$

and one checks that the latter matrix is symplectic, i.e. $\mathcal{J} J \mathcal{J}^T = J$. In the general case (2.61), the latter computation yields $\mathcal{J} J \mathcal{J}^T = (ab)J$.

2.6 Complex Birkhoff variables and harmonic angle-action variables

Let us consider the Hamiltonian of a harmonic oscillator $H = (p^2 + \omega^2 q^2)/2$. The equations of motion are $\dot{q} = p$, $\dot{p} = -\omega^2 q$. One can then introduce the complex variable

$$z := \frac{\omega q + \imath p}{\sqrt{2\omega}} e^{\imath\theta} \ , \tag{2.62}$$

where θ is an arbitrary angle. In terms of the latter variable, the Hamilton equations read $\imath\dot{z} = \omega z = \partial H/\partial\bar{z}$, where $H = \omega|z|^2$, so that $z(t) = z(0)e^{-\imath\omega t}$. One easily checks that such a complex form of the Hamilton equations is actually canonical. Indeed, one easily computes the Poisson bracket

$$\{z, \bar{z}\} = \frac{1}{2\omega}\{\omega q + \imath p, \omega q - \imath p\} = -\imath \ ,$$

while obviously $\{z, z\} = 0 = \{\bar{z}, \bar{z}\}$. Thus, setting $\pi := \imath\bar{z}$, one easily finds that the harmonic oscillator problem can be rewritten in complex canonical form with Hamiltonian $H = -\imath\omega z\pi$ and equations of motion $\dot{z} = \partial H/\partial\pi$ and $\dot{\pi} = -\partial H/\partial z$. The harmonic oscillator can be conveniently treated in terms of another canonical pair of variables, namely the so-called harmonic angle-action variables (φ, I) defined by $z = \sqrt{I}e^{-\imath\varphi}$, or $I = |z|^2$ and $\varphi = \frac{1}{2\imath}\ln(\bar{z}/z)$. One easily shows that the variables (φ, I) are canonical, since $\{\varphi, I\} = 1$. The Hamiltonian of the oscillator becomes $H = \omega I$ and the corresponding equations of motion read $\dot{\varphi} = \partial H/\partial I = \omega$, $\dot{I} = -\partial H/\partial\varphi = 0$, so that $\varphi(t) = \varphi(0) + \omega t$ and $I(t) = I(0)$.

The harmonic oscillator problem treated above motivates the introduction of the canonical transformation $\mathbb{R}^{2n} \ni (q, p) \mapsto (z, \pi) \in \mathbb{C}^{2n}$ to the complex Birkhoff variables defined by

$$z_j := \frac{a_j q_j + \imath p_j}{\sqrt{2a_j}} e^{\imath\theta_j} \quad ; \quad \pi_j := \imath\bar{z}_j \ , \tag{2.63}$$

where $a_1, \dots, a_n$ are positive parameters and $\theta_1, \dots, \theta_n$ angles. The variables (2.63) are canonical: $\{z_j, \pi_k\} = \delta_{jk}$, $\{z_j, z_k\} = 0 = \{\pi_j, \pi_k\}$ for any $j, k = 1, \dots, n$. The Poisson bracket of any pair of functions F, G expressed in terms of the complex Birkhoff variables reads

$$\{F, G\} = \sum_{j=1}^{n} \left(\frac{\partial F}{\partial z_j}\frac{\partial G}{\partial \pi_j} - \frac{\partial F}{\partial \pi_j}\frac{\partial G}{\partial z_j} \right) = -\imath \sum_{j=1}^{n} \left(\frac{\partial F}{\partial z_j}\frac{\partial G}{\partial \bar{z}_j} - \frac{\partial F}{\partial \bar{z}_j}\frac{\partial G}{\partial z_j} \right) \ . \tag{2.64}$$

The Birkhoff variables (2.63) turn out to be the right variables when studying systems of noninteracting harmonic oscillators and their perturbations. In such a case, the parameters a_j are chosen to coincide with the frequencies of the oscillators. In complete analogy with what done for a single harmonic oscillator, one can introduce the canonical, harmonic angle-action variables (φ, I) defined by $z_j = \sqrt{I_j}e^{-\imath\varphi_j}$ for any $j = 1, \dots, n$.

2.7 Hamilton-Jacobi equation

Suppose that one looks for a canonical change of variables $(q, p, H) \mapsto (Q, P, K)$ such that $K = K(P, t)$ depends on the new momenta P and time only. This amounts to look for a canonical transformation, generated by $\hat{S}(q, P, t)$ such that the new momenta P do not evolve in time. More precisely, the Hamilton equations in the new variables read $\dot{Q} = \partial K/\partial P$ and $\dot{P} = -\partial K/\partial Q = 0$, so that P is constant and $Q(t) = Q(0) + \int_0^t \nabla_P K(P, s)\mathrm{d}s$. With this in mind, the first and the third of relations (2.57) yield the equation

$$\frac{\partial\hat{S}}{\partial t} + H\left(q, \frac{\partial\hat{S}}{\partial q}, t\right) = K(P, t) \ , \tag{2.65}$$

a first order PDE in the unknown function $\hat{S}(q, P, t)$ and $K(P, t)$. However,one can eliminate the latter by defining a new generating function $S(q, P, t) := \hat{S}(q, P, t) - \int_0^t K(P, s)\mathrm{d}s$, which satisfies the same relations (2.57) with $K \equiv 0$. Then $S(q, P, t)$ satisfies the *Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}, t\right) = 0 \ , \tag{2.66}$$

introduced in a previous section. Notice that, among the possible solutions of equation (2.66), we are interested in the so-called *complete integrals*, namely those solutions depending on n parameters $P_1, \ldots, P_n$ and such that (2.58) holds.

