2

CANONICAL TRANSFORMATIONS

As for a generic system of differential equations, coordinate transformations can be used in order to bring the system to a simpler form. If the system is Hamiltonian it is desirable to keep the Hamiltonian form of the equations when the system is transformed. The search for a class of transformations satisfying the latter property leads to considering the group of the so called *canonical transformations*.

The condition of canonicity can be expressed in terms of Poisson brackets, Lagrange brackets and differential forms. In these notes I will discuss five different criteria of canonicity. Precisely, a coordinate transformation is canonical in case:

(i) the Jacobian matrix of the transformation is a symplectic matrix;
(ii) the transformation preserves the Poisson brackets;
(iii) the transformation preserves the Lagrange brackets;
(iv) the transformation preserves the differential 2–form $\sum_j dp_j \wedge dq_j$;
(v) the transformation preserves the integral over a closed curve of the 1–form $\sum_j p_j dq_j$.

An useful method for constructing canonicals transformations is furnished by the theory of *generating functions*. This is the basis for further development of the integration methods, eventually leading to the Hamilton–Jacobi's equation.

2.1 Elements of symplectic geometry

Before entering the discussion of canonical transformations we recall a few aspects of symplectic geometry that will be useful. A relevant role is played by the bilinear antisymmetric form induced by the matrix J introduced in sect. 1.1.3, formula (1.16).

Symplectic geometry is characterized by the skew symmetric matrix J in the same sense as Euclidean geometry is characterized by the identity matrix I. Indeed, Euclidean geometry is characterized by transformations preserving the bilinear symmetric form

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_j x_j y_j \ . \tag{2.1}$$

This is called the *inner product*, and is preserved by the group of orthogonal transformations, namely, matrices U satisfying

$$\mathsf{U}^\top \mathsf{U} = \mathsf{I} \ . \tag{2.2}$$

Similarly, symplectic geometry is characterized by transformations preserving the bilinear form

$$[\mathbf{x}, \mathbf{y}] = \langle \mathbf{x}, \mathsf{J}\mathbf{y} \rangle \ , \tag{2.3}$$

where J is the matrix defined by (1.16). This is called the *symplectic product*. It is antisymmetric, i.e., $[\mathbf{x}, \mathbf{y}] = -[\mathbf{y}, \mathbf{x}]$, and non degenerate, i.e., $[\mathbf{x}, \mathbf{y}] = 0$ for all $\mathbf{y} \in \mathbb{R}^{2n}$ implies $\mathbf{x} = 0$.

2.1.1 The symplectic group

Let us consider a linear mapping in $\mathbb{R}^{2n}$

$$x = \mathsf{U} y \ , \tag{2.4}$$

where U is a $2n \times 2n$ nonsingular matrix. The matrix U is said to be *symplectic* if

$$\mathsf{U}^\top \mathsf{J} \mathsf{U} = \mathsf{J} \ . \tag{2.5}$$

This condition is actually equivalent to

$$\mathsf{U} \mathsf{J} \mathsf{U}^\top = \mathsf{J} \ . \tag{2.6}$$

For, by (2.5) we have $\mathsf{U}^\top = \mathsf{J}\mathsf{U}^{-1}\mathsf{J}^{-1}$, and so also $\mathsf{U}\mathsf{J}\mathsf{U}^\top = \mathsf{U}\mathsf{J}^2\mathsf{U}^{-1}\mathsf{J}^{-1} = -\mathsf{J}^{-1} = \mathsf{J}$ in view of $\mathsf{J}^2 = -\mathsf{I}$. By the way, this shows that if U is symplectic then $\mathsf{U}^\top$ is symplectic, too.

The set of symplectic matrices forms a group with respect to matrix multiplication. For, the identity matrix I is clearly symplectic; if U and V are symplectic then $(\mathsf{U}\mathsf{V})^\top \mathsf{J} (\mathsf{U}\mathsf{V}) = \mathsf{V}^\top \mathsf{U}^\top \mathsf{J} \mathsf{U} \mathsf{V} = \mathsf{V}^\top \mathsf{J} \mathsf{V} = \mathsf{J}$, so that $(\mathsf{U}\mathsf{V})$ is symplectic; if U is symplectic then $\mathsf{J} = \mathsf{I}\mathsf{J}\mathsf{I} = (\mathsf{U}^{-1})^\top \mathsf{U}^\top \mathsf{J} \mathsf{U} \mathsf{U}^{-1} = (\mathsf{U}^{-1})^\top \mathsf{J} \mathsf{U}^{-1}$, so that U^{-1} is symplectic. Finally, if U is symplectic so is $\mathsf{U}^\top$.

Going back to the canonically conjugated coordinates (q, p) it is immediately seen that the symplectic product of two $2n$ vectors $\mathbf{z} = (\mathbf{q}, \mathbf{p})$ and $\mathbf{z}' = (\mathbf{q}', \mathbf{p}')$ is given by the expression $\sum_{j=1}^{n} (q_j p'_j - q'_j p_j)$. That is, the symplectic product is obtained by projecting the parallelogram with sides $\mathbf{z}, \mathbf{z}'$ onto each of the planes q_j, p_j, and then adding up algebraically the oriented areas of all these projections.

The set of linear transformations which preserve the symplectic product is characterized as the group of symplectic matrices. For, $[\mathsf{U}\mathbf{x}, \mathsf{U}\mathbf{y}] = \langle \mathsf{U}\mathbf{x}, \mathsf{J}\mathsf{U}\mathbf{y} \rangle = \langle \mathbf{x}, \mathsf{U}^\top \mathsf{J} \mathsf{U} \mathbf{y} \rangle$, and this coincides with $[\mathbf{x}, \mathbf{y}]$ if and only if U satisfies (2.5).

2.1.2 Symplectic spaces and symplectic–orthogonality

A *symplectic space* is a real linear vector space V equipped with a bilinear antisymmetric nondegenerate form $[\cdot, \cdot]$. We shall consider here only vector spaces of finite dimension. Two vectors $\mathbf{x}, \mathbf{y}$ are said to be symplectic–orthogonal in case $[\mathbf{x}, \mathbf{y}] = 0$. We shall write $\mathbf{x} \angle\, \mathbf{y}$. To any subspace W of V we associate the set

$$W^{\angle} = \{\mathbf{x} \in V \mid \mathbf{x} \angle\, \mathbf{y} \text{ for all } \mathbf{y} \in W\} \ ; \tag{2.7}$$

by the bilinearity of the symplectic form $W^{\angle}$ is a subspace, which is said to be *symplectic–orthogonal*[1] to W. We emphasize that the concept of symplectic–orthogonality presents sharp differences with respect to the concept of orthogonality in Euclidean geometry. Here are three main differences.

(i) Every vector is self symplectic–orthogonal, since the symplectic product is antisymmetric; therefore, any one–dimensional subspace is self symplectic–orthogonal, too.

(ii) The restriction of the symplectic product to a subspace is still a bilinear form, but in general it fails to be nondegenerate. For instance, the restriction to a one–dimensional subspace is clearly degenerate.

(iii) The subspaces W and $W^{\angle}$ need not be complementary.

Two basic properties that are common to both geometries are given by the following

Lemma 2.1: *Let W be a subspace of a symplectic space V. Then*

$$\dim W + \dim W^{\angle} = \dim V \ . \tag{2.8}$$

and

$$\left(W^{\angle}\right)^{\angle} = W \tag{2.9}$$

Proof. If $\dim W = 0$ or $\dim W = \dim V$, then the statement is trivial. So, let us suppose that $\dim W = m$ with $0 < m < \dim V$. Denoting $n = \dim V$, let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis of V, with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ a basis of W. Writing a generic vector $\mathbf{v} \in V$ as $\mathbf{v} = \sum_{j=1}^{n} v_j \mathbf{u}_j$, the symplectic product takes the form $[\mathbf{v}, \mathbf{w}] = \sum_{j,k} a_{jk} v_j w_k$, where $a_{jk} = [\mathbf{u}_j, \mathbf{u}_k]$ is an element of an antisymmetric nondegenerate matrix. If $\mathbf{w} \in W$, then $w_{n-m+1} = \ldots = w_n = 0$, because by hypothesis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is a basis of W. If moreover $\mathbf{v} \in W^{\angle}$, then the relation of symplectic orthogonality is

$$\sum_{j=1}^{m} \beta_j w_j = 0 \ , \quad \beta_j = \sum_{k=1}^{n} a_{jk} v_k \ .$$

Since $\mathbf{w} \in W$ is arbitrary, the first equation implies $\beta_1 = \ldots = \beta_m = 0$ and $\beta_{m+1}, \ldots, \beta_n$ arbitrary; so, there are $n - m$ independent solutions. Since the matrix $\{a_{jk}\}$ is nondegenerate, the second relation above guarantees the existence of exactly $n - m$ independent vectors symplectic–orthogonal to the subspace W, so that $\dim W^{\angle} = n - m$, as claimed.

Coming to (2.9), if $\mathbf{w} \in W$, then $\mathbf{v} \angle\, \mathbf{w}$ for all $\mathbf{v} \in W^{\angle}$, because by definition every element of $W^{\angle}$ is symplectic–orthogonal to every element of W, and so also to $\mathbf{w}$, so

[1] Some authors use the name *left–orthogonality*. See for instance [7].

that $W \subset (W^\angle)^\angle$. On the other hand, by (2.8), $\dim(W^\angle)^\angle = \dim W$, so that (2.9) follows. *Q.E.D.*

We now go deeper into the concept of symplectic–orthogonality, pointing out some properties that do not appear in Euclidean geometry. To this end, we first need some definitions.

A subspace W of a symplectic space V is said to be:

(i) *isotropic* in case $W \subset W^\angle$;

(ii) *coisotropic* in case $W \supset W^\angle$, namely, if its symplectic–orthogonal subspace is isotropic;

(iii) *Lagrangian* in case $W = W^\angle$, namely, if it is both isotropic and coisotropic;

(iv) *symplectic* in case the symplectic product restricted to W is still nondegenerate.

These definitions are easily illustrated by considering the space $\mathbb{R}^{2n}$. We denote by $\{\mathbf{e}_1,\ldots,\mathbf{e}_n,\mathbf{d}_1,\ldots,\mathbf{d}_n\}$ the canonical basis of $\mathbb{R}^{2n}$ and by $(q_1,\ldots,q_n,p_1,\ldots,p_n)$ the coordinates, so that a vector $\mathbf{x} \in \mathbb{R}^{2n}$ is represented as $q_1\mathbf{e}_1+\ldots+q_n\mathbf{e}_n+p_1\mathbf{d}_1+\ldots+p_n\mathbf{d}_n$. The symplectic bilinear form is defined by the relations

$$[\mathbf{e}_j,\mathbf{e}_k] = [\mathbf{d}_j,\mathbf{d}_k] = 0 \ , \quad [\mathbf{e}_j,\mathbf{d}_k] = \delta_{j,k} \ , \quad j,k = 1,\ldots,n \ .$$

We shall refer to the basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n,\mathbf{d}_1,\ldots,\mathbf{d}_n\}$ satisfying the latter relations as the *canonical symplectic basis.*

Example 2.1: *Arithmetic planes.* Let us call *arithmetic plane* the subspace spanned by any subset of the vectors $\{\mathbf{e}_1,\ldots,\mathbf{e}_n,\mathbf{d}_1,\ldots,\mathbf{d}_n\}$. Formally, given any subsets J,K of $\{1,\ldots,n\}$, we consider the plane spanned by the vectors $\{\mathbf{e}_j\}_{j\in J} \cup \{\mathbf{d}_k\}_{k\in K}$. The following examples are easily understood:

(i) the arithmetic plane spanned by any subset of the vectors $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ is an isotropic subspace;

(ii) the direct sum of any of the arithmetic planes of the point (i) with the arithmetic plane $\mathrm{span}(\mathbf{d}_1,\ldots,\mathbf{d}_n)$ is a coisotropic subspace;

(iii) the arithmetic plane $\mathrm{span}(\mathbf{e}_1,\ldots,\mathbf{e}_n)$ is a Lagrangian subspace.

(iv) for every $j \in \{1,\ldots,n\}$, the arithmetic 2–dimensional plane $\mathrm{span}(\mathbf{e}_j,\mathbf{d}_j)$ is symplectic; further symplectic subspaces are generated by direct sum of such arithmetic planes.

In the examples (i)–(iii) the role of the vectors $\mathbf{e}_j$ and $\mathbf{d}_j$ can be exchanged, of course.

A more interesting example is the following:

Example 2.2: *Lagrangian arithmetic planes.* Consider any partition of the indexes $\{1,\ldots,n\}$ into two disjoint subsets J and K (i.e., $J \cap K = \emptyset$ and $J \cup K = \{1,\ldots,n\}$). Then the arithmetic plane spanned by the n vectors $\{\mathbf{e}_j\}_{j\in J} \cup \{\mathbf{d}_k\}_{k\in K}$ is a Lagrangian plane. There are 2^n different Lagrangian planes that are generated that way.[2]

[2] There are 2^n different partitions of n objects into two disjoint subsets.

2.1.3 Canonical basis of a symplectic space

The examples above are in fact quite general. For, any symplectic space can be equipped with a canonical symplectic basis. This is stated by the following

Proposition 2.2: *Let V be a symplectic space. Then $\dim V$ is even, say $2n$, and there exists a canonical basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{d}_1, \ldots, \mathbf{d}_n\}$ satisfying*

$$[\mathbf{e}_j, \mathbf{e}_k] = [\mathbf{d}_j, \mathbf{d}_k] = 0 \ , \quad [\mathbf{e}_j, \mathbf{d}_k] = \delta_{jk} \ , \quad j, k = 1, \ldots, n \ . \tag{2.10}$$

The proof depends on some properties that are of independent interest, and are isolated in the following two lemmas.

Lemma 2.3: *A subspace W of V is symplectic if and only if the subspaces W and $W^{\angle}$ are complementary. In that case $W^{\angle}$ is a symplectic subspace.*

Proof. Let W be a symplectic subspace. Then $W \cap W^{\angle} = \{0\}$. For, we have $[\mathbf{v}, \mathbf{w}] = 0$ for all $\mathbf{w} \in W$ in view of $\mathbf{v} \in W^{\angle}$, and this implies $\mathbf{v} = 0$ in view of $\mathbf{v} \in W$ and of the nondegeneracy of the symplectic product. On the other hand, by (2.8), we have $\dim(W \oplus W^{\angle}) = \dim W + \dim W^{\angle} = \dim V$, so that W and $W^{\angle}$ are complementary. Conversely, let W and $W^{\angle}$ be complementary, and let $\mathbf{v} \in W$ be such that $[\mathbf{v}, \mathbf{w}] = 0$ for all $\mathbf{w} \in W$. By definition of $W^{\angle}$ this implies $[\mathbf{v}, \mathbf{w}] = 0$ for all $\mathbf{w} \in V$, and so also $\mathbf{v} = 0$ by nondegeneracy. We conclude that W is symplectic.
It remains to prove that $W^{\angle}$ is symplectic. Since W and $W^{\angle}$ are complementary, if $[\mathbf{v}, \mathbf{w}] = 0$ for all $\mathbf{w} \in W^{\angle}$ then $\mathbf{v} = 0$ by the same argument as above. *Q.E.D.*

Lemma 2.4: *Let V be a symplectic space. Then $\dim V \geq 2$, and there exists a decomposition of V in two complementary symplectic subspaces V_1 and V_2, with $\dim V_1 = 2$ and $V_2 = V_1^{\angle}$. Moreover, in V_1 there exists a symplectic canonical basis $\mathbf{e}_1, \mathbf{d}_1$.*

Proof. Let $\mathbf{e}_1 \neq 0$ be an arbitrary vector. Then, by nondegeneracy, there exists a vector $\mathbf{d}_1$ independent of $\mathbf{e}_1$ such that $[\mathbf{e}_1, \mathbf{d}_1] \neq 0$. This proves that the dimension must be at least 2. In view of linearity, by a trivial rescaling we can choose $\mathbf{d}_1$ such that $[\mathbf{e}_1, \mathbf{d}_1] = 1$.
Let $V_1 = \mathrm{span}(\mathbf{e}_1, \mathbf{d}_1)$, so that $\dim V_1 = 2$ and $\mathbf{e}_1, \mathbf{d}_1$ is a basis of V_1. We prove that it is a symplectic subspace. To this end, first check that the decomposition of any vector $\mathbf{w} \in V_1$ over the basis $\mathbf{e}_1, \mathbf{d}_1$ is $\mathbf{w} = [\mathbf{w}, \mathbf{d}_1]\mathbf{e}_1 - [\mathbf{w}, \mathbf{e}_1]\mathbf{d}_1$; this is elementary. Suppose now that $[\mathbf{w}, \mathbf{v}] = 0$ for all $\mathbf{v} \in V_1$. Then we have in particular $[\mathbf{w}, \mathbf{e}_1] = [\mathbf{w}, \mathbf{d}_1] = 0$, and so also $\mathbf{w} = 0$ in view of the decomposition above. This proves that the restriction of the symplectic form to the subspace V_1 is nondegenerate. Therefore, V_1 is a symplectic subspace, and $\mathbf{e}_1, \mathbf{d}_1$ is a canonical basis of it.
By lemma 2.3, $V_2 = V_1^{\angle}$ is a symplectic subspace complementary to V_1. *Q.E.D.*

Proof of proposition 2.2. If $\dim V = 2$ just apply once lemma 2.4. If $\dim V > 2$, then apply recursively lemma 2.4. With an obvious change of notation, first write $V = V_1 \oplus V'$, where V_1 admits a symplectic canonical basis $\mathbf{e}_1, \mathbf{d}_1$, and V' is symplectic, with $\dim V' = \dim V - 2$. Next, apply again the lemma to V', getting $V' = V_2 \oplus V''$ with V_2 admitting a canonical basis $\mathbf{e}_2, \mathbf{d}_2$. Proceeding the same way, end up with a decomposition $V = V_1 \oplus \cdots \oplus V_n$ (recall that we assume that V is a vector space of finite dimension), where V_j is a 2–dimensional symplectic subspace equipped with a canonical basis $\mathbf{e}_j, \mathbf{d}_j$ satisfying $[\mathbf{e}_j, \mathbf{e}_j] = [\mathbf{d}_j, \mathbf{d}_j] = 0$ and $[\mathbf{e}_j, \mathbf{d}_j] = 1$ $(j = 1, \ldots, n)$. This implies that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{d}_1, \ldots, \mathbf{d}_n\}$ is a basis for V, and so also $\dim V = 2n$. By construction, all subspaces $V_1, \ldots, V_n$ are pairwise symplectic–orthogonal, which implies $[\mathbf{e}_j, \mathbf{e}_k] = [\mathbf{d}_j, \mathbf{d}_k] = [\mathbf{e}_j, \mathbf{d}_k] = 0$ for $j \neq k$. This proves (2.10). *Q.E.D.*

2.1.4 Properties of the subspaces of a symplectic space

We establish two properties that will be relevant in the discussion concerning integrable systems.