A complete integral of the Hamilton-Jacobi equation (2.66) generates a canonical transformation $(q, p, H) \mapsto (Q, P, 0)$.

Exercise 2.8. *Show that if $\hat{S}$ generates a canonical transformation to $K(P, t)$ such that $Q(t) = Q' + \int_0^t \nabla_P K(P, s)\mathrm{d}s$, then $S = \hat{S} - \int_0^t K(P, s)\mathrm{d}s$ generates a canonical transformation to a null Hamiltonian such that $Q(t) = Q'$. Hint: make use of (2.56).*

Now, the Hamilton equations in the new variables read $\dot{Q} = 0$ and $\dot{P} = 0$, so that Q and P are constant. As already seen, by means of equations (2.57), one can get either $q = v(Q, P, t)$ and $p = u(Q, P, t)$ that solve the H-equations with Hamiltonian H, or their inverse $Q = V(q, p, t)$ and $P = U(q, p, t)$. 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:

$$\dot{P}_i = \{U_i, H\}_{q,p} + \frac{\partial U_i}{\partial t} = 0 \quad i = 1, \ldots, n \ . \tag{2.67}$$

Moreover, the functions U_i are new momenta and as a consequence $\{U_i, U_j\}_{q,p} = \{P_i, P_j\}_{Q,P} = 0$. In conclusion, to a complete integral of the H-J equation (2.66) there correspond n first integrals in involution (each pair of them having zero P-bracket) and the H-equations can be solved, in principle.

Remark 2.4. *If no other particular hypothesis is made, solving the H-J equation is as difficult as solving the original Hamilton equations, and such a method turns out to be useless.*

If H is independent of time, and one looks for time-independent canonical transformations $(\partial S/\partial t = 0)$ such that the new Hamiltonian $K = K(P)$ depends on the momenta P only, equation (2.65) yields the time-independent H-J equation

$$H\left(q, \frac{\partial S}{\partial q}(q, P)\right) = K(P) \ . \tag{2.68}$$

Notice that the new Hamiltonian K is an unknown of the problem and cannot be eliminated. In this 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. Here again, a complete integral of equation (2.68) yields n first integrals in involution (namely the $P_i = U_i(q, p)$) and allows to solve the original H-equations, in principle.

Chapter 3

Integrable systems

3.1 Introduction

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.

In order to understand this, we start by supposing that the system 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

$$\det\left(\frac{\partial f}{\partial p}\right) = \det\left(\frac{\partial(f_1,\dots,f_n)}{\partial(p_1,\dots,p_n)}\right) \neq 0 \ , \tag{3.1}$$

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

$$p_1 = u_1(q,t;a) \ ; \ \dots \ p_n = u_n(q,t;a) \ . \tag{3.2}$$

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

$$\left.\frac{\partial u_i}{\partial t} + \sum_{j=1}^{n}\left(\frac{\partial u_i}{\partial q_j} - \frac{\partial u_j}{\partial q_i}\right)\frac{\partial H}{\partial p_j} = -\frac{\partial H}{\partial q_i} - \sum_{j=1}^{n}\frac{\partial u_j}{\partial q_i}\frac{\partial H}{\partial p_j}\right|_{p=u(q,t;a)} \ . \tag{3.3}$$

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

$$\mathrm{rot}(u) := \left(\frac{\partial u}{\partial q}\right) - \left(\frac{\partial u}{\partial q}\right)^T \ , \tag{3.4}$$

$$v(q,t) := \left.\frac{\partial H}{\partial p}\right|_{p=u(q,t;a)} , \tag{3.5}$$

and

$$h(q,t) := H(q, u(q,t;a), t) , \tag{3.6}$$

the equations (3.3) can be rewritten in compact, vector form as

$$\frac{\partial u}{\partial t} + \mathrm{rot}(u)v = -\nabla_q h . \tag{3.7}$$

To such an equation, we can associate a continuity equation for an unknown density ρ associated to the vector velocity field (3.5), i.e. to the vector ODE $\dot{q} = v(q,t)$, namely

$$\frac{\partial \rho}{\partial t} + \nabla_q \cdot (\rho v) = 0 . \tag{3.8}$$

The meaning of $\rho(q,t)$ is that of a probability density to find the representative point of the system "close" to the point q in the configuration space at time t.

We now notice the similarity of the system (3.7)-(3.8) with the Euler equations of ideal hydrodynamics, namely

$$\frac{\partial \boldsymbol{u}}{\partial t} + \mathrm{rot}(\boldsymbol{u})\boldsymbol{u} = -\nabla\left(\frac{|\boldsymbol{u}|^2}{2} + w\right) ; \tag{3.9}$$

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \boldsymbol{u}) = 0 , \tag{3.10}$$

where ϱ is the mass density, $\boldsymbol{u}$ is the velocity field, w is the enthalpy function such that $\nabla w = \nabla p/\varrho$ (assuming a functional dependence of $\varrho(p)$ or $p = p(\varrho)$), 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 (3.7) and (3.9) 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 (3.7) takes the rather simple form

$$\frac{\partial u}{\partial t} + \mathrm{rot}(u)M^{-1}u = -\nabla_q\left(\frac{u \cdot M^{-1}u}{2} + V\right) . \tag{3.11}$$

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 the enthalpy function.