Lemma 2.5: *If W is isotropic, then $\dim W \leq n$; if W is coisotropic, then $\dim W \geq n$; if W is Lagrangian, then $\dim W = n$.*

An immediate consequence of this lemma is the

Corollary 2.6: *An isotropic subspace W is Lagrangian if and only if $\dim W = n$. The same holds true for a coisotropic subspace.*

Proof of lemma 2.5. Just use lemma 2.1, formula (2.8). If W is isotropic, i.e., $W \subset W^{\angle}$, then $\dim W \leq \dim W^{\angle}$, which implies $\dim W \leq n$. If W is coisotropic, just reverse the argument. If W is Lagrangian, it is both isotropic and coisotropic, so that both the inequalities $\dim W \leq n$ and $\dim W \geq n$ apply. *Q.E.D.*

Lemma 2.7: *Let V be a symplectic space and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{d}_1, \ldots, \mathbf{d}_n\}$ be a canonical basis. Then any Lagrangian subspace W of V is complementary to at least one of the Lagrangian arithmetic planes of example 2.2.*

Proof. Consider the n–dimensional Lagrangian plane $D = \mathrm{span}\{\mathbf{d}_1, \ldots, \mathbf{d}_n\}$, and let $P = D \cap W$ and $m = \dim P$. Since P is a subspace of D, we have $0 \leq m \leq n$, and there exist $n - m$ vectors in $\{\mathbf{d}_1, \ldots, \mathbf{d}_n\}$ spanning a $(n - m)$–dimensional arithmetic plane complementary to P in D. That is: there is a subset $K \subset \{1, \ldots, n\}$, with $\#K = n - m$, such that the arithmetic plane $D_K = \mathrm{span}\{\mathbf{d}_k\}_{k \in K}$ satisfies $D_K \cap P = \{0\}$ and $D_K \oplus P = D$ (the set K needs not be unique). Let now $J = \{1, \ldots, n\} \setminus K$, so that $\{J, K\}$ is a partition of $\{1, \ldots, n\}$, and let $E_J = \mathrm{span}\{\mathbf{e}_j\}_{j \in J}$. We prove that the Lagrangian subspace $L = E_J \oplus D_K$ is complementary to W, namely, it is the Lagrangian arithmetic plane we are looking for. Since $\dim W = \dim L = n$, it is enough to prove that $W \cap L = \{0\}$. This is seen as follows. On the one hand, by $P \subset W$ and $W \angle W$ we have $P \angle W$; on the other hand, by $D_K \subset L$ and $L \angle L$ we have $D_K \angle L$. Using these relations we get[3] $D = P \oplus D_K \angle W \cap L$, and so $W \cap L \subset D$ in view of D being Lagrangian. Therefore, $W \cap L = (W \cap D) \cap (L \cap D) = P \cap D_K = \{0\}$, by construction of D_K. *Q.E.D.*

---

[3] Let $\mathbf{v} \in W \cap L$. Since $\mathbf{v} \in W$, we have $\mathbf{v} \angle P$; since $\mathbf{v} \in L$, we have $\mathbf{v} \angle D_K$. We conclude $\mathbf{v} \angle P \oplus D_K$.

2.2 Transformations preserving the Hamiltonian form of the equations.

Let us now turn to the main argument of this chapter, namely to characterize a class of transformation that allows us to stay in the framework of the Hamiltonian formalism.

It may be useful to make again a connection with the lagrangian formalism. It is well known that Lagrange equations have the nice property of being invariant with respect to point transformation (i.e., changes of the coordinates in configuration space, which by differentiation generate the corresponding transformations on the generalized velocities). The Hamiltonian formalism removes the tie between generalized coordinates and velocities, so that arbitrary transformations involving all the canonical coordinates may be devised. However, an arbitrary transformation will likely produce equations which are not in Hamiltonian form, in the sense that the second members are not expressed as derivatives of a unique function.

The problem then is to characterize a restricted class of transformation which keep the form of Hamilton's equations.

2.2.1 Conditions for canonicity

Let us first look for a class of transformations $(q,p) = \mathcal{C}(\overline{q},\overline{p})$ satisfying the following

Condition 1: *to every Hamiltonian function $H(q,p)$ one can associate another function $K(\overline{q},\overline{p})$ such that the canonical system of equations*

$$\dot{q}_j = \frac{\partial H}{\partial p_j}\ , \quad \dot{p}_j = -\frac{\partial H}{\partial q_j}\ , \quad j = 1,\ldots,n\ ,$$

is changed into the system

$$\dot{\overline{q}}_j = \frac{\partial K}{\partial \overline{p}_j}\ , \quad \dot{\overline{p}}_j = -\frac{\partial K}{\partial \overline{q}_j}\ , \quad j = 1,\ldots,n\ ,$$

which is still canonical.

We shall say that such a transformation preserves the canonical form of the equations.

Condition 1 applies both to the case of time–independent and time–dependent transformations

Condition 2: *the transformation preserves the canonical form of the equations with the new Hamiltonian*

$$\overline{H}(\overline{q},\overline{p}) = H(q,p)\big|_{q=q(\overline{q},\overline{p})\,,\,p=p(\overline{q},\overline{p})}\ .$$

The difference with respect to the first condition is that the new Hamiltonian is constructed by a straightforward substitution of the transformation in the old one. Transformations satisfying condition 2 condition will be called *time–independent canonical transformations*, or simply canonical.[4]

---

[4] There is no general agreement about the use of the term *canonical transformation*. Some authors call canonical any transformation satisfying the requisite of preserving the canonical form of the equations, as stated by condition 1 in the text. For example,

The case of time–dependent transformations appears to be a little more complex: we need to go back to condition 1, since the new Hamiltonian is not determined by a mere substitution of variables. I will discuss this matter at the end of the chapter, where it is shown how the case of time–dependent transformations can be reduced to the time–independent one, and how the new Hamiltonian $K(\overline{q}, \overline{p})$ is constructed.

The examples below show that there are transformations satisfying the conditions above. As already anticipated, I will consider here only time–independent transformations.

Example 2.3: *Translation.* The transformation

$$q_j = \overline{q}_j + a_j \ , \quad p_j = \overline{p}_j + b_j \ , \quad 1 \le j \le n \ ,$$

where $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are constants, preserves the canonical form of the equations with the new Hamiltonian

$$\overline{H}(\overline{q}, \overline{p}) = H(q, p)\Big|_{q=\overline{q}+a\,,\,p=\overline{p}+b} \ .$$

Thus, the transformation is canonical.

Example 2.4: *Scaling transformation.* The transformation

$$q_j = \alpha\overline{q}_j \ , \quad p_j = \beta\overline{p}_j \ , \quad 1 \le j \le n \ ,$$

with real constants α and β preserves the canonical form of the equations with the new Hamiltonian

$$K(\overline{q}, \overline{p}) = \frac{1}{\alpha\beta} H(q, p)\Big|_{q=\alpha\overline{q}\,,\,p=\beta\overline{p}} \ .$$

This transformation always satisfies condition 1 above. However, condition 2 is fulfilled only in case $\alpha\beta = 1$, and so we shall call it canonical only in the latter case.[5]

Example 2.5: *Exchange of conjugated coordinates.* The transformation

$$q_j = \overline{p}_j \ , \quad p_j = -\overline{q}_j \ , \quad 1 \le j \le n \ ,$$

preserves the canonical form of the equations with the new Hamiltonian

$$\overline{H}(\overline{q}, \overline{p}) = H(q, p)|_{q=\overline{p}\,,\,p=-\overline{q}} \ .$$

This is a canonical transformation.

---

this is the attitude of Wintner in [81]. Others follow the attitude of the present notes. Example 2.4 below illustrates the difference.

[5] This example shows that restricting the use of the adjective *canonical* only to transformations satisfying the second condition essentially reduces to excluding the class of scaling transformations for which $\alpha\beta \ne 1$. Including all such transformations makes the definition more general, of course. However, this introduces some complications in the expositions that are unnecessary, and this is what I want to avoid. The more general framework is recovered by keeping in mind that a canonical transformation in the sense intended here can always be composed with a scaling transformation, which sometimes is an useful device.

2.2.2 Symplecticity of the Jacobian matrix

In order to find a condition for a transformation to be canonical let us use the compact notation of sect. 1.1.3. Let us write the transformation as $x = u(y)$, and let $K(y) = H \circ u(y)$ be the Hamiltonian expressed in the new coordinates y.

Proposition 2.8: *The transformation $x = u(y)$ is canonical if and only if the Jacobian matrix u_y of the transformation satisfies*

$$u_y^{\top} \mathsf{J} u_y = \mathsf{J} \ , \tag{2.11}$$

Here, J is the matrix defined by (1.16). In the language of sect. 2.2 we say that the Jacobian matrix of the transformation must be symplectic.[6]

Proof. Using coordinates, compute

$$\frac{\partial K}{\partial y_k} = \sum_{j=1}^{2n} \frac{\partial x_j}{\partial y_k} \frac{\partial H}{\partial x_j}\bigg|_{x=u(y)} \ ,$$

or, in compact notation,

$$\partial_y K = u_y^{\top} \partial_x H \circ u \ .$$

Using $\mathsf{J}^2 = -\mathsf{I}$, write the canonical equations for H as $-\mathsf{J}\dot{x} = \partial_x H$, and using also $\dot{x} = u_y \dot{y}$ compute $-\mathsf{J} u_y \dot{y} = \partial_x H \circ u$. Finally, multiply both sides of the latter relation by $u_y^{\top}$, and get $-u_y^{\top} \mathsf{J} u_y \dot{y} = u_y^{\top} \partial_x H \circ u = \partial_y K$. Therefore, the equations for y are written

$$-u_y^{\top} \mathsf{J} u_y \dot{y} = \partial_y K \ .$$

They are in canonical form with the Hamiltonian $K(y)$ provided $u_y^{\top} \mathsf{J} u_y = \mathsf{J}$. *Q.E.D.*

2.2.3 Preservation of Poisson brackets

As we remarked in sect 1.2.2, the Hamiltonian formalism can be expressed in terms of Poisson brackets, saying that *the time evolution of any dynamical variable f is given by equation $\dot{f} = \{f, H\}$* . This leads to a characterization of canonical transformations as possessing the property of leaving invariant the form of the Poisson bracket.

Let $(q, p) = \mathscr{C}(\overline{q}, \overline{p})$ be a coordinate transformation, and denote by $\mathscr{C} f$ the transformed function

$$\bigl(\mathscr{C} f\bigr)(\overline{q}, \overline{p}) = f(q, p)\Big|_{(q,p)=\mathscr{C}(\overline{q},\overline{p})} \ .$$

Also, denote by $\{\cdot, \cdot\}_{q,p}$ and by $\{\cdot, \cdot\}_{\overline{q},\overline{p}}$ the Poisson bracket with respect to the conjugate variables q, p and $\overline{q}, \overline{p}$, respectively. Consider now the class of transformations

[6] In the recent literature the name “canonical transformation” is frequently replaced by “symplectic transformation”. In these notes I prefer to use the old fashioned name ”canonical”. However, I will sometimes use the adjective “symplectic” when dealing with linear transformations, which involve symplectic matrices.

satisfying the condition that the following diagram is commutative for any functions f and g:

$$
\begin{tikzcd}
f, g \arrow[r, "\mathscr{C}"] \arrow[d, "\{\cdot,\cdot\}"'] & \mathscr{C}f, \mathscr{C}g \arrow[d, "\{\cdot,\cdot\}"] \\
\{f,g\}_{q,p} \arrow[r, "\mathscr{C}"'] & \mathscr{C}\bigl(\{f,g\}_{q,p}\bigr) = \{\mathscr{C}f, \mathscr{C}g\}_{\overline{q},\overline{p}}\ .
\end{tikzcd}
\tag{2.12}
$$

In words, one obtains the same result both (a) by computing the Poisson bracket with respect to the variables q, p and then changing the variables in the result, or (b) by changing the variables and then computing the Poisson bracket with respect to the new variables $\overline{q}, \overline{p}$. If this happens to be true, we shall say that the transformations *preserves the Poisson brackets*.

Proposition 2.9: *A transformation $(q, p) = \mathscr{C}(\overline{q}, \overline{p})$ is canonical if and only if it preserves the Poisson brackets, i.e., the diagram (2.12) is commutative.*

The difficult part of the proof is the "only if", i.e., that the condition is necessary. In order to see it, we need to investigate to which extent we can transform the criterion expressed by the latter proposition to a practically applicable criterion.

Let us consider the coordinates q, p as functions on the phase space; it is an easy matter to check that the relations

$$
\begin{aligned}
&\{q_j, q_k\} = \{p_j, p_k\} = 0 \\
&\{q_j, p_k\} = \delta_{jk}
\end{aligned}
\qquad 1 \le j \le n\ ,\ 1 \le k \le n
\tag{2.13}
$$

hold true, where δ_{jk} is the Kronecker symbol. These expressions are sometimes called the *fundamental Poisson brackets*.

We prove the following

Lemma 2.10: *A transformation preserves the Poisson bracket between any two functions if and only if it preserves the fundamental Poisson brackets.*

As a direct consequence, proposition 2.9 can be reformulated in a more useful manner as

Corollary 2.11: *A transformation $(q, p) = \mathscr{C}(\overline{q}, \overline{p})$, is canonical if and only if it preserves the fundamental Poisson brackets, i.e.,*

$$
\begin{aligned}
&\{q_j, q_k\}_{\overline{q},\overline{p}} = \{p_j, p_k\}_{\overline{q},\overline{p}} = 0 \\
&\{q_j, p_k\}_{\overline{q},\overline{p}} = \delta_{jk}\ , \qquad 1 \le j \le n\ , \quad 1 \le k \le n\ .
\end{aligned}
\tag{2.14}
$$

As a matter of fact, a direct proof of this statement is achieved by just writing explicitly the condition of proposition 2.8 in terms of the old canonical coordinates q, p and of the new ones $\overline{q}, \overline{p}$. The reader will see that they are just different formulations of the same thing. However, this arguments brings no light on the apparently stronger condition that the Poisson bracket between *any* two function must be preserved. Since the formulation of the Hamiltonian formalism in terms of Poisson brackets is of interest in itself, let us proceed by giving a complete proof, independent of that of proposition 2.8.

Proof of lemma 2.10. Assume that the Poisson bracket between any two functions is preserved; then the fundamental Poisson brackets are preserved, too. So, we must prove only the converse. To this end, first check that if we are given a function $f(\varphi_1, \ldots, \varphi_r)$, where, in turn, $\varphi_1 \ldots, \varphi_r$ are functions of the canonical variables q, p, then

$$\{f, g\} = \sum_{l=1}^{r} \frac{\partial f}{\partial \varphi_l} \{\varphi_l, g\} ,$$

where $g(q, p)$ is any function. This is just matter of straightforward calculations. Put now $r = 2n$ and $\varphi_1(\overline{q}, \overline{p}) = q_1(\overline{q}, \overline{p}), \ldots, \varphi_{2n}(\overline{q}, \overline{p}) = p_n(\overline{q}, \overline{p})$, and consider also $g(q, p)$ as function of the new variables $\overline{q}, \overline{p}$ through $\varphi_1, \ldots, \varphi_{2n}$. Then, using the identity above, compute

$$\begin{aligned} \{f, g\}_{\overline{q},\overline{p}} = \sum_{j,k} \biggl(&\frac{\partial f}{\partial q_j} \frac{\partial g}{\partial q_k} \{q_j, q_k\}_{\overline{q},\overline{p}} + \frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_k} \{q_j, p_k\}_{\overline{q},\overline{p}} \\ &+ \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_k} \{p_j, q_k\}_{\overline{q},\overline{p}} + \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial p_k} \{p_j, p_k\}_{\overline{q},\overline{p}} \biggr) . \end{aligned}$$

In view of the preservation of the fundamental Poisson bracket we immediately get $\{f, g\}_{\overline{q},\overline{p}} = \{f, g\}_{q,p}$, namely that the Poisson bracket between f and g is preserved. *Q.E.D.*

Proof of proposition 2.9. Let the transformation to preserve the Poisson brackets. Denoting $\overline{q} = \overline{q}(q, p)$, $\overline{p} = \overline{p} = (q, p)$ the inverse transformation, let $f(q, p)$ be any of the new coordinates, e.g., $f(q, p) = \overline{q}_j(q, p)$ for some j. Then $\dot{f} = \{f, H\}_{q,p}$. On the other hand, by preservation of the Poisson brackets we also have, after changing the variables, $\dot{f} = \{f, H\}_{\overline{q},\overline{p}}$, that is, $\dot{\overline{q}}_j = \frac{\partial H}{\partial \overline{p}_j}$ where $H(\overline{q}, \overline{p}) = H(q, p)\big|_{(q,p)=\mathscr{C}(\overline{q},\overline{p})}$. Therefore, the transformed equations keep canonical form by just transforming the hamiltonian, as required by condition 2. Conversely, let the transformation be canonical, and let, e.g., $f = q_j$ and $H = p_k$ for some j and k. Thus, $\dot{f} = \{q_j, p_k\} = \delta_{j,k}$. On the other hand, after transforming to new variables we have $\dot{f} = \{f, H\}_{\overline{q},\overline{p}}$, because in the new variables the equations are still in canonical form. Since the time derivative of f must be the same after the transformation, we conclude $\{q_j, p_k\}_{\overline{q},\overline{p}} = \delta_{j,k}$. The argument applies to any pair of canonical coordinates q, p, and this means that the fundamental Poisson bracket are preserved. By lemma 2.10 this implies that the Poisson brackets are preserved. *Q.E.D.*

Example 2.6: *The case of one degree of freedom.* In the case $n = 1$ the canonicity condition in order the transformation to be canonical.may be written as

$$\{q, p\} = \det \begin{pmatrix} \dfrac{\partial q}{\partial \overline{q}} & \dfrac{\partial q}{\partial \overline{p}} \\ \dfrac{\partial p}{\partial \overline{q}} & \dfrac{\partial p}{\partial \overline{p}} \end{pmatrix} = 1 ,$$

which means that the transformation must be area–preserving. For instance, the scaling transformation

$$q = \alpha \overline{q} \ , \quad p = \frac{1}{\alpha}\overline{p} \tag{2.15}$$

is canonical. An example of a transformation which is common in geometry but is not canonical is the transformation to polar coordinates (in the phase plane), which is not area preserving. A similar transformation which however is canonical is

$$q = \sqrt{2I}\cos\varphi \ , \quad p = \sqrt{2I}\sin\varphi \ . \tag{2.16}$$

The variables I, φ thus defined are called *action–angle variables for the harmonic oscillator.*

Example 2.7: *Harmonic oscillators.* The Hamiltonian of a system of harmonic oscillators is

$$H(q,p) = \sum_{j=1}^{n} \frac{1}{2}\left(p_j^2 + \omega_j^2 q_j^2\right) \ ,$$

with $\omega = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n$. A more symmetric Hamiltonian is constructed by applying the rescaling transformation

$$q_j = \frac{\overline{q}_j}{\sqrt{\omega_j}} \ , \quad p_j = \overline{p}_j\sqrt{\omega_j} \ , \quad 1 \le j \le n \ ,$$

which is clearly canonical. The transformed Hamiltonian is

$$\overline{H}(\overline{q},\overline{p}) = \sum_j \frac{\omega_j}{2}\left(\overline{p}_j^2 + \overline{q}_j^2\right) \ .$$

The form of the Hamiltonian can be further simplified by using the transformation to action–angle variables, namely

$$\overline{q}_j = \sqrt{2I_j}\cos\varphi_j \ , \quad \overline{p}_j = \sqrt{2I_j}\sin\varphi_j \ , \quad 1 \le j \le n \ .$$

By this, the Hamiltonian is transformed to

$$H(I,\varphi) = \sum_j \omega_j I_j \ ,$$

which is trivially integrated.

2.2.4 Preservation of Lagrange brackets

Suppose that the canonical coordinates q, p are given as functions of two variables u, v. The Lagrange bracket $[u, v]$ is defined as[7]

$$[u,v] = \sum_{j=1}^{n}\left(\frac{\partial q_j}{\partial u}\frac{\partial p_j}{\partial v} - \frac{\partial q_j}{\partial v}\frac{\partial p_j}{\partial u}\right) \ . \tag{2.17}$$

[7] Lagrange, *Mém. de l'Institut de France* (1808).

There is a strict relation between Lagrange brackets and Poisson brackets. Let the canonical coordinates q, p be expressed as differentiable and invertible functions of $2n$ independent variables $u_1, \ldots, u_{2n}$ as

$$q_j = q_j(u_1, \ldots, u_{2n}) \ , \quad p_j = p_j(u_1, \ldots, u_{2n}) \ , \quad j,k = 1, \ldots, n \ , \tag{2.18}$$

so that the variables u can be expressed as differentiable functions of the canonical coordinates q, p, by inversion.

Lemma 2.12: *Let the $2n \times 2n$ matrices $\mathsf{A} = A_{j,k}$ and $\mathsf{B} = \{B_{j,k}\}$, $j,k = 1, \ldots, 2n$ be defined as $A_{jk} = \{u_j, u_k\}$ and $B_{jk} = [u_j, u_k]$. Then we have*

$$\mathsf{A}\mathsf{B}^\top = \mathsf{I} \ , \tag{2.19}$$

the identity matrix.

Proof. For $r, s = 1, \ldots, 2n$ compute

$$\begin{aligned}
&\sum_{l=1}^{2n} \{u_l, u_r\}[u_l, u_s] \\
&\qquad = \sum_{l=1}^{2n} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\frac{\partial u_l}{\partial q_j} \frac{\partial u_r}{\partial p_j} - \frac{\partial u_l}{\partial p_j} \frac{\partial u_r}{\partial q_j} \right) \left(\frac{\partial q_k}{\partial u_l} \frac{\partial p_k}{\partial u_s} - \frac{\partial q_k}{\partial u_s} \frac{\partial p_k}{\partial u_l} \right) \\
&\qquad = \sum_{j=1}^{n} \sum_{k=1}^{n} \Bigg[\left(\sum_{l=1}^{2n} \frac{\partial u_l}{\partial q_j} \frac{\partial q_k}{\partial u_l} \right) \frac{\partial u_r}{\partial p_j} \frac{\partial p_k}{\partial u_s} - \left(\sum_{l=1}^{2n} \frac{\partial u_l}{\partial q_j} \frac{\partial p_k}{\partial u_l} \right) \frac{\partial u_r}{\partial p_j} \frac{\partial q_k}{\partial u_s} \\
&\qquad\qquad - \left(\sum_{l=1}^{2n} \frac{\partial u_l}{\partial p_j} \frac{\partial q_k}{\partial u_l} \right) \frac{\partial u_r}{\partial q_j} \frac{\partial p_k}{\partial u_s} + \left(\sum_{l=1}^{2n} \frac{\partial u_l}{\partial p_j} \frac{\partial p_k}{\partial u_l} \right) \frac{\partial u_r}{\partial q_j} \frac{\partial q_k}{\partial u_s} \Bigg] \\
&\qquad = \sum_{j=1}^{n} \left(\frac{\partial u_r}{\partial p_j} \frac{\partial p_j}{\partial u_s} + \frac{\partial u_r}{\partial q_j} \frac{\partial q_j}{\partial u_s} \right) = \delta_{rs} \ ;
\end{aligned}$$

here, use has been made of the identities

$$\begin{aligned}
\sum_{l=1}^{2n} \frac{\partial u_l}{\partial q_j} \frac{\partial p_k}{\partial u_l} &= \sum_{l=1}^{2n} \frac{\partial u_l}{\partial p_j} \frac{\partial q_k}{\partial u_l} = 0 \ , \\
\sum_{l=1}^{2n} \frac{\partial u_l}{\partial q_j} \frac{\partial q_k}{\partial u_l} &= \sum_{l=1}^{2n} \frac{\partial u_l}{\partial p_j} \frac{\partial p_k}{\partial u_l} = \delta_{jk} \ .
\end{aligned}$$

Q.E.D.

It is now convenient to rename $(u_1, \ldots, u_{2n})$ as $(\overline{q}_1, \ldots, \overline{q}_n, \overline{p}_1, \ldots, \overline{p}_n)$. If the transformation (2.18) is the identity, then we have

$$\begin{aligned}
&[\overline{q}_j, \overline{q}_k] = [\overline{p}_j, \overline{p}_k] = 0 \\
&[\overline{q}_j, \overline{p}_k] = \delta_{jk} \ , \quad 1 \le j \le n \ , \ 1 \le k \le n \ .
\end{aligned} \tag{2.20}$$

These expressions are called the *fundamental Lagrange brackets.*

Proposition 2.13: *A transformation $(q,p) = \mathscr{C}(\overline{q},\overline{p})$ is canonical if and only if it preserves the fundamental Lagrange brackets.*

Proof. Recall that a canonical transformation preserves the fundamental Poisson brackets (corollary 2.11), and use lemma 2.12. If the fundamental Poisson brackets are preserved, then $\mathsf{A} = \mathsf{J}$, the symplectic matrix defined by (1.16); using $\mathsf{J}\mathsf{J}^\top = \mathsf{I}$, by (2.19) we conclude $\mathsf{B} = \mathsf{J}$. Conversely, if the fundamental Lagrange brackets are preserved, then $\mathsf{B} = \mathsf{J}$, and by (2.19) we conclude $\mathsf{A} = \mathsf{J}$. *Q.E.D.*

2.3 Poincaré's integral invariants

We consider the differential form

$$\omega^2 = \sum_{j=1}^{n} dq_j \wedge dp_j \ . \tag{2.21}$$

Proposition 2.14: *A transformation $(q,p) = \mathscr{C}(\overline{q},\overline{p})$ is canonical if and only if it preserves the 2–form $\omega^2 = \sum_j dq_j \wedge dp_j$.*

Proof. Using the formula for changing variables in a differential form, compute

$$\begin{aligned} \sum_j dq_j \wedge dp_j &= \sum_j \sum_{k,l} \left(\frac{\partial q_j}{\partial \overline{q}_k} d\overline{q}_k + \frac{\partial q_j}{\partial \overline{p}_k} d\overline{p}_k \right) \wedge \left(\frac{\partial p_j}{\partial \overline{q}_l} d\overline{q}_l + \frac{\partial p_j}{\partial \overline{p}_l} d\overline{p}_l \right) \\ &= \sum_{k<l} \left([\overline{q}_k, \overline{q}_l] d\overline{q}_k \wedge d\overline{q}_l + [\overline{p}_k, \overline{p}_l] d\overline{p}_k \wedge d\overline{p}_l \right) \\ &\qquad + \sum_{k,l} [\overline{q}_k, \overline{p}_l] d\overline{q}_k \wedge d\overline{p}_l \ , \end{aligned}$$

which shows that the coefficients of the transformed differential form are the Lagrange brackets. If the transformation is canonical, then by proposition 2.13 we get

$$\sum_j dp_j \wedge dq_j = \sum_j d\overline{p}_j \wedge d\overline{q}_j \ .$$

Conversely, if the latter identity is fulfilled, then the Lagrange brackets are preserved. By proposition 2.13 the claim follows. *Q.E.D.*

Starting form the 2–form ω^2 we can construct further differential forms of increasing order $\omega^4, \ldots, \omega^{2n}$. The following corollaries hold true, the proof of which is left to the reader.

Corollary 2.15: *A canonical transformation preserves all 2k–forms ω^{2k}, $k = 1, \ldots, n$.*

Corollary 2.16: *A canonical transformation preserves the phase space volume.*

The $2k$–forms ω^{2k} have been named by Poincaré *absolute invariant integrals*. The latter corollary is usually called Liouville's theorem.

Using Stokes theorem one proves

Proposition 2.17: *A transformation $(q,p) = \mathscr{C}(\overline{q},\overline{p})$ is canonical if and only if*

$$\int_\gamma \sum_j p_j dq_j = \int_{\mathscr{C}(\gamma)} \sum_j \overline{p}_j d\overline{q}_j \ , \tag{2.22}$$

where γ is a closed curve.

`*** Questo paragrafo deve essere ampliato in modo consistente in una prossima versione. ***`

2.4 Generating functions

The characterization of canonical transformations by propositions 2.14 and 2.17 offers us an explicit method for constructing canonical transformations.

Let us write the transformation as $q = q(\overline{q},\overline{p}), p = p(\overline{q},\overline{p})$, and let us assume that the relation $q = q(\overline{q},\overline{p})$ can be inverted (at least locally) with respect to $\overline{p}$, so that

$$\det \frac{\partial(q_1,\ldots,q_n)}{\partial(\overline{p}_1,\ldots,\overline{p}_n)} \neq 0 \ . \tag{2.23}$$

Transformations satisfying this condition are called *free canonical transformations*. By proposition 2.17 we have

$$\int_\gamma \sum_j p_j dq_j = \int_{\overline{\gamma}} \sum_j \overline{p}_j d\overline{q}_j \ ,$$

where γ is an arbitrary closed curve. This means that there exists a function $S(q,\overline{q})$ such that

$$\sum_j \bigl(p_j \mathrm{d}q_j - \overline{p}_j \mathrm{d}\overline{q}_j \bigr) = \mathrm{d}S \ , \tag{2.24}$$

The function $S(q,\overline{q})$ is called the *generating function*. We have indeed the following

Proposition 2.18: *Let the function $S(q,\overline{q})$ satisfy the condition*

$$\det \left(\frac{\partial^2 S}{\partial q_j \partial \overline{q}_k} \right) \neq 0 \ . \tag{2.25}$$

Then the transformation defined by

$$p_j = \frac{\partial S}{\partial q_j}(q,\overline{q}) \ , \quad \overline{p}_j = -\frac{\partial S}{\partial \overline{q}_j}(q,\overline{q}) \ , \quad 1 \leq j \leq n \tag{2.26}$$

is a free canonical transformation.

Proof. The transformation clearly satisfies the condition of proposition 2.17 for canonicity and condition (2.23) for being a free canonical transformation. *Q.E.D.*

It should be remarked that the transformation is given in implicit form; however, by condition (2.25), we are allowed to do the inversions needed in order to express either the old variables q,p as functions of the new ones $\overline{q},\overline{p}$ or the new variables as functions of the old ones. For, inverting the second of (2.26) with respect to q we get $q = q(\overline{q}, \overline{p})$, and replacing this in the first of (2.26) we get also $p = p(q, \overline{q})\big|_{q=q(\overline{q},\overline{p})}$, as required. Conversely, inverting the first of (2.26) with respect to $\overline{q}$ and replacing the result in the second of (2.26) we obtain the inverse transformation.

Example 2.8: *Exchange of conjugated coordinates.* The generating function

$$S(q, \overline{q}) = \sum_j q_j \overline{q}_j \tag{2.27}$$

generates the canonical transformation

$$p_j = \overline{q}_j \ , \quad \overline{p}_j = -q_j \ , \quad 1 \le j \le n \ , \tag{2.28}$$

exchanging the coordinates with the momenta.

The class of free canonical transformation does not exhaust all possibilities. For instance, one will immediately realize that the identity is not free, so that it can not be represented by a generating function of the form above. A different form of the canonical transformation can be constructed using the Legendre transformation.

`*** Aggiungere trasformata di Legendre ***`

Proposition 2.19: *Let the generating function $S(\overline{p}, q)$ satisfy the condition*

$$\det\left(\frac{\partial^2 S}{\partial \overline{p}_j \partial q_k}\right) \ne 0 \ . \tag{2.29}$$

Then the transformation implicitly defined by

$$p_j = \frac{\partial S}{\partial q_j}(\overline{p}, q) \ , \quad \overline{q}_j = \frac{\partial S}{\partial \overline{p}_j}(\overline{p}, q) \ , \quad 1 \le j \le n \ . \tag{2.30}$$

Proof. We check that the canonicity condition of proposition 2.17 is satisfied. For, introducing the Legendre transform of $S(\overline{p}, q)$, namely $\tilde{S} = S - \sum_j \overline{p}_j \overline{q}_j$, we compute

$$\begin{aligned} \mathrm{d}\tilde{S} &= \sum_j \left(\frac{\partial S}{\partial \overline{p}_j} \mathrm{d}\overline{p}_j + \frac{\partial S}{\partial q_j} \mathrm{d}q_j - \overline{p}_j \mathrm{d}\overline{q}_j - \overline{q}_j \mathrm{d}\overline{p}_j \right) \\ &= \sum_j \left(p_j \mathrm{d}q_j - \overline{p}_j \mathrm{d}\overline{q}_j \right) \ , \end{aligned}$$

so that (2.22) follows. *Q.E.D.*

The form $S(\overline{p}, q)$ of the generating function is actually the most common.[8] This because many useful transformations are expressed with a generating function of this form. The examples which follows illustrate some interesting cases.

Example 2.9: *The identity and the scaling transformations.* The generating function $S(\overline{p}, q) = \alpha \sum_j \overline{p}_j q_j$ generates the scaling transformation

$$p_j = \alpha \overline{p}_j \ , \quad \overline{q}_j = \alpha q_j \ , \quad 1 \le j \le n \ .$$

---

[8] Several books report only this form.

For $\alpha = 1$ this is the identity.

Example 2.10: *Extended point transformation.* Suppose that we are given a point transformation $\overline{q} = \overline{q}(q)$ which is a diffeomorphism,[9] so that it admits an inverse $\overline{q} = \overline{q}(q)$ and

$$\det \frac{\partial(\overline{q}_1, \ldots, \overline{q}_n)}{\partial(q_1, \ldots, q_n)} \neq 0 \ , \quad \det \frac{\partial(q_1, \ldots, q_n)}{\partial(\overline{q}_1, \ldots, \overline{q}_n)} \neq 0 \ . \tag{2.31}$$

A corresponding canonical transformation can be constructed using the generating function

$$S(\overline{p}, q) = \sum_k \overline{p}_k \overline{q}_k \big|_{\overline{q} = \overline{q}(q)} \ .$$

For, the complete transformation is

$$q_j = q_j(\overline{q}) \ , \quad p_j = \sum_k \overline{p}_k \frac{\partial \overline{q}_k}{\partial q_j}(q) \ , \quad 1 \leq j \leq n \ .$$

On the other hand, the invertibility condition (2.29) of proposition 2.19 is satisfied in view of (2.31), since

$$\det\left(\frac{\partial^2 S}{\partial \overline{p}_k \partial q_j}\right) = \det\left(\frac{\partial \overline{q}_k}{\partial q_j}\right) \neq 0 \ .$$

This extension is not unique. The most general extended point transformation is generated by the function

$$S(\overline{p}, q) = \sum_k \overline{p}_k \overline{q}_k \big|_{\overline{q} = \overline{q}(q)} + W(q) \ , \tag{2.32}$$

where $W(q)$ is an arbitrary function.

Example 2.11: *Near the identity canonical transformations.* Consider the generating function

$$S(\overline{p}, q) = \sum_j \overline{p}_j q_j + \varepsilon f(\overline{p}, q) \ , \tag{2.33}$$

where $f(\overline{p}, q)$ is an arbitrary function and ε a real parameter, which is assumed to be small. The invertibility condition (2.29) of proposition 2.19 is clearly satisfied for ε small enough. The corresponding canonical transformation in implicit form is

$$p_j = \overline{p}_j + \varepsilon \frac{\partial f}{\partial q_j}(\overline{p}, q) \ , \quad \overline{q}_j = q_j + \varepsilon \frac{\partial f}{\partial \overline{p}_j}(\overline{p}, q) \ , \quad 1 \leq j \leq n \ .$$

[9] This is the class of transformation which are allowed in the framework of the Lagrangian formalism.