Remark 3.1. *Attention has to be paid to the fact that for the Euler equation (3.9) the enthalpy w depends on the density ϱ, while nothing similar holds in equation (3.7). A dependence of the effective potential energy on the probability density ρ is instead characteristic of quantum mechanics.*

Now, by analogy with the case of fluids, we look for curl-free, i.e. irrotational solutions of the Euler-like equation (3.7) (we recall that in fluid dynamics, looking for a solution of the Euler equation (3.9) of the form $\boldsymbol{u} = \nabla\phi$ leads to the Bernoulli equation for the velocity potential ϕ, namely $\partial\phi/\partial t + |\nabla\phi|^2/2 + w = constant$). 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 (3.7) and lifting a gradient, one gets

$$\frac{\partial S}{\partial t} + H(q, \nabla_q S, t) = K(t; a) \ . \tag{3.12}$$

One can set $K(t; a) \equiv 0$ without any loss of generality, and the latter equation becomes the time-dependent Hamilton-Jacobi equation (if $\varphi \not\equiv 0$ then $\widetilde{S} := S - \int K\mathrm{d}t$ satisfies equation (3.12) 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

$$f_i(q, u(q, t; a), t) \equiv a_i \ , \tag{3.13}$$

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 (3.7). 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 (3.13), 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 (3.12) with $K = 0$, generates a canonical transformation $(q,p,H) \mapsto (b,a,0)$ to a zero Hamiltonian, transformation defined by the implicit equations $p = \nabla_q S(q,t;a)$, $b := \nabla_a S(q,t;a)$. The restriction to the case where H, $f_1,\dots,f_n$ are independent of time is completely analogous to that just treated and is left as an exercise.

What has just been shown above motivates the following definition of integrable Hamiltonian system.

Definition 3.1. *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 $f_1(q,p,t),\dots,f_n(q,p,t)$ 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 $j = 1,\dots,n$;*
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; e.g. (3.1) holds) for any (q,p,t);*
3. *$\{f_j, f_k\} = 0$ for any $j,k = 1,\dots,n$.*

The following theorem holds.

Theorem 3.1 (Liouville-Jacobi)**.** *Let the Hamiltonian system defined by $H(q,p,t)$ is Liouville-integrable and evolves on the n-dimensional time-dependent invariant manifold*

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

iff there exists a complete integral $S(q,t;a)$ of the Hamilton-Jacobi equation.

$\lhd$ PROOF. It has essentially been given in all details and left as an exercise. $\rhd$

Remark 3.2. *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 $H(\mathscr{C}^{-1}(b,a)) = K(a)$.*

Example 3.1. *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 3.2. *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 3.2 (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*

$$M_a := \{(q,p) \in \Gamma \ : \ f_1(q,p) = a_1, \dots, f_n(q,p) = a_n\} \tag{3.14}$$

Ma 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* $\mathbb{T}^n = \overline{\mathbb{T}^1 \times \cdots \times \mathbb{T}^1}$ *(*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

$$\omega(I) := \frac{\partial E(I)}{\partial I} \ . \tag{3.15}$$

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 (3.15) depending, in general, on the value of the action I_0 labeling the torus $\mathbb{T}^n$.

Example 3.3. *Consider a system of* n *noninteracting harmonic oscillators with Hamiltonian* $H = \sum_{k=1}^{n} (p_k^2 + \omega_k^2 q_k^2)/2$. *In this case,* $f_k = (p_k^2 + \omega_k^2 q_k^2)/2$, *and the invariant manifold* M_a *is the product of* n *ellipses (which ones?). Passing to Birkhoff comeplex variables first, defined by* $z := (\omega_k q_k + \imath p_k)/\sqrt{2\omega_k}$, *one gets* $H = \sum_k \omega_k |z_k|^2$. *Then, canonical, harmonic angle-action variables* (φ, I) *are introduced by setting* $z_k = \sqrt{I_k} e^{-\imath \varphi_k}$, *which yields* $H = \sum_k \omega_k I_k$, *so that the dynamics takes place on the torus* $\mathbb{T}^n$ *labelled by the constant vector* I. *Observe that if a term* $+\lambda \sum_k q_k^4$, $\lambda > 0$, *is added to* H, *then one can still pass to the harmonic angle-action variables, but the Hamiltonian will no longer depend only on* I*: the harmonic angle-action variables are not those predicted by the Arnol'd theorem in this case.*

3.2 Separable systems

Integrable systems (in the sense of Liouville-Arnol'd) display a "trivial" dynamics, linear in the co-ordinates (angles), the conjugate momenta (actions) being constant. However, even when one knows a priori that a given system is integrable, i.e. one is able to write down the n first integrals in involution, finding the canonical transformation to action-angle variables may be completely non trivial. A class of integrable systems for which action-angle variables can be found "in principle", i.e. solving integrals, is that of *separable systems*.

Definition 3.2. *A Liouville integrable, autonomous Hamiltonian system defined by H is said to be separable if the corresponding HJ equation (2.68) admits a (local) complete integral of the form*

$$S(q,P) = \sum_{i=1}^{n} S_i(q_i, P) \ . \tag{3.16}$$

Separability of course depends both on the Hamiltonian and on the particular choice of the canonical coordinates: for a given H, a system can result separable with certain coordinates and not with other ones. On the other hand, integrability is independent of the coordinates. Separability is equivalent to state that $p_i = u_i(q_i, P) := \partial S_i/\partial q_i$ depends only on q_i.