The explicit form can be found, e.g., by inverting the second relation with respect to q and replacing the result in the first one. This gives

$$\begin{aligned} q_j &= \overline{q}_j - \varepsilon \frac{\partial f}{\partial \overline{p}_j}(\overline{p}, \overline{q}) + \varepsilon^2 \ldots \\ p_j &= \overline{p}_j + \varepsilon \frac{\partial f}{\partial q_j}(\overline{p}, \overline{q}) + \varepsilon^2 \ldots \ . \end{aligned}$$

For $\varepsilon = 0$ the transformation is the identity, while for $\varepsilon \neq 0$ the coordinates are changed by a little amount. Such a kind of transformations is the basic tool for the development of perturbation theory. However, it can be remarked that the inversion required in order to put the transformation in explicit form is a quite unpleasant aspect, mainly if one plans to perform an explicit calculation. We shall see that inversions can be avoided by using the algorithm of Lie transforms.

The generating functions discussed till now do not actually exhaust the class of canonical transformations. The following example illustrates this point.

Example 2.12: 2^n *canonical transformations.* Let J, K be a partition of the set $\{1, \ldots, n\}$ into two disjoint subsets, $J \cup K = \{1, \ldots, n\}$, $J \cap K = \emptyset$, and consider the canonical transformation

$$\begin{aligned} p_j &= \overline{q}_j \ , \quad \overline{p}_j = -q_j \quad \text{for} \quad j \in J \\ p_k &= \overline{p}_k \ , \quad \overline{q}_k = q_k \quad \text{for} \quad k \in K \ . \end{aligned} \tag{2.34}$$

There are 2^n different transformation of this type. In particular, the exchange of conjugated coordinates of example 2.8 is found by setting $J = \{1, \ldots, n\}$, $K = \emptyset$, and the identity is found by setting $J = \emptyset$, $K = \{1, \ldots, n\}$. We know that the latter two examples are covered by propositions 2.18 and 2.19, respectively, but a generating function of the form above can not be found in all other cases.

All the examples above are actually covered by the following

Proposition 2.20: *Take any partition of the integers $\{1, \ldots, n\}$ into two disjoint sets J, K. Assume that the generating function $S = S(\overline{q}_J, \overline{p}_K, q)$ satisfy the condition*

$$\det \left(\frac{\partial^2 S}{\partial(\overline{q}_J, \overline{p}_K) \partial q} \right) \neq 0 \ . \tag{2.35}$$

Then the transformation implicitly defined by

$$\begin{aligned} \overline{p}_j &= -\frac{\partial S}{\partial \overline{q}_j} \quad \text{for} \quad j \in J \\ \overline{q}_k &= \frac{\partial S}{\partial \overline{p}_k} \quad \text{for} \quad k \in K \\ p_l &= \frac{\partial S}{\partial q_l} \quad \text{for} \quad 1 \leq l \leq n \end{aligned} \tag{2.36}$$

is canonical. Conversely, for any canonical transformation one can find a partition J, K and a generating function of the form above.

Proof. The proof that the transformation is canonical requires checking that proposition 2.17 applies. This is just a minor modification of the proof of proposition 2.19, and is left to the reader.
We prove that all canonical transformations are covered. To this end first remark that the canonical transformation (2.36) is characterized by the condition

$$\det\left(\frac{\partial(p_1,\ldots,p_n)}{\partial(\overline{q}_J,\overline{p}_K)}\right)\neq 0\ . \tag{2.37}$$

If a canonical transformation is given, then the n functions $p_1(\overline{q},\overline{p}),\ldots,p_n(\overline{q},\overline{p})$ are independent, which means that the $n\times 2n$ Jacobian matrix

$$\frac{\partial(p_1,\ldots,p_n)}{\partial(\overline{q}_1,\ldots,\overline{q}_n,\overline{p}_1,\ldots,\overline{p}_n)} \tag{2.38}$$

has rank n. Moreover, in view of $\{p_j,p_k\}=0$ for $j,k=1,\ldots,n$, the n vectors $\mathsf{J}\left(\frac{\partial p_j}{\partial\overline{q}_1},\ldots,\frac{\partial p_j}{\partial\overline{q}_n},\frac{\partial p_j}{\partial\overline{p}_1},\ldots,\frac{\partial p_j}{\partial\overline{q}_n}\right)$ span a n–dimensional Lagrangian subspace of the tangent space to $\mathscr{F}$ at every point $(\overline{q},\overline{p})$. On the other hand, at every point $(\overline{q},\overline{p})$ we can define a canonical basis $\{\mathbf{e}_1,\ldots,\mathbf{e}_n,\mathbf{d}_1,\ldots,\mathbf{d}_n\}$ by just setting

$$\mathbf{e}_j=\left(\frac{\partial\overline{q}_j}{\partial\overline{q}_1},\ldots,\frac{\partial\overline{q}_j}{\partial\overline{q}_n},0,\ldots,0\right)\ ,\quad \mathbf{d}_j=\left(0,\ldots,0,\frac{\partial\overline{p}_j}{\partial\overline{p}_1},\ldots,\frac{\partial\overline{p}_j}{\partial\overline{p}_n}\right)\ ,\quad j=1,\ldots,n\ .$$

By lemma 2.7, the Lagrangian subspace above is complementary to at least one of the Lagrangian arithmentic planes of the canonical basis. This means that there exists a partition $\{J,K\}$ of $\{1,\ldots,n\}$ such that the $2n\times 2n$ matrix obtained by adding to (2.38) the n rows $\{\mathbf{e}_k\}_{k\in K}\cup\{\mathbf{d}_j\}_{j\in J}$ has non zero determinant. On the other hand, the determinant turns out to be exactly (2.37), so that we conclude that J,K is the wanted partition. *Q.E.D.*

2.5 Time–depending canonical transformations

We show here how the theory of canonical transformations can be generalized so that the cases of non autonomous Hamiltonians and of time depending transformations are taken into account.

We use the extension of the phase space discussed in sect. 1.1.1, i.e., we introduce two further canonical variables q_0,p_0 and for a given Hamiltonian $H(q,p,t)$ we consider the Hamiltonian in the extended phase space[10]

$$\tilde{H}(q,p,q_0,p_0)=H(q,p,q_0)+p_0\ . \tag{2.39}$$

On the extended space phase we can perform canonical transformations of the form $q=q\,(\overline{q},\overline{p},\overline{q}_0,\overline{p}_0)$, $p=p\,(\overline{q},\overline{p},\overline{q}_0,\overline{p}_0)$, $q_0=q_0(\overline{q},\overline{p},\overline{q}_0,\overline{p}_0)$, $p_0=p_0(\overline{q},\overline{p},\overline{q}_0,\overline{p}_0)$, to which

[10] We emphasize the particular role played by the variables q_0,p_0 by denoting the canonical coordinates in phase space as (q,p,q_0,p_0), where $q=(q_1,\ldots,q_n)$ and $p=(p_1,\ldots,p_n)$.

the theory developed till now applies. However, this means that we change also the time variable, in the sense that the new variable $\overline{q}_0$ will not evolve uniformly in time.

The natural choice is to consider a restricted class of transformations which keeps the coordinate q_0 invariant, namely $q_0 = \overline{q}_0$; in turn, $p_0 = p_0(\overline{q}, \overline{p}, \overline{q}_0, \overline{p}_0)$ will be determined so as to fulfill the canonicity conditions. A first consequence is that the condition $\{q_0, p_0\} = 1$ implies $p_0 = \overline{p}_0 + f(\overline{q}, \overline{p}, q_0)$, with some function f. A second consequence is that the canonicity conditions $\{q_0, q_j\} = \{q_0, p_j\} = 0$ for $1 \le j \le n$ imply that q_j, p_j do not depend on $\overline{p}_0$. This also means that the Poisson brackets $\{q_j, q_k\}$, $\{q_j, p_k\}$ and $\{p_j, p_k\}$ are actually computed by differentiating only with respect to the variables $\overline{q}, \overline{p}$.

The general scheme is the following: in order to perform a time dependent transformation we first consider the extended phase space and the Hamiltonian (2.39), and perform a transformation satisfying the conditions above. This means that the transformed Hamiltonian will take the form

$$\overline{H}(\overline{q}, \overline{p}, \overline{q}_0, \overline{p}_0) = H(q, p, q_0)\big|_{q=q(\overline{q},\overline{p},\overline{q}_0),\, p=p(\overline{q},\overline{p},\overline{q}_0), q_0=\overline{q}_0} + \overline{p}_0 + f(\overline{q}, \overline{p}, \overline{q}_0)\ .$$

In view of the linear dependence on $\overline{p}_0$ we remove the extension of the phase space by setting again $\overline{q}_0 = t$ and removing the term $\overline{p}_0$, thus obtaining the transformed Hamiltonian

$$\overline{H}(\overline{q}, \overline{p}, t) = H(q, p, t)\big|_{q=q(\overline{q},\overline{p},t),\, p=p(\overline{q},\overline{p},t)} + f(\overline{q}, \overline{p}, t)\ .$$

We emphasize that the new Hamiltonian is not merely the transformed function of the old one: there is an extra term that must be computed. The following propositions show that the canonicity conditions to be checked are the ones discussed till now, that must be fulfilled identically in t, and explain how to determine the extra term in the Hamiltonian.

Proposition 2.21: *Let $q = q(\overline{q}, \overline{p}, t)$, $p = p(\overline{q}, \overline{p}, t)$ be a time dependent transformation which preserves the fundamental Poisson brackets identically in t. Then the transformation is canonical, and there exists a function $F(q, p, t)$ such that the transformed Hamiltonian is*

$$\overline{H}(\overline{q}, \overline{p}, t) = \big[H(q, p, t) - F(q, p, t)\big]_{q=q(\overline{q},\overline{p},t),p=p(\overline{q},\overline{p},t)} \tag{2.40}$$

Proof. We just prove that in the extended phase space there is a function $F(q, p, \overline{q}_0)$ such that the extended transformation

$$\begin{aligned} q &= q(\overline{q}, \overline{p}, \overline{q}_0)\ , & p &= p(\overline{q}, \overline{p}, \overline{q}_0)\ , \\ q_0 &= \overline{q}_0 \qquad , & p_0 &= \overline{p}_0 - F(q, p, \overline{q}_0)\big|_{q=q(\overline{q},\overline{p},\overline{q}_0)\,,\, p=p(\overline{q},\overline{p},\overline{q}_0)} \end{aligned} \tag{2.41}$$

is canonical. Differentiating with respect to $\overline{q}_0$ the relations $\{q_j, q_k\} = \{p_j, p_k\} = 0\,,\ \{q_j, p_k\} = \delta_{jk}$ for $1 \le j \le n$ (true in view of the assumed preservation of the fundamental Poisson brackets) we get

$$\left\{\frac{\partial q_j}{\partial \overline{q}_0}, q_k\right\} + \left\{q_j, \frac{\partial q_k}{\partial \overline{q}_0}\right\} = 0 \ ,$$
$$\left\{\frac{\partial q_j}{\partial \overline{q}_0}, p_k\right\} + \left\{q_j, \frac{\partial p_k}{\partial \overline{q}_0}\right\} = 0 \ ,$$
$$\left\{\frac{\partial p_j}{\partial \overline{q}_0}, p_k\right\} + \left\{p_j, \frac{\partial p_k}{\partial \overline{q}_0}\right\} = 0 \ .$$

Recalling that q, p do not depend on $\overline{p}_0$, and denoting

$$f_j = \frac{\partial q_j}{\partial \overline{q}_0} \ , \quad g_j = \frac{\partial p_j}{\partial \overline{q}_0} \ ,$$

we write the identities above as

$$\frac{\partial f_j}{\partial p_k} - \frac{\partial f_k}{\partial p_j} = 0 \ , \quad \frac{\partial f_j}{\partial q_k} - \frac{\partial g_k}{\partial p_j} = 0 \ , \quad \frac{\partial g_j}{\partial q_k} - \frac{\partial g_k}{\partial q_j} = 0 \ .$$

This implies the (local) existence of a function $F(q, p, \overline{q}_0)$ such that

$$\frac{\partial q_j}{\partial \overline{q}_0} = \frac{\partial F}{\partial p_j} \ , \quad \frac{\partial p_j}{\partial \overline{q}_0} = -\frac{\partial F}{\partial q_j} \ . \tag{2.42}$$

With this function we complete the transformation as in (2.41). The canonicity of the extended transformation is checked by remarking that $\{q_0, q_j\} = \{q_0, p_j\} = 0, \{q_0, p_0\} = 1$ in view of the assumed preservation of the fundamental Poisson brackets for all t, and so for all $\overline{q}_0$, and that

$$\{p_0, q_j\} = \frac{\partial q_j}{\partial \overline{q}_0} - \{F, q_j\} \ , \quad \{p_0, p_j\} = \frac{\partial p_j}{\partial \overline{q}_0} - \{F, p_j\}$$

are zero in view of (2.42). Replacing the transformation in the Hamiltonian (2.39) and removing the extension of the phase space the claim follows. *Q.E.D.*

Proposition 2.22: *Let $S(\overline{p}, q, t)$ be a function satisfying*

$$\det\left(\frac{\partial^2 S}{\partial \overline{p}_j \partial q_k}\right) \neq 0 \ .$$

Then the transformation implicitly defined by

$$\overline{q}_j = \frac{\partial S}{\partial \overline{p}_j} \ , \quad p_j = \frac{\partial S}{\partial q_j} \ , \quad 1 \leq j \leq n$$

is canonical, and the transformed Hamiltonian takes the form

$$\overline{H}(\overline{q}, \overline{p}, t) = H(q, p, t)\Big|_{q=q(\overline{q},\overline{p},t),\, p=p(\overline{q},\overline{p},t)} + \frac{\partial S}{\partial t}(\overline{p}, q, t)\Big|_{q=q(\overline{q},\overline{p},t)} \ . \tag{2.43}$$

Proof. In the extended phase space consider the generating function

$$\tilde{S}(\overline{p}, q, \overline{p}_0, q_0) = \overline{p}_0 q_0 + S(\overline{p}, q, q_0) \ .$$

The corresponding transformation is

$$(2.44)\qquad \begin{aligned} \overline{q}_j &= \frac{\partial S}{\partial \overline{p}_j}\ , \quad p_j = \frac{\partial S}{\partial q_j}\ , \quad 1 \le j \le n\ ,\\ \overline{q}_0 &= q_0\ , \qquad p_0 = \overline{p}_0 + \frac{\partial S}{\partial q_0}\ . \end{aligned}$$

Replacing the transformation in the Hamiltonian (2.39) and removing the extension of the phase space the claim follows. *Q.E.D.*

2.6 The equation of Hamilton–Jacobi

The integration of the canonical equations can be performed by looking for a generating function of a canonical transformation giving the Hamiltonian a particularly simple form. It is customary to use the formalism of time dependent canonical transformation.

Having given the Hamiltonian $H(q, p, t)$ we look for a function S which is a solution of the *Hamilton–Jacobi equation* ([33], [35], [39], [40])

$$(2.45)\qquad H\left(q, \frac{\partial S}{\partial q}, t\right) + \frac{\partial S}{\partial t} = 0\ .$$

We are actually looking for the generating function of a transformation such that the transformed Hamiltonian is identically zero. The problem is to find a solution of eq. (2.45) depending on $q_1, \ldots, q_n, t$ and on n arbitrary parameters $\alpha_1, \ldots, \alpha_n$; this is said to be a *complete integral.*

Proposition 2.23: *Consider the Hamiltonian $H(q, p, t)$, and assume that we are given a complete integral $S(\alpha, q, t)$ of Hamilton–Jacobi's equation (2.45), depending on n arbitrary parameters $\alpha_1, \ldots, \alpha_n$ and satisfying*

$$\det\left(\frac{\partial^2 S}{\partial \alpha_j \partial q_k}\right) \ne 0\ .$$

Then the solutions of the canonical equations are written in implicit form as

$$(2.46)\qquad \beta_j = \frac{\partial S}{\partial \alpha_j}(\alpha, q, t)\ , \quad p_j = \frac{\partial S}{\partial q_j}(\alpha, q, t)\ , \quad 1 \le j \le n\ ,$$

where $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ are constants depending on the initial data.

Proof. The function $S(\alpha, q, t)$ satisfies the conditions of proposition 2.22; therefore, it is the generating function of a canonical transformation, actually the transformation (2.46). Since the transformed Hamiltonian is identically zero, the corresponding canonical equations are

$$\dot{\alpha}_j = 0\ , \quad \dot{\beta}_j = 0\ , \quad 1 \le j \le n\ ,$$

i.e., α, β are constants depending on the initial data. By inversion of (2.46) with respect to q, p one gets functions

$$q = q(\alpha, \beta, t) \ , \quad p = p(\alpha, \beta, t) \ ,$$

i.e., the wanted solutions of the canonical equations. *Q.E.D.*

Example 2.13: *Free particle* Let the Hamiltonian be

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \ .$$

The corresponding Hamilton–Jacobi's equation is

$$\frac{1}{2m}\left[\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] + \frac{\partial S}{\partial t} = 0 \ .$$

We use the method of separation of variables. We look for a solution of the form

$$S(x, y, z, t) = X(x) + Y(y) + Z(z) + T(t) \ ,$$

so that the equation is rewritten as

$$\frac{1}{2m}\left[\left(\frac{dX}{dx}\right)^2 + \left(\frac{dY}{dy}\right)^2 + \left(\frac{dZ}{dz}\right)^2\right] + \frac{dT}{dt} = 0 \ . \tag{2.47}$$

Therefore, we obtain the equations[11]

$$\frac{dX}{dx} = \alpha_x \ , \quad \frac{dY}{dy} = \alpha_y \ , \quad \frac{dZ}{dz} = \alpha_z \ , \quad \frac{dT}{dt} = -\frac{\alpha_x^2 + \alpha_y^2 + \alpha_z^2}{2m} \ ,$$

with $\alpha_x, \alpha_y, \alpha_z$ arbitrary constants. By integration we construct the generating function

$$S(\alpha_x, \alpha_y, \alpha_z, x, y, z, t) = \alpha_x x + \alpha_y y + \alpha_z z - \frac{\alpha_x^2 + \alpha_y^2 + \alpha_z^2}{2m} t \ ,$$

so that the transformation is

$$\begin{aligned} p_x &= \alpha_x \ , & p_y &= \alpha_y \ , & p_z &= \alpha_z \\ \beta_x &= x - \frac{\alpha_x}{m} t \ , & \beta_y &= y - \frac{\alpha_y}{m} t \ , & \beta_z &= z - \frac{\alpha_z}{m} t \ . \end{aligned}$$

This is the solution of the canonical equations.

[11] Differentiating (2.47) with respect to x we get $\frac{d^2X}{dx^2} = 0$, so that $\frac{dX}{dx}$ must be constant. Similarly, differentiating with respect to y, z and t we get that $\frac{dY}{dy}$, $\frac{dZ}{dz}$ and $\frac{dT}{dt}$ are constants, too. In view of (2.47), only three of these constants are arbitrary.

3

INTEGRABLE SYSTEMS

The present chapter deals with general methods that should allow us to integrate the Hamilton's equations, using the tools developed in the previous chapters. This was indeed the dream of the great mathematicians, after Newton, till the end of the XIX century. Nowadays we are well aware that integrable systems are rather exceptional, and this is indeed a good justification of the fact that all textbooks contain the same examples — just a few which are the classical and more interesting ones. Nevertheless, integrable systems represent an excellent first approximation of interesting mechanical systems, and are the starting point of classical perturbation theory. It is not far from reality to say that in most cases the goal of perturbation methods is to reduce a system of differential equations to a form as close as possible to an integrable one.

Before entering the discussion we should make an agreement on the meaning of the term *integrable system*. In view of the theorem of existence and uniqueness of the solutions of a system of differential equations every Hamiltonian system can be said to be integrable provided some mild regularity conditions are satisfied by the Hamiltonian function. This is useful, of course, if one is interested in computing the orbit corresponding to a given initial datum, e.g., with numerical methods.[1] However, we should keep in mind that the theorem has a *local* character: in our case it assures only the existence of the solution for some time interval. The process of continuation of a given solution may be used in order to establish the *existence* of the solution for larger time intervals, but it gives essentially no information about the *global behaviour* of the orbits. In the framework of Hamiltonian systems (although this is not a true

[1] The most common numerical methods for solving differential equations are indeed based on the possibility of writing the first few terms of the Taylor expansion of the solution. The computation of the orbit is performed by repeating an elementary iteration step: starting from the initial point at time 0 one computes the (approximate) point at time τ; then the new point is used as initial point for the next step, and so on. However, the expansion is only *local*. Moreover, small error which are unavoidably introduced at every step may accumulate. Proving that the computed orbit remains close to the true orbit for a long time interval is, generally speaking, an hard problem.

restriction) it is customary to assign a more precise meaning to the word *integrability*. In some sense, one asks for being able to write the solution for all times.

The traditional interpretation involves the concept of *integrability by quadratures*. This means that the solution has to be found via a finite number of algebraic operations, including inversion of functions, and of computation of integrals of known functions (quadrature). The integration method discussed in sect. 1.3 for systems with one degree of freedom is a good example. In the framework of Hamiltonian theory Liouville's theorem can be considered as the most advanced general result in this direction. The paradigm model is represented by an Hamiltonian depending only on the momenta $p_1, \ldots, p_n$, i.e., $H = H(p_1, \ldots, p_n)$, which is trivially integrable. In short, Liouville's theorem says that if a Hamiltonian system possesses enough first integrals then the Hamiltonian can be given the form above with a suitable coordinate transformation.

In more recent times more attention is paid to the *global* description of the behavior of the solutions, with particular attention to the existence of *periods*, or *frequencies*. Thus, most authors impose the sharper condition that the coordinates $q_1, \ldots, q_n$ conjugated to $p_1, \ldots, p_n$ are actually angles, that is, $q \in \mathbb{T}^n$. In the latter case the canonically conjugated variables p, q are called *action–angle variables*. This seems to be a very strong condition: for instance, the problem of a mass point freely moving on the space can not be described by action–angle variables in strict sense because there are no periods. However, such a strong attitude can be justified *a posteriori*. Indeed, small perturbations of an integrable system that admits action–angle variables typically produce a very complicated dynamical behavior, which is still not completely understood.[2]

In view of this discussion, we consider the following problem: *Assume we are given a Liouville–integrable system. Can we introduce action–angle variables?* The answer to this question is furnished by the theorem of Arnold–Jost.[3]

The theorems of Liouville and of Arnold–Jost constitute the main contents of this chapter. A general discussion of the dynamical behaviour of an integrable system is also included.

---

[2] A typical situation arising in Mechanics is the study of a system of many particles, that may be either free of moving all around the space under the mutual interactions, as is the case of our planetary system or of an atomic system, or may be subjected to some constraints, as, e.g., in the case of a rigid body. The first step usually consists in exploiting the conservation of the total momentum by eliminating the motion of the center of mass, which is a trivial one being that of a free particle. Then in many cases action–angle variables may be introduced in some approximation.

[3] The use of action–angle variables was well known in connection with classical problems like the planetary motions and the motion of a rigid body. During the first decades of this century it has become relevant also in connection with the first developments of quantum theory. A classical and valuable reference is M. Born's treatise [15]. The theorem of Arnold–Jost states that integrability in classical mechanics is strongly connected with the existence of action–angle variables.

3.1 Involution systems

The investigation of integrability is mainly based on the existence of independent first integrals. In the framework of Hamiltonian system a relevant role is played by first integrals with vanishing mutual Poisson bracket.

A system of r functions $\{\Phi_1(q,p),\ldots,\Phi_r(q,p)\}$ is said to be an *involution system* if the functions are independent, i.e.,

$$\mathrm{rank}\left(\frac{\partial(\Phi_1,\ldots,\Phi_r)}{\partial(q_1,\ldots,q_n,p_1,\ldots,p_n)}\right)=r\ ,$$

and the Poisson bracket between any two functions vanishes, i.e., $\{\Phi_j,\Phi_k\}=0$ for $j,k=1,\ldots,r$.

3.1.1 Some geometrical properties.

We prove some lemmas that will be used in the rest of the chapter.

Lemma 3.1: *An involution system contains at most n independent functions, where n is the number of degrees of freedom.*

Proof. It is convenient to use the compact notation. At any point $\mathbf{z}\in\mathscr{F}$ the symplectic gradients $(\mathsf{J}\partial_z\Phi_1,\ldots,\mathsf{J}\partial_z\Phi_r)$ span a r–dimensional subspace which is isotropic, due to the involution property satisfied by the functions. By lemma 2.5 the dimension of such a subspace can not exceed n. We conclude that $r\le n$. *Q.E.D.*

Example 3.1: *Involution systems contructed using the canonical coordinates.* The most trivial but useful example is given by the canonical coordinates themselves. Consider any partition J,K of $\{1,\ldots,n\}$; then the n functions $\{q_j\}_{j\in J}\cup\{p_k\}_{k\in K}$ form an involution system.

Other examples may be easily constructed by making reference to known integrable systems. For instance, let the phase space be $\mathbb{R}^3\times\mathbb{R}^3$, with canonical coordinates x,y,z and momenta p_x,p_y,p_z. The latter three quantities are the components of the momentum, and form an involution system. This reminds us the case of a free particle in the ordinary euclidean space. On the other hand, it is just a particular case of the first example, since it corresponds to a partition which selects only the momenta.

Example 3.2: *Using the angular momentum in spherical coordinates* It seems spontaneous to try to construct involution systems by using the three components of the angular momentum, namely $M_x=yp_z-zp_y$, $M_y=zp_x-xp_z$, $M_z=xp_y-yp_x$. However, it is immediately seen that the latter three quantities are independent, but not in involution: this has been shown in example 1.10. Replacing some components of the angular momentum with some components of the momentum does not help, for the same reason.

An involution system may be constructed by considering one of the components of the angular momentum, for instance M_z, and the quantity $\Gamma^2=M_x^2+M_y^2+M_z^2$, namely the square of the norm of the angular momentum. The latter two quantities are indeed in involution. A third function in involution with M_x and Γ^2 is, e.g.,

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(r) \ ,$$

where $r = \sqrt{x^2 + y^2 + z^2}$, and $V(r)$ an arbitrary (differentiable) function.
The same example can be reformulated using spherical coordinates. The phase space is $(0, +\infty) \times (0, \pi) \times \mathbb{T} \times \mathbb{R}^3$, with canonical coordinates $r, \vartheta, \varphi, p_r, p_\vartheta, p_\varphi$. The three functions

$$J = p_\varphi \ , \quad \Gamma^2 = p_\vartheta^2 + \frac{J^2}{\sin^2\vartheta} \ , \quad E = \frac{1}{2m}\left(p_r^2 + \frac{\Gamma^2}{r^2}\right) + V(r)$$

form an involution system. One will recognize here the first integrals of the problem of motion under central forces, example 1.12.

Example 3.3: *Harmonic oscillators* A last example is constructed by considering a system of harmonic oscillators. Let the phase space to be $\mathbb{R}^{2n}$, with canonical coordinates x, y. The n functions

$$\Phi_1 = \frac{x_1^2 + y_1^2}{2}, \ldots, \Phi_n = \frac{x_n^2 + y_n^2}{2}$$

form an involution system.
Although apparently trivial this example plays a central role in studying the small oscillations of a system in the neighbourhood of an equilibrium, since it represents a remarkable first approximation.

In the rest of the chapter we shall need the following technical

Lemma 3.2: *Let $\Phi_1, \ldots, \Phi_n$ be an involution system on the phase space $\mathscr{F}$. Then at every point $P \in \mathscr{F}$ there is a partition J, K of $\{1, \ldots, n\}$ such that*

$$\det\left(\frac{\partial(\Phi_1, \ldots, \Phi_n)}{\partial(q_J, p_K)}\right) \neq 0 \ .$$

The proof is a straightforward adaptation of that of proposition 2.20, and is left to the reader.

3.1.2 The Hamiltonian flow as a canonical transformation

Consider a point (q, p) of the phase space, and let $(q_t, p_t) = \phi^t(q, p)$ be the transformed point under the flow ϕ^t generated by a canonical system with Hamiltonian $H(q, p)$. For fixed t we can consider (q_t, p_t) as new coordinates of the point (q, p). That is, we consider the flow as generating a coordinate transformation on the phase space.

Lemma 3.3: *Let ϕ^t denote the flow generated by the Hamiltonian $H(q, p)$. Then for every fixed t the transformation $\bigl(q_t(q, p), p_t(q, p)\bigr) = \phi^t(q, p)$ is canonical.*

Proof. It is enough to prove that for every t the fundamental Poisson brackets are preserved, namely that

$$\begin{aligned} &\{q_{j,t}, q_{k,t}\}_{q,p} = \{p_{j,t}, p_{k,t}\}_{q,p} = 0 \ , \\ &\{q_{j,t}, p_{k,t}\}_{q,p} = \delta_{j,k} \ . \end{aligned}$$

This is true for $t = 0$, because $\phi^0(q,p) = (q,p)$ is the identity. Let us prove that at every point (q,p) of the phase space one has

$$\frac{d}{dt}\{q_{j,t}, q_{k,t}\} = \frac{d}{dt}\{p_{j,t}, p_{k,t}\} = \frac{d}{dt}\{q_{j,t}, p_{k,t}\} = 0 \ ,$$

To this end, recalling again that $q_{j,0} = q_j$, $p_{j,0} = p_j$ and using the Hamilton's equations we have

$$q_{j,t} = q_j + t\frac{\partial H}{\partial p_j} + \ldots \ , \quad p_{j,t} = p_j - t\frac{\partial H}{\partial q_j} + \ldots \ ,$$

where the dots stand for terms of higher order in t. Thus we have

$$\{q_{j,t}, q_{k,t}\} = \{q_j, q_k\} + t\left[\left\{\frac{\partial H}{\partial p_j}, q_k\right\} + \left\{q_j, \frac{\partial H}{\partial p_k}\right\}\right] + \ldots \ ,$$

which in turn means that one has

$$\frac{d}{dt}\{q_{j,t}, q_{k,t}\} = \left\{\frac{\partial H}{\partial p_j}, q_k\right\} + \left\{q_j, \frac{\partial H}{\partial p_k}\right\} = -\frac{\partial}{\partial p_k}\frac{\partial H}{\partial p_j} + \frac{\partial}{\partial p_j}\frac{\partial H}{\partial p_k} = 0 \ .$$

With a similar calculation we get

$$\begin{aligned}
\frac{d}{dt}\{p_{j,t}, p_{k,t}\} &= -\left\{\frac{\partial H}{\partial q_j}, p_k\right\} - \left\{p_j, \frac{\partial H}{\partial q_k}\right\} = 0 \ , \\
\frac{d}{dt}\{q_{j,t}, p_{k,t}\} &= \quad \left\{\frac{\partial H}{\partial p_j}, p_k\right\} - \left\{q_j, \frac{\partial H}{\partial q_k}\right\} = 0 \ .
\end{aligned}$$

Since the fundamental Poisson brackets have a zero time derivative at every point, they keep a constant value along every orbit. *Q.E.D.*

The proposition has a suggestive geometrical interpretation: the Hamiltonian flow can be seen as the unfolding of a canonical transformation parametrically depending on time.

3.1.3 *Variational equations and first integrals*

Let a system of differential equations (which needs not be Hamiltonian)

$$\dot{x}_j = X_j(x_1, \ldots, x_n) \ , \quad 1 \le j \le n \ , \tag{3.1}$$

be given and let $x(t)$ be an orbit with initial point x_0. Let also $x_0 + \delta x_0$ be a point close to x_0, with an infinitesimal increment δx_0, and let $x(t) + \delta x(t)$ be the corresponding orbit, so that it is a solution of the differential equations

$$\begin{aligned}
\frac{d}{dt}(x_j + \delta x_j) &= X_j(x_1 + \delta x_1, \ldots, x_n + \delta x_n) \\
&= X_j(x_1, \ldots, x_n) + \sum_{l=1}^{n} \frac{\partial X_j}{\partial x_l}(x_1, \ldots, x_n)\, \delta x_l + \ldots \ .
\end{aligned}$$

where the dots denote terms of higher order in δx. Since $x(t)$ is assumed to be a solution of the equation $\dot{x} = X(x)$, one immediately gets that $\delta x(t)$ obeys the so called *variational equation*

$$\frac{d}{dt}\delta x_j = \sum_{l=1}^{n} \frac{\partial X_j}{\partial x_l}\delta x_l \ , \quad 1 \le j \le n \ , \tag{3.2}$$

where the functions $\frac{\partial X_j}{\partial x_l}(x_1, \ldots, x_n)$ must be evaluated along the known solution $x(t)$.

A similar procedure applies to the Hamiltonian case. Let us do it in detail, recalling that the canonical equations have the rather particular form

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \ , \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} \ , \quad 1 \le j \le n \ . \tag{3.3}$$

Let us denote by $\delta q_j, \delta p_j$ respectively the increments with respect to the variables q_j, p_j. Then the variational equations are

$$\begin{aligned} \frac{d}{dt}\delta q_j &= \sum_{l=1}^{n} \left(\frac{\partial^2 H}{\partial p_j \partial q_l}\delta q_l + \frac{\partial^2 H}{\partial p_j \partial p_l}\delta p_l \right) \\ \frac{d}{dt}\delta p_j &= -\sum_{l=1}^{n} \left(\frac{\partial^2 H}{\partial q_j \partial q_l}\delta q_l + \frac{\partial^2 H}{\partial q_j \partial p_l}\delta p_l \right) \end{aligned} \tag{3.4}$$

An interesting relation between first integrals and variational equations is given by the following

Proposition 3.4: *Let Φ be a first integral of the canonical system with Hamiltonian $H(q,p)$. Then a solution of the variational equations (3.4) is*

$$\delta q_j = \tau \frac{\partial \Phi}{\partial p_j} \ , \quad \delta p_j = -\tau \frac{\partial \Phi}{\partial q_j} \ , \quad 1 \le j \le n \ , \tag{3.5}$$

where $\tau \neq 0$ is an arbitrary constant.