Example 3.4. *Of course, if $n = 1$ the system is always separable. Indeed, $H(q, p) = E$ implies $p = u(q, E) = \partial S/\partial q$, locally (think of a Hamiltonian quadratic in the momenta: then in writing p as a function of q one has to make a choice of sign). Here $E = K(P) = P$ is the new Hamiltonian, coinciding with the new momentum; then $\dot{Q} = \partial K/\partial P = 1$ implies that the new coordinate Q is time, up to translation.*

Example 3.5. *Suppose that $H = \sum_{i=1}^{n} H_i(q_i, p_i)$ is the sum of one degree of freedom Hamiltonians. The n Hamiltonians H_i are clearly first integrals in involution, since H_i and H_j are functions of different pairs of canonical coordinates for $i \neq j$. Thus $H_i(q_i, p_i) = P_i$, and $p_i = u_i(q_i, Pi) = \partial S_i/\partial q_i$, locally. In this case $H = K(P) = \sum_{i=1}^{n} P_i$. As a consequence, $\dot{Q}_i = \partial K/\partial P_i = 1$ and $Q_i = t + Q_i(0)$.*

Example 3.6. *Suppose $H(q, p) = K(\Phi_1(q_1, p_1), \ldots, \Phi_n(q_n, p_n))$. Then the Φ_i are n fist integrals in involution (prove it). Setting $\Phi_i(q_i, p_i) = P_i$ one gets $p_i = u_i(q_i, P_i) = \partial S_i/\partial q_i$ locally, and the new Hamiltonian $K(P_1, \ldots, P_n)$. As a consequence, $\dot{Q}_i = \partial K/\partial P_i$ and $Q_i = (\partial K/\partial P_i)t + Q_i(0)$.*

Example 3.7. *A less trivial example is the following. Suppose that the Hamiltonian of the system has the* matryoshka *structure*

$$H(q,p) = \Phi_n(q_n, p_n, \xi_{n-1}) \ ; \ \xi_j = \Phi_j(q_j, p_j, \xi_{j-1}) \ , \ j = 1, \ldots, n-1 \ , \ \xi_0 = 0 \ . \tag{3.17}$$

Now, $\Phi_i(q_i, p_i, \xi_{i-1})$ depends only on the canonical pairs $(q_1, p_1), \ldots, (q_i, p_i)$, so that

$$\begin{aligned}
\{\Phi_i, \Phi_j\}_{i>j} &= \sum_{k=1}^{n} \left(\frac{\partial \Phi_i}{\partial q_k}\frac{\partial \Phi_j}{\partial p_k} - \frac{\partial \Phi_i}{\partial p_k}\frac{\partial \Phi_j}{\partial q_k} \right) = \sum_{k=1}^{j} \left(\frac{\partial \Phi_i}{\partial q_k}\frac{\partial \Phi_j}{\partial p_k} - \frac{\partial \Phi_i}{\partial p_k}\frac{\partial \Phi_j}{\partial q_k} \right) = \\
&= \sum_{k=1}^{j} \left(\frac{\partial \Phi_i}{\partial \xi_{i-1}}\frac{\partial \Phi_{i-1}}{\partial \xi_{i-2}} \cdots \frac{\partial \Phi_k}{\partial q_k}\frac{\partial \Phi_j}{\partial \xi_{j-1}}\frac{\partial \Phi_{j-1}}{\partial \xi_{j-2}} \cdots \frac{\partial \Phi_k}{\partial p_k} + \right. \\
&- \left. \frac{\partial \Phi_i}{\partial \xi_{i-1}}\frac{\partial \Phi_{i-1}}{\partial \xi_{i-2}} \cdots \frac{\partial \Phi_k}{\partial p_k}\frac{\partial \Phi_j}{\partial \xi_{j-1}}\frac{\partial \Phi_{j-1}}{\partial \xi_{j-2}} \cdots \frac{\partial \Phi_k}{\partial q_k} \right) \equiv 0 \ .
\end{aligned} \tag{3.18}$$

Thus the Φ_i are first integrals in involution (why they are first integrals?). Let us set $\Phi_1(q_1,p_1) = P_1$, $\Phi_2(q_2,p_2,P_1) = P_2,\dots$, $\Phi_i(q_i,p_i,P_{i-1}) = P_i,\dots,\Phi_n(q_n,p_n,P_{n-1}) = P_n$. Then $p_1 = u_1(q_1,P_1) = \partial S_1/\partial q_1$, $p_i = u_i(q_i,P_{i-1},P_i) = \partial S_i/\partial q_i$. The new Hamiltonian is thus $K(P) = P_n$, so that $Q_i(t) = \delta_{i,n}t + Q_i(0)$.

In all the treated examples $S(q,P) = \sum_{i=1}^n S_i(q_i,P) = \sum_{i=1}^n \int^{q_i} u_i(s,P)\mathrm{d}s$.

3.2.1 The Stäckel theorem

An interesting question naturally arises: is it possible to characterize the separable systems? In other words, do necessary and sufficient conditions exist for separability? A positive answer exists in the case of those Hamiltonian systems whose kinetic energy that is a diagonal quadratic form in the momenta, with coefficients possibly depending on the coordinates, namely:

$$H(q,p) = K(q,p) + U(q) := \sum_{i=1}^n c_i(q)\frac{p_i^2}{2} + U(q) \ . \tag{3.19}$$

Such systems will be simply referred to as diagonal or orthogonal. The reason for the latter definition arises in the Lagrangian formalism for constrained systems whose constraint manifold is defined by $\mathbb{R}^n \ni q \mapsto X(q) \in \mathbb{R}^N$ ($n \le N$; if $n = N$ the system is not constrained and the latter map just defines a change of coordinates). Then, if M denotes the diagonal matrix of the particle masses, the kinetic energy of the system is

$$K = \frac{\dot{X}\cdot M\dot{X}}{2} = \sum_{i,j=1}^n \frac{1}{2}\left(\frac{\partial X}{\partial q_i}\cdot M\frac{\partial X}{\partial q_j}\right)\dot{q}_i\dot{q}_j \ .$$

The system of Lagrangian coordinates q is orthogonal, by definition, if

$$\frac{\partial X(q)}{\partial q_i}\cdot M\frac{\partial X(q)}{\partial q_j} = \frac{1}{c_i(q)}\ \delta_{i,j} \ ,$$

i.e. if the coordinate curves on the manifold intersect orthogonally to each other (with respect to the scalar product defined by M). In this case $K = \sum_{i=1}^n \dot{q}_i^2/(2c_i)$ and the Hamiltonian corresponding to the Lagrangian $L = K - U$, where $U(q) := U(X(q))$, takes the form (3.19) (show it explicitly). Let us denote by $c(q)$ the n-dimensional vector whose components are the coefficients $c_i(q)$ entering K in (3.19); let us also denote by $\hat{e}_k$ the k-th vector of the canonical basis of $\mathbb{R}^n$: $(\hat{e}_k)_j = \delta_{j,k}$. The following interesting theorem characterizing diagonal systems holds.

Theorem 3.3 (Stäckel)**.** *The diagonal Hamiltonian system (3.19) is separable iff a vector $\psi(q)$ and a non singular matrix $A(q)$ can be found such that*

1. *$\psi_i = \psi_i(q_i)$, $A_{ij} = A_{ij}(q_i)$, for any $i,j = 1,\dots,n$;*
2. *$c(q)\cdot\psi(q) = U(q)$, $A^T(q)c(q) = \hat{e}_k$, for some $k = 1,\dots,n$.*

$\lhd$ PROOF. As a preliminary comment, let us observe that the index k in the statement is arbitrary: if $A^T c = \hat{e}_k$, then suitably exchanging the rows of A^T one moves k from 1 to n.

Let us first prove necessity. Suppose that the system is separable. Then there exists a complete integral $S(q,P) = \sum_{i=1}^n S_i(q_i,P)$ of the HJ equation:

$$\sum_{i=1}^n c_i(q) \frac{1}{2} \left(\frac{\partial S_i}{\partial q_i} \right)^2 + U(q) = K(P) \ . \tag{3.20}$$

We can always redefine the constant parameters $P \mapsto \alpha$ in such a way that $\alpha_k = K(P)$, $\alpha_i = P_i$ for $i = 1,\dots,n$, $i \neq k$, where the index k is so chosen that $\partial K/\partial P_k \neq 0$. Then the inverse transformation $\alpha \mapsto P$ is obtained by making P_k explicit from $K(\dots,\alpha_{k-1},P_k,\alpha_{k+1},\dots) = \alpha_k$. The HJ equation (3.20) reads then

$$\sum_{i=1}^n c_i(q) \frac{1}{2} \left(\frac{\partial S_i}{\partial q_i}(q_i,\alpha) \right)^2 + U(q) = \alpha_k \ . \tag{3.21}$$

Let now $\bar{\alpha}$ be any fixed value of the vector α of the parameter for which $S(q,\bar{\alpha})$ solves (3.21). By a suitable translation in the space of parameters we can always assume $\bar{\alpha} = 0$. Let us set $G_i(q_i,\alpha) := (\partial S_i/\partial q_i)^2$, and expand the left hand side of (3.21) around $\alpha = \bar{\alpha} = 0$. This yields

$$\sum_{i=1}^n \frac{c_i(q)}{2} \left[G_i(q_i,0) + \sum_{j=1}^n \alpha_j \frac{\partial G_i}{\partial \alpha_j}(q_i,0) + O(|\alpha|^2) \right] + U(q) = \alpha_k$$

which, upon defining

$$\psi_i(q_i) := -\frac{1}{2} G_i(q_i,0) \quad ; \quad A_{ij}(q_i) := \frac{1}{2} \frac{\partial G_i}{\partial \alpha_j}(q_i,0) \ , \tag{3.22}$$

reads

$$-c(q) \cdot \psi(q) + U(q) + c(q) \cdot A(q) \alpha + O(|\alpha|^2) = \alpha_k \ .$$

The latter identity implies $c \cdot \psi = U$ and (by the arbitrariness of α) $A^T c = \hat{e}_k$. It is left as an exercise to check that the completeness of S implies that the matrix A, defined in (3.22), is non singular. This proves the necessity of the above conditions 1. and 2. to separability.