Proof. In view of the linearity of the variational equations it is enough to prove the statement for $\tau = 1$. By differentiating the relation $\{\Phi, H\} = 0$ we immediately get

$$\left\{ \frac{\partial \Phi}{\partial p_j}, H \right\} + \left\{ \Phi, \frac{\partial H}{\partial p_j} \right\} = 0 \ , \quad \left\{ \frac{\partial \Phi}{\partial q_j}, H \right\} + \left\{ \Phi, \frac{\partial H}{\partial q_j} \right\} = 0 \ , \quad 1 \le j \le n \ ,$$

that is

$$\frac{d}{dt}\frac{\partial \Phi}{\partial p_j} = \left\{ \frac{\partial H}{\partial p_j}, \Phi \right\} \ , \quad \frac{d}{dt}\frac{\partial \Phi}{\partial q_j} = \left\{ \frac{\partial H}{\partial q_j}, \Phi \right\} \ .$$

Writing in explicit form the r.h.s. of these equations we get

$$\begin{aligned} \frac{d}{dt}\frac{\partial \Phi}{\partial p_j} &= \sum_{l=1}^{n} \left(\frac{\partial^2 H}{\partial p_j \partial q_l}\frac{\partial \Phi}{\partial p_l} - \frac{\partial^2 H}{\partial p_j \partial p_l}\frac{\partial \Phi}{\partial q_l} \right) \ , \\ \frac{d}{dt}\frac{\partial \Phi}{\partial q_j} &= \sum_{l=1}^{n} \left(\frac{\partial^2 H}{\partial q_j \partial q_l}\frac{\partial \Phi}{\partial p_l} - \frac{\partial^2 H}{\partial q_j \partial p_l}\frac{\partial \Phi}{\partial q_l} \right) \ , \end{aligned}$$

which in view of (3.5) coincides with (3.4). *Q.E.D.*

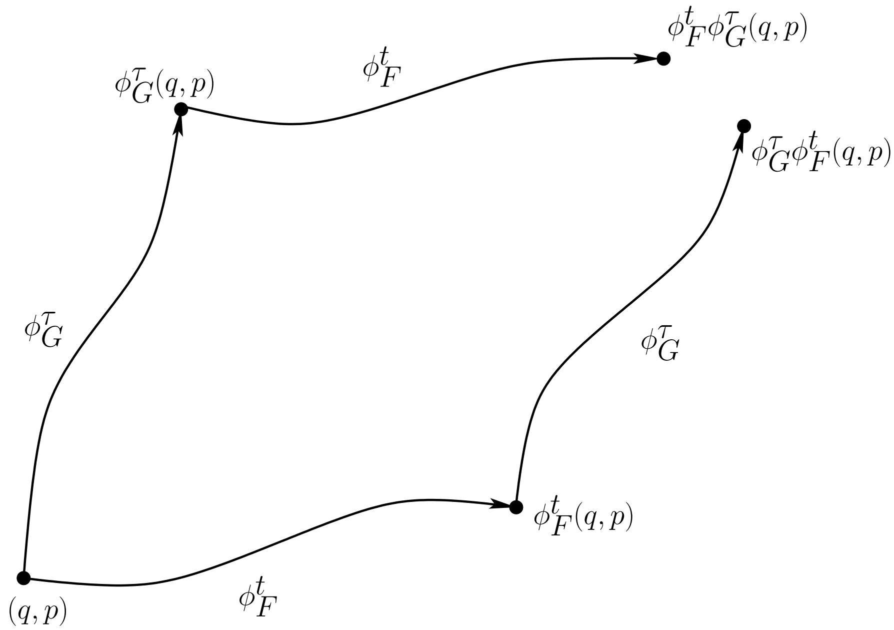

Figure 3.1. Illustrating the commutation of flows.

`*** Questo va spiegato meglio ***` Let me add a remark. The proposition essentially says that the increment $(\delta q, \delta p)$ in (3.5) is the linear approximation in τ of the canonical flow due to $\Phi(q,p)$. In compact notation we may write the initial increment as $\delta \mathbf{z}_0 = \tau \mathsf{J} \partial_z \Phi\big|_{\mathbf{z}_0}$. Then the statement of the theorem may be rewritten as

$$\phi^t\bigl(\tau \mathsf{J} \partial_z \Phi\big|_{\mathbf{z}}\bigr) = \tau \mathsf{J} \partial_z \Phi\big|_{\phi^t \mathbf{z}} \ . \tag{3.6}$$

That is: the linear approximation of the increment at the point $\phi^t \mathbf{z}$ is still the approximated flow at time τ of the Hamiltonian field generated by Φ. This remark will be used in the next section.

3.1.4 Commutation of canonical flows

I come now to consider the commutation of canonical flows. If two functions $F(q,p)$ and $G(q,p)$ are given we can consider both of them on the same foot, i.e., both generate a Hamiltonian vector field, and so a canonical flow. Now, the question is the following, as illustrated in fig. 3.1. Given an initial point (q,p), let us follow the flow of F up to a time t, thus getting at the point $\phi_F^t(q,p)$, and then follow the flow of G up to time τ, ending up in $\phi_G^\tau \circ \phi_F^t(q,p)$. Then exchange the order of the flow, thus moving from (q,p) to $\phi_G^\tau(q,p)$ along the flow of G and then to $\phi_F^t \circ \phi_G^\tau(q,p)$ along the flow of F, the times t and τ being unchanged. In general one cannot expect the two end points to be the same. The interesting fact is that *if F and G are in involution then the final points coincide.*

A suggestive intepretation is the following. Consider an orbit of F, and apply the flow ϕ_G^τ to every point of the orbit. The resulting set is again an orbit of F. The role of F and G may be exchanged, of course.

Proposition 3.5: *Let the dynamical variables $F(q,p)$ and $G(q,p)$ be in involution, and consider the canonical flows at times t and τ generated by F and G, respectively, i.e.*

$$(q_t, p_t) = \phi_F^t(q,p) \ , \quad (q_\tau, p_\tau) = \phi_G^\tau(q,p) \ . \tag{3.7}$$

Then the following statements hold true.

(i) The function G is invariant for the canonical flow generated by F; conversely, the function F is invariant for the canonical flow generated by G. That is:

$$F(q_\tau, p_\tau)\Big|_{(q_\tau,p_\tau)=\phi_G^\tau(q,p)} = F(q,p) \ , \quad G(q_t, p_t)\Big|_{(q_t,p_t)=\phi_F^t(q,p)} = G(q,p) \ .$$

(ii) The flows (3.7) do commute, i.e., for every (q,p) we have

$$\phi_G^\tau \circ \phi_F^t(q,p) = \phi_F^t \circ \phi_G^\tau(q,p) \ .$$

Proof. (i) It is just a reformulation of the claim that G is a first integral for F and vice versa.

`*** Migliorare la dimostrazione ***`

(ii) It is convenient to use the compact notation. Let t, τ be infinitesimal quantities. Then for any $\mathbf{z}$ we have

$$\phi_F^t \mathbf{z} = \mathbf{z} + t\mathsf{J}\partial_z F\big|_{\mathbf{z}} + \ldots \ , \quad \phi_G^\tau \mathbf{z} = \mathbf{z} + \tau\mathsf{J}\partial_z G\big|_{\mathbf{z}} + \ldots \ ,$$

the dots denoting terms of higher order in t and τ. Thus, replacing $\mathbf{z}$ with $\phi_G^\tau\mathbf{z}$ or $\phi_F^t\mathbf{z}$ as appropriate, we get

$$\begin{aligned} \phi_F^t \circ \phi_G^\tau \mathbf{z} &= \phi_G^\tau \mathbf{z} + t\mathsf{J}\partial_z F\big|_{\phi_G^\tau \mathbf{z}} + \ldots \ , \\ \phi_G^\tau \circ \phi_F^t \mathbf{z} &= \phi_F^t \mathbf{z} + \tau\mathsf{J}\partial_z G\big|_{\phi_F^t \mathbf{z}} + \ldots \ . \end{aligned} \tag{3.8}$$

On the other hand we have also

$$\phi_F^t \circ \phi_G^\tau \mathbf{z} = \phi_F^t\big(\mathbf{z} + \tau\mathsf{J}\partial_z G\big|_{\mathbf{z}} + \ldots\big) \ ,$$

i.e., we are considering the orbit with initial point $z + \tau\mathsf{J}\partial_z G\big|_{\mathbf{z}}$. By proposition 3.4, rewritten as in (3.6), we have $\phi^t\big(\tau\mathsf{J}\partial_z G\big|_{\mathbf{z}}\big) = \tau\mathsf{J}\partial_z G\big|_{\phi^t\mathbf{z}}$ so that we can write

$$\phi_F^t \circ \phi_G^\tau \mathbf{z} = \phi_F^t \mathbf{z} + \tau\mathsf{J}\partial_z G\big|_{\phi_F^t \mathbf{z}} + \ldots \ . \tag{3.9}$$

With a similar calculation, exchanging t with τ and F with G, we have also

$$\phi_G^\tau \circ \phi_F^t \mathbf{z} = \phi_G^\tau \mathbf{z} + t\mathsf{J}\partial_z F\big|_{\phi_G^\tau \mathbf{z}} + \ldots \ . \tag{3.10}$$

Thus, subtracting (3.9) and (3.10) from the second and the first of (3.8), respectively, we see that the contributions which are linear in t and τ do vanish. Thus at every point x we have

$$\frac{\partial}{\partial t}[\phi_F^t, \phi_G^\tau]\mathbf{z} = \frac{\partial}{\partial \tau}[\phi_F^t, \phi_G^\tau]\mathbf{z} = 0$$

which in view of $[\phi_F^0, \phi_G^0]\mathbf{z} = 0$ proves the claim. *Q.E.D.*

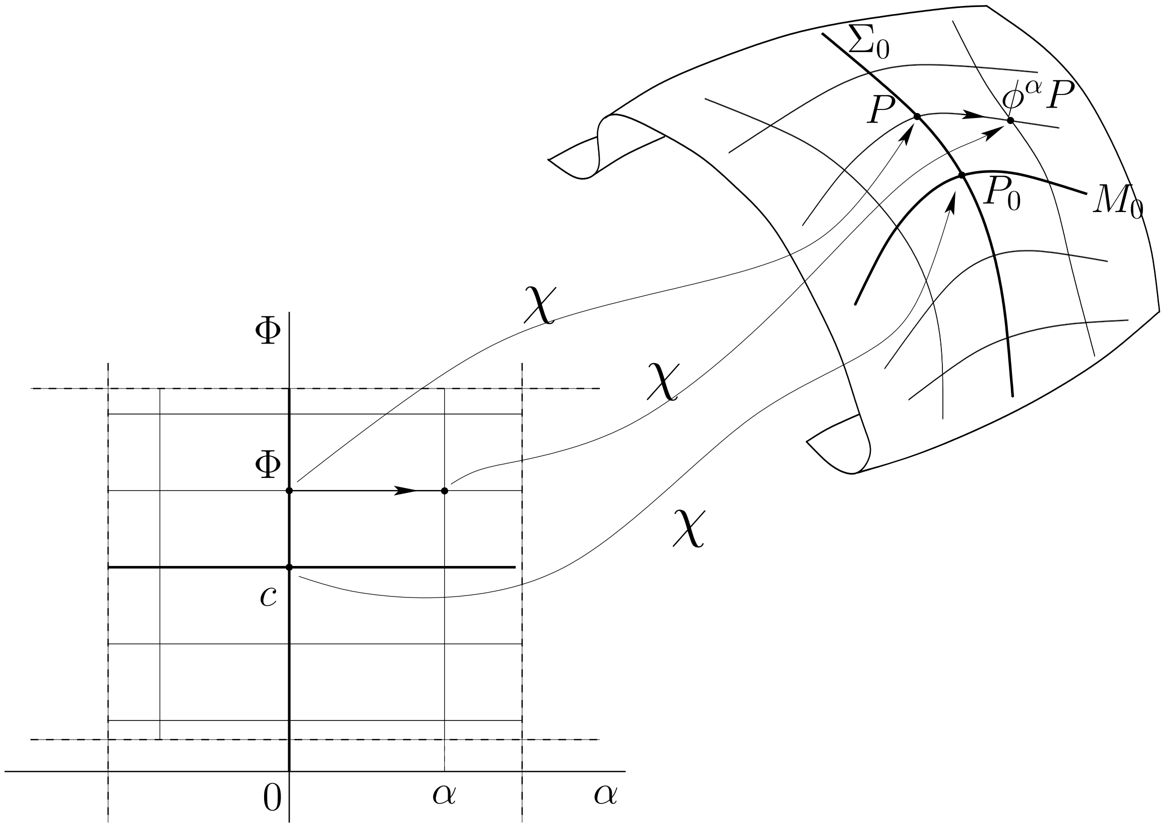

Figure 3.2. Illustrating the local coordinates induced by the flow of the involution system $\Phi_1(q,p),\dots,\Phi_n(q,p)$. For graphical reasons the notation $P_0 = (q_0,p_0)$ and $P=(q,p)$ is used in the figure.

3.1.5 Complete involution systems and coordinates induced by the flow

Proposition 3.5 turns out to be very useful when we have a complete involution system $\Phi_1(q,p),\dots,\Phi_n(q,p)$ on a phase space $\mathscr{F}$. Indeed this enables us to introduce *local* coordinates constructed through the canonical flows of the functions. This is illustrated in fig. 3.2. The crucial point is the following. Let $\phi_1^{\alpha_1},\dots,\phi_n^{\alpha_n}$ be the flows of $\Phi_1,\dots,\Phi_n$ up to time $\alpha_1,\dots,\alpha_n$, respectively. Apply the flows $\phi_1^{\alpha_1},\dots,\phi_n^{\alpha_n}$ to a point $(q,p)\in\mathscr{F}$ in any order: the result will always be the same, in view of the property that the flows do commute. Thus, we will simply denote

$$\phi^\alpha = \phi_1^{\alpha_1}\circ\ldots\circ\phi_n^{\alpha_n}$$

for α in some neighbourhood of the origin of $\mathbb{R}^n$.

Local coordinates can be constructed using the existence and of the commuting property of the flows, as illustrated in fig. 3.2. The claim is that the values of the functions $\Phi_1,\dots,\Phi_n$ and the times $\alpha_1,\dots,\alpha_n$ of the corresponding canonical flows define a local coordinate system.

The formal statement is given by[4]

Lemma 3.6: *Let $\Phi_1(q,p),\ldots,\Phi_n(q,p)$ be a complete involution system. Pick a point $(q_0,p_0)\in\mathscr{F}$ and let $\Phi_1(q_0,p_0)=c_1,\ldots,\Phi_n(q_0,p_0)=c_n$, with $c\in\mathbb{R}^n$. Then there exist a neigbourhood $V_\Phi\subset\mathbb{R}^n$ of c, a neighbourhood $V_\alpha\subset\mathbb{R}^n$ of the origin and a neighbourhood $U\subset\mathscr{F}$ of (q_0,p_0) such that the following holds true: there exists a diffeomorphism $\chi\,:\,V_\alpha\times V_\Phi\to U$ mapping $(\alpha,\Phi)\in V_\alpha\times V_\Phi$ to $(q,p)=\chi(\alpha,\Phi)\in U$ satisfying $\chi(0,c)=(q_0,p_0)$ and $\chi(\alpha,\Phi)=\phi^\alpha\chi(0,\Phi)$.*

`*** Dimostrazione troppo sintetica. Migliorare. ***`

Proof. Let

$$M_0=\bigl\{(q,p)\in\mathscr{F}\,:\,\Phi_1(q,p)=c_1,\ldots,\Phi_n\bigl(q,p)=c_n\bigr\}\ ,$$

so that $(q_0,p_0)\in M_0$. At the point (q_0,p_0) the Hamiltonian vector fields (in compact notation) $\mathsf{J}\partial_z\Phi_1,\ldots,\mathsf{J}\partial_z\Phi_n$ are independent, due to the independence of the functions Φ, and generate a plane $\mathrm{span}(\mathsf{J}\partial_z\Phi_1,\ldots,\mathsf{J}\partial_z\Phi_n)$ which is Lagrangian, in view of the Φ's being in involution. By lemma 2.7 this plane is complementary to at least one of the arithmetic planes of example 2.2 generated by the coordinates q,p. Let Π be one such plane. Then in a neighbourhood of (q_0,p_0) there is a n–dimensional manifold Σ_0 transversal to M_0 and tangent to Π in (q_0,p_0) parameterized by the coordinates $\Phi_1,\ldots,\Phi_n$. Denote by $(q,p)=\chi_0(\Phi)$ the map from a neighbourhood V_Φ of $\Phi=c$ to $\mathscr{F}$. We may always arrange that χ_0 is a differentiable map.[5] Let now the map $\chi\,:\,V_\alpha\times V_\Phi\to\mathscr{F}$ be defined as

$$\chi(0,\Phi)=\chi_0(\Phi)\ ,\quad \chi(\alpha,\Phi)=\phi^\alpha\chi(0,\Phi)$$

and let $U=\chi(V_\alpha\times V_\Phi)$. In view of the smoothness and differentiability with respect to parameters of the solutions of differential equations the map is a diffeomorphism between $V_\alpha\times V_\Phi$ and U. *Q.E.D.*

[4] A similar statement is easily made for a generic n–dimensional manifold where n independent and commuting vector fields are defined. The peculiar aspect of the lemma is that the Hamiltonian structure allows us to make an effective use of the Hamiltonian vector fields generated by a complete involution system, combining both the existence of invariant surfaces and the flow along the surfaces.

[5] An example may be useful. Suppose, e.g., that $\det\Bigl(\frac{\partial\Phi_j}{\partial p_k}\Bigr)\neq 0$; this means that the arithmetic Lagrangian plane tangent to the coordinate lines $p_1,\ldots,p_n$ is complementary to the manifold M_0 in $P_0=(q_0,p_0)$. By the implicit function theorem the relations $\Phi_1(q,p)=c_1,\ \ldots,\ \Phi_n=c_n(q,p)$ can be inverted in a neighbourhood of P_0, so as to give $p_1=p_1(\Phi,q),\ \ldots,\ p_n=p_n(\Phi,q)$. Set now $q=q_0$ and let $\Phi\in V_\Phi$. Then the functions $p_1=p_1(\Phi,q_0),\ \ldots,\ p_n=p_n(\Phi,q_0)$ determine the wanted local manifold Σ_0 with coordinates Φ. I.e., the map χ_0 is defined as $\chi_0(\Phi)=\bigl(q_0,p(\Phi,q_0)\bigr)$, and Σ_0 is the image on $\mathscr{F}$ of V_Φ. By construction the map χ_0 is differentiable. The choice of the plane Π is quite arbitrary, of course.

3.1.6 Complete involution systems and canonical transformations

The following proposition claims that the map χ of proposition 3.6 is a canonical transformation. Furthermore it states that the transformation can be constructed by quadratures.

Proposition 3.7: *Let $\{\Phi_1(q,p),\ldots,\Phi_n(q,p)\}$ be an involution system. Then there exists a local canonical transformation to new variables α,Φ*

$$q = q(\alpha,\Phi)\ ,\quad p = p(\alpha,\Phi)\ .$$

With the non restrictive hypothesis

$$\det\left(\frac{\partial(\Phi_1,\ldots,\Phi_n)}{\partial(p_1,\ldots,p_1)}\right) \neq 0\ . \tag{3.11}$$

the generating function of the canonical transformation is constructed by quadrature as

$$S(\Phi,q) = \int \sum_j p_j(\Phi,q)dq_j\ , \tag{3.12}$$

where $p_1(\Phi,q),\ldots,p_n(\Phi,q)$ are obtained by inversion of $\Phi_1(q,p),\ldots,\Phi_n(q,p)$.