On the contrary, let us suppose that conditions 1. and 2. are satisfied. Let us define $G_i(q) := (\partial S/\partial q_i)^2$. The HJ equation then reads

$$c(q) \cdot \left[\frac{1}{2} G(q) + \psi(q) \right] = \alpha_k \ , \tag{3.23}$$

where we have set $K(P) = \alpha_k$, the index k being given by the hypothesis (this defines a transformation $P \mapsto \alpha$ of the parameters, as above). Now, $A^T c = \hat{e}_k$ implies $c = A^{-T} \hat{e}_k$; moreover $\alpha_k = \hat{e}_k \cdot \alpha$. Equation (3.23) then becomes

$$\hat{e}_k \cdot \left[A^{-1} \left(\frac{1}{2} G + \psi \right) - \alpha \right] = 0 \ . \tag{3.24}$$

By the arbitrariness of k, the square bracket must vanish identically, which in turn implies $G = -2\psi + 2A\alpha$ or, in components

$$G_i(q) := \left(\frac{\partial S}{\partial q_i}\right)^2 = -2\psi_i(q_i) + 2\sum_{j=1}^{n} A_{ij}(q_i)\alpha_j \ . \tag{3.25}$$

Thus $\partial S/\partial q_i$ is a function of q_i only, so that $S(q,\alpha) = \sum_{i=1}^{n} S_i(q_i,\alpha)$, i.e. the system is separable. $\triangleright$

Example 3.8. *Let us consider the Hamiltonian of a particle of mass m moving in the 3-dimensional space subject to a conservative force of potential energy $U(x)$. In cartesian co-ordinates the Hamiltonian reads $H(x,p) = |p|^2/(2m) + U(x)$. Here $c_i = 1/m$, $i = 1,2,3$. By the Stäckel theorem the system is separable if there exist $\psi_i(x_i)$ such that $U(x) = c\cdot\psi = \sum_{i=1}^{3}\psi_i(x_i)/m$, i.e. if the potential energy is the sum of terms each depending on one of the coordinates only. This is the kind of separability of the Example 3.5: calling $v_i(x_i) := \psi_i(x_i)/m$, the Hamiltonian reads $H = \sum_{i=1}^{3}[\frac{p_i^2}{2m} + v_i(x_i)] := \sum_{i=1}^{3} H_i(x_i,p_i)$. One has $H_i = P_i$ constant. Then, setting $P_1 + P_2 + P_3 = \alpha_1$, $P_2 = \alpha_2$ and $P_3 = \alpha_3$, the HJ equation $H(x, \partial S/\partial x) = \alpha_1$, with $S = S_1 + S_2 + S_3$, splits into the three HJ equations*

$$\left(\frac{\partial S_1}{\partial x_1}\right)^2 = -2mv_1(x_1) + 2m(\alpha_1 - \alpha_2 - \alpha_3) \ ;$$
$$\left(\frac{\partial S_2}{\partial x_2}\right)^2 = -2mv_2(x_2) + 2m\alpha_2 \ ;$$
$$\left(\frac{\partial S_3}{\partial x_3}\right)^2 = -2mv_3(x_3) + 2m\alpha_3 \ .$$

Comparing with (3.25), one easily finds that

$$A = \begin{pmatrix} m & -m & -m \\ 0 & m & 0 \\ 0 & 0 & m \end{pmatrix}$$

satisfies $A^T c = \hat{e}_1$, consistent with the choice $H = K = \alpha_1$.

As already stressed, separability depends on the choice of the coordinates.

Example 3.9. *Let us reconsider the problem of the latter example in spherical polar coordinates (r,θ,φ), such that $x = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)^T$. One easily checks that $\partial x/\partial r$, $\partial x/\partial\theta$ and $\partial x/\partial\varphi$ are mutually orthogonal. The square of the infinitesimal displacement is $(\mathrm{d}s)^2 = (\mathrm{d}r)^2 + (r\mathrm{d}\theta)^2 + (r\sin\theta\mathrm{d}\varphi)^2$, so that the Kinetic energy is $K = m\dot{s}^2/2$ and the Lagrangian is*

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) - U(r,\theta,\varphi) \ .$$

The corresponding Hamiltonian (show it) is

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\varphi^2}{2mr^2\sin^2\theta} + U(r,\theta,\varphi) \ . \tag{3.26}$$

Here $c = (1/m, 1/(mr^2), 1/(mr^2\sin^2\theta))^T$. *The Stäckel theorem in this case requires, for separability, the existence of a vector* $\psi = (\psi_r(r), \psi_\theta(\theta), \psi_\varphi(\varphi))$ *such that* $U = c \cdot \psi$, *namely*

$$U(r,\theta,\varphi) = \frac{\psi_r(r)}{m} + \frac{\psi_\theta(\theta)}{mr^2} + \frac{\psi_\varphi(\varphi)}{mr^2\sin^2\theta} := u(r) + \frac{v(\theta)}{r^2} + \frac{w(\varphi)}{r^2\sin^2\theta} \ . \tag{3.27}$$