Corollary 3.8: *The canonical coordinates of proposition 3.7 are determined up to a canonical transformation with generating function*

$$W(\overline{\Phi},\alpha) = \sum_j \overline{\Phi}_j\alpha_j + f(\overline{\Phi})\ , \tag{3.13}$$

where $f(\overline{\Phi})$ is an arbitrary function. Equivalently, one can add an arbitrary function $f(\Phi)$ to the generating function $S(\Phi,q)$ defined by (3.12).

By going back to the proof of proposition 3.6 we may realize that the arbitrary choice of the function $f(\Phi)$ corresponds to the arbitrary choice of the surface Σ_0 corresponding to $\alpha = 0$.

The proof of the corollary is trivial, and is left to the reader.

Corollary 3.9: *The coordinates $\alpha_1,\ldots,\alpha_n$ of proposition 3.7 coincide with the coordinates $\alpha_1,\ldots,\alpha_n$ of proposition 3.6, up to a Φ depending translation as in corollary 3.8.*

Proof. In the new canonical coordinates α,Φ the canonical equations for Φ_k are $\dot\alpha_j = \delta_{j,k}$ and $\dot\Phi_j = 0$, with $j = 1,\ldots,n$. Thus α_k is the time of the flow generated by Φ_k, for $k = 1,\ldots,n$. The translation comes fron the arbitrary choice of the n–dimensional menifold corresponding to $\alpha = 0$. *Q.E.D.*

Proof of proposition 3.7. In view of lemma 3.2 condition (3.11) is not restrictive, since it can always be fulfilled by exchanging some of the coordinates with the conjugated momenta. In view of that condition, we can invert the functions $\Phi_1,\ldots,\Phi_n$ with respect to $p_1,\ldots,p_n$, thus getting $p_1 = p_1(\Phi,q),\ldots,p_n = p_n(\Phi,q)$. Consider the differential form $\sum_j p_j(\Phi,q)dq_j$; we prove that it is exact. To this end, let us differentiate the identity

$$\Phi_j = \Phi_j(q,p)\big|_{p=p(\Phi,q)}$$

with respect to Φ, q, namely taking into account that in the r.h.s. p must be replaced by its expression in terms of Φ, q. This gives

$$d\Phi_j = \sum_{k,l} \frac{\partial \Phi_j}{\partial p_k}\left(\frac{\partial p_k}{\partial \Phi_l} d\Phi_l + \frac{\partial p_k}{\partial q_l} dq_l\right) + \sum_l \frac{\partial \Phi_j}{\partial q_l} dq_l \ .$$

By comparison of the coefficients of dq, $d\Phi$ we get the identities

$$\begin{aligned} &\sum_k \frac{\partial \Phi_j}{\partial p_k}\frac{\partial p_k}{\partial \Phi_l} = \delta_{j,l} \ , \\ &\sum_k \frac{\partial \Phi_j}{\partial p_k}\frac{\partial p_k}{\partial q_l} = -\frac{\partial \Phi_j}{\partial q_l} \ , \quad j,l = 1,\ldots,n \ . \end{aligned}$$

Replace now the second of these identities in the relation $\{\Phi_j, \Phi_m\} = 0$, which holds true because the functions are assumed to be in involution. With a few calculations we get

$$\begin{aligned} \{\Phi_j, \Phi_m\} &= \sum_l \left(\frac{\partial \Phi_j}{\partial q_l}\frac{\partial \Phi_m}{\partial p_l} - \frac{\partial \Phi_j}{\partial p_l}\frac{\partial \Phi_m}{\partial q_l}\right) \\ &= -\sum_{l,k} \frac{\partial \Phi_m}{\partial p_l}\frac{\partial \Phi_j}{\partial p_k}\frac{\partial p_k}{\partial q_l} + \sum_{l,k} \frac{\partial \Phi_j}{\partial p_l}\frac{\partial \Phi_m}{\partial p_k}\frac{\partial p_k}{\partial q_l} \\ &= -\sum_l \frac{\partial \Phi_m}{\partial p_l} \sum_k \frac{\partial \Phi_j}{\partial p_k}\left(\frac{\partial p_k}{\partial q_l} - \frac{\partial p_l}{\partial q_k}\right) = 0 \end{aligned}$$

(note that in the second sum on the second line the indexes l and k can be exchanged). By condition (3.11) this implies

$$\frac{\partial p_k}{\partial q_l} - \frac{\partial p_l}{\partial q_k} = 0 \ , \quad l,k = 1,\ldots,n \ ,$$

so that the differential form $\sum_j p_j dq_j$ is exact, as claimed. By integration we construct the generating function (3.12) which, in view of (3.11), satisfies the invertibility condition (2.29) of proposition 2.19. Therefore, the wanted canonical transformation is implicitly defined by

$$\alpha_j = \frac{\partial S}{\partial \Phi_j} \ , \quad p_j = \frac{\partial S}{\partial q_j} \ , \quad j = 1,\ldots,n \ .$$

Q.E.D.

3.2 The theorem of Liouville

For a generic system of differential equations on a n–dimensional manifold a complete integration by quadrature can be performed when $n-1$ independent first integrals are known, n being the dimension of the space. Thus, one expects that in the Hamiltonian case, the dimension of phase space being $2n$, one needs to know $2n-1$ first integrals. However, the canonical structure allows us to perform the complete integration if only n first integrals are known, provided they fulfill the further condition of being in involution.

Theorem 3.10: *Assume that an autonomous canonical system with n degrees of freedom and with Hamiltonian $H(q,p)$ possesses n independent first integral $\{\Phi_1(q,p),\ldots,\Phi_n(q,p)\}$ forming a complete involution system. Then the system is integrable by quadratures. More precisely, one can construct the generating function $S(\Phi,q)$ of a canonical transformation $(q,p)=\chi(\alpha,\Phi)$ such that the transformed Hamiltonian depends only on the new momenta $\Phi_1,\ldots,\Phi_n$, and the solutions are expressed as*

$$\alpha_j(t) = \alpha_{j,0} + t\frac{\partial H}{\partial \Phi_j}\bigg|_{(\Phi_{1,0},\ldots,\Phi_{n,0})} , \quad j=1,\ldots,n ,$$

with $\alpha_{j,0}$ and $\Phi_{j,0}$ determined by the initial data.

3.2.1 Proof of Liouville's theorem

By proposition 3.7 we can construct by quadratures a canonical transformation $(q,p)=\chi(\alpha,\Phi)$ such that $\Phi_1,\ldots,\Phi_n$ are the new momenta. In view of preservation of Poisson brackets, we can compute the Poisson bracket $\{H,\Phi_j\}$ with respect to the new variables α,Φ. Since $\Phi_1,\ldots,\Phi_n$ are first integrals, this gives

$$\{H,\Phi_j\} = \frac{\partial H}{\partial \alpha_j} = 0 , \quad j=1,\ldots,n .$$

This means that the transformed Hamiltonian depends only on the momenta, i.e., $H=H(\Phi)$. Therefore, the canonical equations are

$$\dot\alpha_j = \frac{\partial H}{\partial \Phi_j} , \quad \dot\Phi_j = 0 , \quad j=1,\ldots,n ,$$

and are trivially integrable, as stated. This concludes the proof.

3.2.2 Integration procedure

I emphasize that Liouville's theorem actually furnishes an explicit integration algorithm. Here is the procedure.

(i) If necessary, exchange some pairs of canonical variables so that the condition

$$\det\left(\frac{\partial(\Phi_1,\ldots,\Phi_n)}{\partial(p_1,\ldots,p_n)}\right) \neq 0$$

is fulfilled. Then perform an inversion, finding $p=p(\Phi,q)$.

(ii) By a quadrature, construct the generating function

$$S(\Phi,q) = \int \sum_j p_j(\Phi,q)dq_j .$$

(iii) By substitution, determine the transformed Hamiltonian $H(\Phi)$.

(iv) The solutions of the canonical equations are

$$(3.14)\qquad \Phi_j(t) = \Phi_{j,0}\ ,\quad \alpha_j(t) = \frac{\partial H}{\partial \Phi_j}\Big|_{\Phi_j=\Phi_{j,0}} t + \alpha_{j,0}\ ,\quad j = 1,\ldots,n\ ,$$

where Φ_{j_0} and $\alpha_{j,0}$ are the initial values that can be computed from the initial data.

(v) By inversion of the canonical transformation find $q = q(\Phi,\alpha)$ and $p = p(\Phi,\alpha)$.

(vi) The solutions $q(t), p(t)$ in the original variables are found by substitution of $\Phi(t), \alpha(t)$ given by (3.14).

Example 3.4: *Systems with one degree of freedom.* Let us consider the Hamiltonian

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

describing the motion of a mass point on a straight line under the action of the potential $V(x)$. The condition at point (i) reduces to $\frac{\partial H}{\partial p} \neq 0$, which is fulfilled for $p \neq 0$. Setting $H(x,p) = E$, we invert the relation above with respect to p, getting

$$(3.15)\qquad p = \pm\sqrt{2m[E - V(x)]}\ .$$

The generating function is

$$S(E,x) = \sqrt{2m}\int \sqrt{E - V(x)}dx\ ,$$

and the canonical transformation in implicit form is written as

$$p = \pm\sqrt{2m[E - V(x)]}\ ,\quad \alpha = \sqrt{\frac{m}{2}}\int \frac{dx}{\sqrt{E - V(x)}}\ .$$

The first of these relations coincides with (3.15), as expected. The transformed Hamiltonian is, trivially, $H(E) = E$, and the solutions of Hamilton's equations in coordinates α, E are

$$E(t) = E_0\ ,\quad \alpha(t) = t - t_0\ ,$$

E_0 and t_0 being the initial values of energy and time, respectively. Therefore, in order to actually compute the solutions we need to compute the integral

$$(3.16)\qquad t - t_0 = \sqrt{\frac{m}{2}}\int_{x_0}^{x} \frac{d\xi}{\sqrt{E_0 - V(\xi)}}\ ,$$

where $x_0 = x(0)$ is the initial datum and E_0 is the initial energy. The latter formula actually coincides with (1.36), of sect. 1.3.1, as should be expected.

Example 3.5: *Harmonic oscillators.* For the Hamiltonian

$$(3.17)\qquad H = \frac{1}{2}\sum_{l=1}^{n}(y_l^2 + \omega_l^2 x_l^2)$$

the quantities

$$\Phi_l = \frac{1}{2}(y_l^2 + \omega_l^2 x_l^2) \ , \quad 1 \le l \le n \tag{3.18}$$

form an involution system, and moreover we have

$$H = \sum_l \Phi_l \ . \tag{3.19}$$

By inversion of (3.18) with respect to y_l we get

$$y_l = \sqrt{2\Phi_l - \omega_l^2 x_l^2} \ ,$$

and the generating function is

$$S(\Phi, x) = \sum_{l=1}^{n} F_l(\Phi_l, x_l) \ ,$$

where

$$F_l(\Phi_l, x_l) = \int \sqrt{2\Phi_l - \omega_l^2 x_l^2} dx_l \ , \quad 1 \le l \le n \ .$$

The canonical transformation is completed by the new coordinates

$$\alpha_l = \frac{\partial S}{\partial \Phi_l} = \int \frac{dx_l}{\sqrt{2\Phi_l - \omega_l^2 x_l^2}} = \frac{1}{\omega_l} \mathrm{arcos} \left(\frac{\omega_l x_l}{\sqrt{2\Phi_l}} \right) \ .$$

The Hamiltonian is given by (3.19), and the canonical equations

$$\dot{\Phi}_l = 0 \ , \quad \dot{\alpha} = \frac{\partial H}{\partial \Phi_l} = 1$$

have solutions

$$\Phi_l(t) = \Phi_{l,0} \ , \quad \alpha(t) = t - t_0 \ ,$$

where t_0 is the initial time, and $\Phi_{l,0}$ are constants to be computed by the initial data. Finally, by inversion we obtain the solution

$$x_l = \frac{\sqrt{2\Phi_{l_0}}}{\omega_l} \cos \omega_l (t - t_0) \ .$$

Example 3.6: *Motion under central field on a plane.* As a further example, let us consider the Hamiltonian

$$H = \frac{1}{2m} \left(p_r^2 + \frac{p_\vartheta^2}{r^2} \right) + V(r) \ ,$$

describing the planar motion of a mass point in a central force field. As first integrals we can use

$$\Gamma = p_\vartheta \ , \quad \Phi = \frac{1}{2m} \left(p_r^2 + \frac{\Gamma^2}{r^2} \right) + V(r) \ . \tag{3.20}$$

By inversion, compute

$$p_\vartheta = \Gamma\ ,\quad p_r = \left[2m\bigl(\Phi - V(r)\bigr) - \frac{\Gamma^2}{r^2}\right]^{1/2}\ , \tag{3.21}$$

so that the generating function is

$$S(\Phi,\Gamma,r,\vartheta) = \int \left[2m\bigl(\Phi - V(r)\bigr) - \frac{\Gamma^2}{r^2}\right]^{1/2} dr + \int \Gamma d\vartheta\ .$$

Denoting by φ and γ the canonical variables conjugated to Φ and Γ, respectively, the transformation in implicit form is given by (3.21) and

$$\begin{aligned}
\varphi &= \frac{\partial S}{\partial \Phi} = m \int \left[2m\bigl(\Phi - V(r)\bigr) - \frac{\Gamma^2}{r^2}\right]^{-1/2} dr\\
\gamma &= \frac{\partial S}{\partial \Gamma} = -m\Gamma \int \frac{1}{r^2}\left[2m\bigl(\Phi - V(r)\bigr) - \frac{\Gamma^2}{r^2}\right]^{-1/2} dr + \int d\vartheta\ .
\end{aligned}$$

Once $V(r)$ is known, by quadrature we can compute

$$\varphi = f(\Phi,\Gamma,r)\ ,\quad \gamma = g(\Phi,\Gamma,r) + \vartheta - \vartheta_0\ , \tag{3.22}$$

where ϑ_0 is given by the initial conditions, and the functions f and g are given by the integrals in the formula above. The transformed Hamiltonian is $H = \Phi$, so that the solution of Hamilton's equations are

$$\begin{aligned}
\varphi &= t - t_0\ ,\quad &\gamma &= \gamma_0\ ,\\
\Phi &= \Phi_0\ , &\Gamma &= \Gamma_0\ .
\end{aligned} \tag{3.23}$$

Here, γ_0, Φ_0 and Γ_0 must be computed from the initial data $r_0, \vartheta_0, p_{r,0}, p_{\vartheta,0}$ at time t_0 using (3.20) and (3.22). The solutions $r(t), \vartheta(t), p_r(t), p_\vartheta(t)$ in the original coordinates are computed by inverting (3.22) so as to obtain

$$r = r(\Phi,\Gamma,\varphi)\ ,\quad \vartheta = \vartheta_0 + \gamma - g(\Phi,\Gamma,r)\Big|_{r=r(\Phi,\Gamma,\varphi)}\ .$$

By substitution of (3.23) in the latter expressions we get r and ϑ as functions of time and of the initial values, namely

$$r = r(\Phi_0,\Gamma_0,t-t_0)\ ,\quad \vartheta = \vartheta_0 + \gamma_0 - g(\Phi_0,\Gamma_0,r)\Big|_{r=r(\Phi_0,\Gamma_0,t-t_0)}\ .$$

Finally, the momenta p_r and p_ϑ as functions of time are computed by replacing (3.23) and the latter expressions in (3.21). This completes the solution of the problem.

Exercise 3.1: Apply Liouville's theorem to the case of a central field of forces in space, reducing it to quadrature.

3.2.3 Some comments on Liouville's theorem

The form (3.14) of the solutions for the Hamiltonian in the new variables appears to be quite simple. However, the example above of the harmonic oscillators shows that all phenomena related to the periodicity of the motion are hidden, and show up only when the transformation back to the original variables is performed.

For comparison, if we apply to the Hamiltonian (3.17) the transformation to action–angle variables for the harmonic oscillators, namely

$$x_l = \sqrt{2I_l}\cos\varphi_l \ , \quad y_l = \sqrt{2I_l}\sin\varphi_l \ , \quad l = 1,\ldots,n \ ,$$

we get the transformed Hamiltonian

$$H(I_1,\ldots,I_n) = \sum_l \omega_l I_l \ .$$

The canonical equations then are

$$\dot\varphi_l = \omega_l \ , \quad \dot I_l = 0 \ ,$$

and the evolution of the phases $\varphi_l(t) = \omega_l t + \varphi_{l,0}$ is still uniform with velocity ω. The remarkable difference with respect to (3.14) is that the new coordinates φ are *angles* representing the phases of the oscillators. Therefore, the periodic character of the evolution is evident, and the velocity ω of the phases is the angular frequency.

The same problem shows up also in the discussion of example 3.6: the algorithm is well defined, but the fact that, e.g., in the Keplerian case the motion is periodic will not be recognized until the complete solution is explicitly calculated.

That this is a general problem is further illustrated by the following

Example 3.7: *Free rotator.* Let the phase space be $\mathbb{T}\times\mathbb{R}$, with coordinates $q\in\mathbb{T}$ and $p\in\mathbb{R}$, and let the Hamiltonian be

$$H = \frac{p^2}{2} \ . \tag{3.24}$$

The system is actually trivial: the equations are $\dot q = p$, $\dot p = 0$, with solutions $p(t) = p_0$, $q(t) = p_0 t + q_0$, where p_0 , q_0 are the initial data. Recalling that q is an angle, the motion is immediately seen to be periodic with angular frequency p_0 depending on the initial data. If, however, we forget this fact, and apply the procedure suggested by Liouville's theorem, using the Hamiltonian H as a first integral, then we get that the new canonical coordinates are the time t, which flows uniformly, and the energy H, which is constant. We see again that all informations concerning the periodicity of the motion are lost, and are recovered only after writing the solutions for the original variables.

The common aspect to all these examples is that the choice of the first integrals to be used in order to apply Liouville's procedure is quite arbitrary. However, the examples of the harmonic oscillator and of the free rotator suggest that there should be a particular choice which is the best one, and that it is connected with the fact that the coordinates conjugated to the momenta $\Phi_1,\ldots,\Phi_n$ should be *angles*.