This is the most general form of separable potential energy in spherical polar coordinates. Observe that it includes the fundamental case of central motions, with $U = u(r)$ *(i.e.* $v = 0 = w$*). Upon rewriting the Hamiltonian (3.26) as follows*

$$H = \frac{p_r^2}{2m} + u(r) + \frac{1}{r^2}\left[\frac{p_\theta^2}{2m} + v(\theta) + \frac{1}{\sin^2\theta}\left(\frac{p_\varphi^2}{2m} + w(\varphi)\right)\right] \ , \tag{3.28}$$

one realizes that this is a separable system of the matryoshka *kind, treated in Example 3.7. Setting then* $S = S_r(r,\alpha) + S_\theta(\theta,\alpha) + S_\varphi(\varphi,\alpha)$, *the HJ equation splits into*

$$\begin{aligned}
&\frac{1}{2m}\left(\frac{\partial S_r}{\partial r}\right)^2 + u(r) + \frac{\alpha_2}{r^2} = \alpha_3 (=: K) \ ; \\
&\frac{1}{2m}\left(\frac{\partial S_\theta}{\partial \theta}\right)^2 + v(\theta) + \frac{\alpha_1}{\sin^2\theta} = \alpha_2 \ ; \\
&\frac{1}{2m}\left(\frac{\partial S_\varphi}{\partial \varphi}\right)^2 + w(\varphi) = \alpha_1 \ ;
\end{aligned}$$

The latter system can be rewritten in vector form and compared with (3.25), which yields

$$\begin{pmatrix} (\partial S_r/\partial r)^2 \\ (\partial S_\theta/\partial \theta)^2 \\ (\partial S_\varphi/\partial \varphi)^2 \end{pmatrix} = -2 \underbrace{\begin{pmatrix} mu(r) \\ mv(\theta) \\ mw(\varphi) \end{pmatrix}}_{\psi} + 2 \underbrace{\begin{pmatrix} 0 & -m/r^2 & m \\ -m/\sin^2\theta & m & 0 \\ m & 0 & 0 \end{pmatrix}}_{A} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \ .$$

One easily checks that $A^T c = \hat{e}_3$, *consistent with the choice* $H = K = \alpha_3$.

3.2.2 Action-angle variables

We now show how to construct the action-angle variables for separable systems, whose existence is guaranteed, under the hypothesis of compactness of the level set of the first integrals, by the Theorem 3.2. Now, first of all, the separability conditions (3.16) is equivalent to say that the n first integrals in involution of the system, the momenta P_i or suitable functions of them, are each a function of a pair of canonical variables (q_i, p_i). The sufficiency of such a condition to separability is trivial, its necessity is a bit less trivial to be proven. This implies that the level set of first integrals of the separable system is of the form $f_i(q_i, p_i, P) = \alpha_i$, $i = 1, \dots, n$. The compactness condition, in this case, implies two possibilities only: either the level curve $f_i = \alpha_i$ is closed, or is open and periodic and q_i is essentially an angle (viceversa, if q_i is an angle, the level curve can be closed). In the latter case, regarding the motion of q_i on the circle, one gets again a closed (thus compact) curve. Thus, a compact level set $f_i = \alpha_i$ of a separable system consists of the product of n closed curves $\gamma_i : s \mapsto (q_i(s), p_i(s))$, or *cycles*, that admit local representations of the graph form $p_i = u_i(q_i) = \partial S_i/\partial q_i$, for any $i = 1, \dots, n$.

Definition 3.3. *The quantities*

$$I_i := \frac{1}{2\pi} \oint_{\gamma_i} p_i \mathrm{d}q_i \tag{3.29}$$

are called action variables of the separable system at hand.

The action variables thus defined are functions of the parameters α defining the level set of the first integrals, i.e. $I = F(\alpha)$, so that $\alpha = G(I)$. Such relations allow to pass from the momenta α to the new ones, the action variables I. If the canonical transformation $(q, p) \mapsto (\beta, \alpha)$ is generated by $S(q, \alpha)$ such that

$$\mathrm{d}S = p \cdot \mathrm{d}q + \beta \cdot \mathrm{d}\alpha \ ,$$

then, the transformation $(\beta, \alpha) \mapsto (\varphi, I)$ is generated by $W(q, I) = S(q, G(I))$, namely

$$\mathrm{d}W = p \cdot \mathrm{d}q + \beta \cdot \frac{\partial G}{\partial I} \mathrm{d}I = p \cdot \mathrm{d}q + \varphi \cdot \mathrm{d}I \ ,$$

where $\varphi := (\partial G/\partial I)^T \beta = \partial W/\partial I$. In order to prove that the new coordinates φ are angles, one proves that the variation of φ_j along any cycle γ_i amounts to 2π if $j = i$ and zero otherwise. Indeed,

$$\oint_{\gamma_i} \mathrm{d}\varphi_j = \oint_{\gamma_i} \frac{\partial \varphi_j}{\partial q_i} \mathrm{d}q_i = \frac{\partial}{\partial I_j} \oint_{\gamma_i} \frac{\partial W}{\partial q_i} \mathrm{d}q_i = \frac{\partial}{\partial I_j} 2\pi I_i = 2\pi \delta_{ij} \ ,$$

where use has been made of $\varphi_j = \partial W/\partial I_j$ and $p_i = \partial S/\partial q_i = \partial W/\partial q_i$.

Example 3.10. *Autonomous Hamiltonian systems with one degree of freedom ($n = 1$) are obviously integrable. Any connected and compact component of the level curve $H(q, p) = E$, not containing critical points of H, is a periodic orbit. In that case the action $I = \frac{1}{2\pi} \oint p\mathrm{d}q$. The latter quantity is obviously a function of the energy level E: $I = f(E)$ and $E = g(I)$. In the simplest mechanical case where $H = p^2/2 + V(q)$, one has*

$$I = \frac{1}{\pi} \int_{q_-}^{q_+} \sqrt{2(E - V(q))} \ \mathrm{d}q \ , \tag{3.30}$$

where $V(q_\pm) = E$ and $V'(q_\pm) \neq 0$. As an example, for the harmonic oscillator, with $V(q) = \omega^2 q^2/2$, one finds $I = E/\omega$ (show this).

Example 3.11. *A further example is the pendulum of length ℓ, whose potential energy is $V(q) = \omega^2(1 - \cos q)$, where $\omega^2 = g/\ell$ and q is the angle with respect to the vertical axis. Here one has to distinguish between the values of the energy $E < 2\omega^2$ and $E > 2\omega^2$. In the former case the motion consists of oscillations between the turning points $\pm\bar{q}$, being $V(\pm\bar{q}) = E$ and the formula for I is given by (3.30); in the limit $E \to 0^+$ one finds $I \to E/\omega$ (why? show this). In the latter case the pendulum rotates and one has $I = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sqrt{2(E - V(q))} \ \mathrm{d}q$. In the limit $E \to +\infty$ one finds $I \sim \sqrt{2E}$, or $E \sim I^2/2$.*

Example 3.12. *The case of the (attracting) Kepler problem goes as follows. The Hamiltonian of the system in spherical polar coordinates is*

$$H = \frac{p_r^2}{2m} + \frac{1}{2mr^2}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) - \frac{k}{r} \ , \tag{3.31}$$

which is of the matryoshka type. The three first integrals in involution are p_φ *(observe that* φ *is ignorable:* $\partial H/\partial\varphi = 0$*),* $p_\theta^2 + p_\varphi^2/\sin^2\theta$ *and* H*. One finds by a direct computation that* $p_\varphi = L_z$*, the* z *component of the angular momentum* $L = x \wedge p$*, whereas* $p_\theta^2 + p_\varphi^2/\sin^2\theta = |L|^2$*. As a consequence, the three action variables of such a separable system can be computed by means of elementary (though long) integrations, yielding*

$$I_\varphi = \frac{1}{2\pi}\oint p_\varphi \mathrm{d}\varphi = L_z \ ; \tag{3.32}$$

$$I_\theta = \frac{1}{2\pi}\oint p_\theta \ \mathrm{d}\theta = \frac{1}{\pi}\int_{\pi/2-\theta_0}^{\pi/2+\theta_0} \sqrt{|L|^2 - \frac{L_z^2}{\sin^2\theta}} = |L| - L_z \ , \tag{3.33}$$

where $\sin^2(\pi/2 \pm \theta_0) = \cos^2(\theta_0) = L_z^2/|L|^2$*;*

$$I_r = \frac{1}{2\pi}\oint p_r \ \mathrm{d}r = \frac{\sqrt{2m}}{\pi}\int_{r_-}^{r_+} \sqrt{E + \frac{k}{r} - \frac{|L|^2}{\sin^\theta}} \ \mathrm{d}r = k\sqrt{\frac{m}{-2E}} - |L| \ , \tag{3.34}$$

where $r_\pm$ *are the zeroes of the integrand square root. The overall result, expressing the energy* E *as a function of the actions, is*

$$E(I) = -\frac{mk^2}{2(I_r + I_\theta + I_\varphi)^2} \ . \tag{3.35}$$

Exercise 3.1. *Prove (3.33). Hints. Make use a first change of variables introducing the angle* ψ *such that* $\cos\theta = \sin\theta_0 \ \sin\psi$*. Make then use of a second change of variables introducing* $u = \tan\psi$*.*

Exercise 3.2. *Find first, explicitly* $r_\pm(E)$*; call* $c := (r_+ + r_-)/(r_+ - r_-)$*. Then rewrite the integral in (3.34) as*

$$I_r = \frac{sqrt{-}2mE}{\pi}\int_{r_-}^{r_+} \sqrt{(r_+ - r)(r - r_-)} \ \mathrm{d}r \ .$$

Introduce the changes of variable

$$r = \frac{r_+ + r_-}{2} + \frac{r_+ - r_-}{2}\xi \ ,$$

then $\xi = \cos\phi$*, and finally* $\tau = \tan(\phi/2)$*. One thus finds* $I_r = \pi[c - \sqrt{c^2 - 1}]$ *which, by the explicit expression of* c*, leads to (3.34).*

The energy in (3.35) depends on the sum of three actions. One can thus transform to new angle-action variables (θ, J) *defined by* $J_1 = I_r + I_\theta + I_\varphi$, $J_2 = I_\theta + I_\varphi$, $J_3 = I_\varphi$, *the corresponding angles being obtained by canonically completing the transformation. Then*

$$E(J) = -\frac{mk^2}{2J_1^2} \quad ; \quad \dot{\theta}_1 = \frac{\partial E}{\partial J_1} = \frac{mk^2}{J_1^3} \ , \tag{3.36}$$

all the other canonical variables (the three actions and the other two angles θ_2 *and* θ_2*) being constant. One finds in this way that the Kepler motions corresponding to negative energy values are all periodic in time, i.e. the torus predicted by the Arnol'd theorem in this case has dimension one.*

Exercise 3.3. *Find the Kepler law* $T^2/a^3 = const.$, *where* T *is the period of the orbit and* $a = (r_+ + r_-)/2$ *its linear dimension (the semi-major axis of the ellipse). Hint:* $T = 2\pi/\omega = 2\pi/\dot{\theta}_1$; *express* a *as a function of* J_1.