3.2.4 Action–angle variables for systems with one degree of freedom

The construction of action–angle variables turns out to be particularly simple but instructive in the case of a system with one degree of freedom. This is always a Liouville–integrable system, since the Hamiltonian is a first integral. Moreover, the orbits of the system are implicitly defined by the equation $H(q,p) = E$. Let $\overline{q}, \overline{p}$ be an extremum for $H(q,p)$, so that $q(t) = \overline{q}, p(t) = \overline{p}$ is a solution of Hamilton's equations. Then there is an open interval $\mathscr{E}$ such that for $E \in \mathscr{E}$ the set of point satisfying $H(q,p) = E$ contains a continuous family of closed curves surrounding the point $(\overline{q}, \overline{p})$. Let γ_E be one such curve; it can be described via a coordinate $\varphi \in \mathbb{T}$, in many ways. With such a coordinate it is easy to account for the periodicity of the motion, since a period is completed when φ is incremented by 2π. It is quite natural to ask if there exists a canonical momentum I, conjugated to the coordinate φ, which parameterizes the family γ_E of closed curves.

If such a quantity I exists, it must be constant on every curve γ_E; this implies that it must be a first integral. Thus, let us look for a function $I(q,p)$ which is in involution with the Hamiltonian $H(q,p)$ and satisfies $\frac{\partial I}{\partial p} \neq 0$. If such a function exists, by proposition 3.7 we can construct a further function

$$S(I,q) = \int p(I,q)dq \ , \tag{3.25}$$

the latter being the generating function of a canonical transformation which defines I as the new momentum. Thus, there is also a coordinate, that we denote again by φ, conjugated to I; on the other hand, φ must be periodic, because it is a coordinate on a closed curve, and we can always manage so that the period is 2π. Let us now see how we can construct $I(q,p)$. If φ, I are canonical variables, by proposition 2.17 we must have

$$\oint_{\gamma_E} p dq = \oint_{\gamma_E} I d\varphi \ .$$

Since $I(q,p)$ must be constant on γ_E and φ is periodic, the integral on the right hand side is easily calculated to be $2\pi I$. Therefore, it must be

$$I = \frac{1}{2\pi} \oint_{\gamma_E} p dq \ . \tag{3.26}$$

This quantity has been named the *action* of the system. The integral must be computed after expressing p as a function of E and q, and gives a function $I(E)$; replacing $E = H(q,p)$ gives I as a function of q,p, as required. We conclude that $I(q,p)$ can be computed through a quadrature. It will be noticed that this function represents the area enclosed by the curve γ_E passing through the point q,p divided by 2π.

Since $I(q,p)$ is a first integral, we can apply the theorem of Liouville. With a further quadrature we can compute the generating function $S(I,q)$ given by (3.25), and by differentiation we determine the angle φ. The canonical coordinates I, φ are called *action–angle variables.* The transformed Hamiltonian $H(I)$ is independent of φ, and Hamilton's equations read

$$\dot{I} = 0 \ , \quad \dot{\varphi} = \omega(I) \ , \tag{3.27}$$

where $\omega(I) = \frac{\partial H}{\partial I}$. Having fixed the initial conditions $I(0) = I_0, \varphi(0) = \varphi_0$ the corresponding solution is

$$I(t) = I_0 \ , \quad \varphi(t) = \omega_0 t + \varphi^{(0)} \ , \tag{3.28}$$

where $\omega_0 = \omega(I_0)$. The periodicity of the motion is now evident, because φ is an angle, and the period clearly is $T = 2\pi/\omega_0$. The period can be easily computed as

$$T = 2\pi \frac{dI}{dE} \ . \tag{3.29}$$

For, from $H = E$ one has $\frac{dH}{dE} = \frac{dH}{dI}\frac{dI}{dE} = \omega(I)\frac{dI}{dE} = 1$, and the claim follows using $T = 2\pi/\omega(I)$.

Example 3.8: *The harmonic oscillator.* The Hamiltonian

$$H(p, x) = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 x^2$$

has an equilibrium for $p = x = 0$, and for $E > 0$ the curve $H(p, x) = E$ is an ellipse centered on the origin and with semi axes $\sqrt{2E}/\omega$ and $\sqrt{2E}$ (see example 1.9). The action is easily computed as $I = E/\omega$, the period is $T = 2\pi/\omega$, and the angle φ represents the phase of the oscillator. It is not necessary to proceed to an explicit calculation of the generating function, since the explicit form of the canonical transformation $x = \sqrt{2I}\cos\varphi$, $y = \sqrt{2I}\sin\varphi$ can be obtained by elementary considerations.

Example 3.9: *Oscillations around a stable equilibrium.* Consider the Hamiltonian

$$H(p, x) = \frac{p^2}{2} + V(x) \ , \quad (x, p) \in \mathbb{R}^2 \tag{3.30}$$

describing the motion of a point with unitary mass moving on a straight line under the action of a potential $V(x)$. Let $V(x)$ have a point of relative minimum at $x = 0$, and let $V(0) = 0$ (this is not restrictive, of course, since it is always possible to introduce the displacement from equilibrium as a coordinate, and the potential is defined up to a constant). Then there exists an open interval of positive values of E such that the level set of points satisfying $H(p, x) = E$ contains a closed curve around the origin. For such values of E the action is defined as

$$I = \frac{\sqrt{2}}{\pi}\int_{x_{\min}}^{x_{\max}} \sqrt{E - V(x)}dx \ , \tag{3.31}$$

$x_{\min}, x_{\max}$ being the extrema of the interval of oscillation, to be computed as solutions of the equation $E - V(x) = 0$. The period can be computed as

$$T = \sqrt{2}\int_{x_{\min}}^{x_{\max}} \frac{\mathrm{d}x}{\sqrt{E - V(x)}} \ . \tag{3.32}$$

The calculation of the action and of the period is therefore reduced to a quadrature. If an explicit expression for the angle φ is wanted, then one must compute the generating function

$$S(I, x) = \sqrt{2}\int \sqrt{E - V(x)}dx$$

and replace $E = E(I)$ as computed from (3.31). The canonical transformation is written in implicit form as

$$p = \frac{\partial S}{\partial x} \ , \quad \varphi = \frac{\partial S}{\partial I} \ .$$

Writing the transformation explicitly requires an inversion.

Example 3.10: *The pendulum.* As a more specific example, consider the Hamiltonian

$$H(p, \vartheta) = \frac{p^2}{2} - \cos\vartheta \ , \quad (\vartheta, p) \in \mathbb{T} \times \mathbb{R} \ ,$$

which describes the frictionless motion of a unit mass point constrained to a circle in the vertical plane, subject to gravity. The Hamiltonian has a minimum for $p = \vartheta = 0$ (recall that ϑ is defined $\bmod 2\pi$), and for $-1 < C < 1$ the equation $H(p, \vartheta) = C$ determines a closed curve around the equilibrium. According to (3.31) the action variable is computed as

$$I = \frac{2\sqrt{2}}{\pi} \int_0^{\vartheta_{\max}} \sqrt{C + \cos\vartheta}\, d\vartheta \ ,$$

where $\vartheta_{\max} = \arccos(-C)$. Remark that the symmetry $V(-\vartheta) = V(\vartheta)$ of the potential has been taken into account in the latter formula. The calculation of the integral can be reduced to that of elliptic integrals as follows. Transform $\vartheta = 2\varphi$, thus getting

$$I = \frac{8}{\pi} \int_0^{\varphi_{\max}} \sqrt{\frac{C+1}{2} - \sin^2\varphi}\, d\varphi \ .$$

Denote now $k^2 = (C+1)/2$, and perform the further transformation $\sin\varphi = k \sin\psi$; this gives

$$\begin{aligned} I &= \frac{8k^2}{\pi} \int_0^{\pi/2} \frac{\cos^2\psi}{\sqrt{1 - k^2 \sin^2\psi}}\, d\psi \\ &= \frac{16(k^2+1)}{\pi} K(k^2) + \frac{16}{\pi} E(k^2) \ , \end{aligned}$$

where $K(k^2)$ and $E(k^2)$ are the complete elliptic integrals of first and second kind, namely

$$\begin{aligned} K(k^2) &= \int_0^{\pi/2} \frac{d\psi}{\sqrt{1 - k^2 \sin^2\psi}} \\ E(k^2) &= \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2\psi}\, d\psi \ . \end{aligned}$$

These integrals are computed via a series expansion which is convergent for $|k|^2 < 1$. Therefore, the action I as a function of C is defined only for $C < 1$. This value corresponds indeed to the separatrix. The transformed Hamiltonian $H(I)$ can be computed by inversion of the function given by the integral, putting $H = C$. The frequency corresponding to a given initial value of the action is computed as $\omega(I) = \frac{\partial H}{\partial I}$.

Exercise 3.2: Show how to calculate the action for the rotating pendulum, i.e, when $|C| > 1$.

3.3 On manifolds with nonsingular vector fields

Let us now start investigating the global geometric character of the manifolds defined by a complete involution system. In this section I will pay attention to a single connected manifold implicitly defined as

$$M_0 = \{(q,p) \in \mathscr{F} \ : \ \Phi_1(q,p) = 0, \ldots, \Phi_n(q,p) = 0\} \ . \tag{3.33}$$

If more than one connected manifold is defined in the formula above, just select one. The value 0 might be replaced by any suitable value $c \in \mathbb{R}^n$, of course.

The independence of the functions $\Phi_1, \ldots, \Phi_n$ implies that the corresponding Hamiltonian vector fields are independent at every point. Moreover, in view of the involution property, the vector fields are tangent to M_0 which is a n–dimensional Lagrangian manifold. The relevant property here is the existence of n independent vector fields at every point, since it puts severe constraints on the topology of the manifold.

I state the result in general terms in view of its independent interest and generality.

Proposition 3.11: *Let M be a connected differentiable manifold of dimension n, and let $X_1, \ldots, X_n$ be n vector fields which are independent at every point of M and satisfy*[6] $[X_j, X_k] = 0$ *on M, for $j,k = 1,\ldots,n$; assume moreover that the flows of $X_1, \ldots, X_n$ can be indefinitely continued on M. Then there is a non negative $k \le n$ such that M is diffeomorphic to $\mathbb{T}^k \times \mathbb{R}^{n-k}$.*

Corollary 3.12: *If M is compact, then it is diffeomorphic to a n–dimensional torus.*

The proof requires several technical steps, which are separately worked out in the next subsections.

3.3.1 Local coordinates induced by the flow

The first step for the proof of proposition 3.11 is the fact that the n independent and commuting flows generate a n–parameter group acting on M. As a local property, the group defines a local coordinate system in the neighbourhood of any point $P \in M$; this is essentially a restriction to M of the statement of proposition 3.6. The global property is that every two points can be connected by the action of the group, so that the image of $\mathbb{R}^n$ by the group covers M .

Lemma 3.13: *Under the hypotheses of proposition 3.11 there exists a n–parameter group of diffeomorphisms $\phi^t \ : \ \mathbb{R}^n \times M \to M$ with the following properties:*

(i) for every $t \in \mathbb{R}^n$ and every $P \in M$ there is a unique $\phi^t P \in M$;

(ii) for every $P \in M$ there are a neighbourhood V of the origin of $\mathbb{R}^n$ and a neighbourhood U of P which are diffeomorphic;

(iii) for every pair P, Q of points of M there is $t \in \mathbb{R}^n$ such that $\phi^t P = Q$. Here, t needs not be unique.

---

[6] The symbol $[X_j, X_k]$ denotes the commutator between the vector fields X_j and X_k . If the vector fields are Hamiltonians and are generated by $F(q,p)$ and $G(q,p)$, say, then the commutator is the vector field generated by $\{F, G\}$.

Proof. (i) Let $t \in \mathbb{R}^n$, and define $\phi^t = \phi^{t_n}_{X_n} \circ \ldots \circ \phi^{t_1}_{X_1}$. For every point $P \in M$ the point $\phi^t P$ is defined, because the flows can be indefinitely prolonged, and is of course unique; in particular, $\phi^0 P = P$. Since the flows do commute, the result is independent of the order of application of the flows $\phi^{t_j}_{X_j}$, so that it depends only on $t \in \mathbb{R}^n$. On the other hand, using the commutativity and the group property of each flow, we have $\phi^t \circ \phi^s = \phi^{t+s}$. We conclude that ϕ^t is a group. Since each vector field X_j defines a differentiable flow $\phi^{t_j}_{X_j}$, the composition ϕ^t is differentiable; that is, ϕ^t is a diffeomorphism.
(ii) Let $x_1(P), \ldots, x_n(P)$ be local coordinates in the neighbourhood of any given point $P \in M$. The n–parameter group of point (i) defines a differentiable map

$$\bigl(x_1(t), \ldots, x_n(t)\bigr) = \bigl(x_1(\phi^t P), \ldots, x_n(\phi^t P)\bigr) .$$

At the origin $t = 0$ the rows of the Jacobian matrix of the map

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(t_1, \ldots, t_n)}$$

are the vector fields $X_1, \ldots, X_n$ evaluated at the point P. Therefore, the Jacobian determinant is non zero at $t = 0$ in view of the independence of the vector fields, and by continuity is non zero in a neigbourhood of $t = 0$. We conclude that the map is a local diffeomorphism, as claimed.
(iii) Since M is connected, there is a curve γ connecting P with Q. By (ii), to every point of $P' \in \gamma$ we can associate a neighbourhood $U(P')$ which is diffeomorphic to a neighbourhood of the origin of $\mathbb{R}^n$. This family is a covering of γ. Since γ is compact, we can extract a finite sequence, $P_0, \ldots, P_m$ say, of points of γ such that $P_0 = P$, $P_m = Q$ and $P_j \in U(P_{j-1})$, the neighbourhood of P_{j-1} diffeomorphic to a neighbourhood of the origin of $\mathbb{R}^n$. Therefore, there exists a finite sequence $t_1, \ldots, t_m$ such that $P_j = \phi^{t_j} P_{j-1}$ for $j = 1, \ldots, m$. By the group property we conclude $\phi^t P = Q$ with $t = t_1 + \ldots + t_m$. *Q.E.D.*

3.3.2 The stationary group

Let us now investigate the global properties of the n–parameters group ϕ^t of diffeomorphism, the existence of which has been stated in the previous section. Accordig to lemma 3.13 we may regard the flows ϕ^t as a differentiable map from $\mathbb{R}^n \times M$ to M the image of which is the whole manifold M. However, this is not a diffeomorphism. For, it may happen that $\phi^t P = P$ for some $t \in \mathbb{R}^n$ and for some $P \in M$. The relevant fact is that if $\phi^t P = P$ then we also have $\phi^t Q = Q$ for all $Q \in M$. For, by (iii) of lemma 3.13 we have $Q = \phi^s P$ for some $s \in \mathbb{R}^n$; on the other hand, by the group property we have $\phi^t Q = \phi^t \circ \phi^s P = \phi^{s+t} P = \phi^s \circ \phi^t P = \phi^s P = Q$.

The *stationary set* of ϕ^t is defined as

$$G = \{t \in \mathbb{R}^n \mid \phi^t P = P \quad \text{for all } P \in M\} . \tag{3.34}$$

Lemma 3.14: *The stationary set G defined by (3.34) is a non empty discrete subgroup of $\mathbb{R}^n$; that is: it possesses the group property, and contains no accumulation points.*

Proof. The set G is not empty, since it contains at least the origin of $\mathbb{R}^n$.
If $t \in G$ then $\phi^{-t}P = \phi^{-t} \circ \phi^t P = P$, so that $-t \in G$. If $t_1, t_2 \in G$, then we have $\phi^{t_1+t_2}P = \phi^{t_1} \circ \phi^{t_2}P = \phi^{t_1}P = P$, so that $t_1 + t_2 \in G$. We conclude that G has the group property.
The origin is an isolated point of G, because in a neighbourhood V_0 of the origin ϕ^t is a diffeomorphism, which implies that $G \cap V_0 = \{0\}$. Let now $0 \neq t \in G$, and consider the neighbourhood $V_t = t + V_0 = \{t + s \mid s \in V_0\}$. Let $t' \in G \cap V_t$, i.e., $\phi^{t'}P = P$. By the group property, we also have $\phi^{t'-t}P = P$, i.e., $t' - t \in G \cap V_0$, which implies $t' - t = 0$. This proves that t is an isolated point. *Q.E.D.*

3.3.3 Angular coordinates

Having determined the k–dimensional stationary group G, we proceed now to the construction of a diffeomorphism between M and $\mathbb{T}^k \times \mathbb{R}^{n-k}$. To this end we need a statement of algebraic character, namely that a discrete subgroup G of $\mathbb{R}^n$ has a basis. That is, there exist k independent vectors $e_1, \ldots, e_k$ in G, where k is the dimension of G, such that

$$G = \{m_1 e_1 + \ldots + m_k e_k \,:\, (m_1, \ldots, m_k) \in \mathbb{Z}^k\} \ .$$

The proof of this claim is found in Appendix A.

Proof of proposition 3.11. Take any basis $\{e_1, \ldots, e_k\}$ of G, and complete it with $n-k$ independent vectors $u_1, \ldots, u_{n-k}$ so that $\{e_1, \ldots, e_k, u_1, \ldots, u_{n-k},\}$ is a basis of $\mathbb{R}^n$. Writing an arbitrary point $t \in \mathbb{R}^n$ as

$$t = \frac{\psi_1}{2\pi}e_1 + \ldots + \frac{\psi_k}{2\pi}e_k + \tau_1 u_1 + \ldots + \tau_{n-k}u_k \ ,$$

we establish a global one-to-one correspondence between $\mathbb{T}^k \times \mathbb{R}^{n-k}$ and M by taking an arbitrary point $P \in M$ as corresponding to the origin and setting $P(\psi_1, \ldots, \psi_k, \tau_1, \ldots, \tau_{n-k}) = \phi^t P$, with t given above. Since in the neighbourhood of any point $Q \in M$ this mapping if a local diffeomorphism we conclude that M is diffeomorphic to $\mathbb{T}^k \times \mathbb{R}^{n-k}$. *Q.E.D.*

Proof of corollary 3.12. By proposition 3.11 M is diffeomorphic to $\mathbb{T}^k \times \mathbb{R}^{n-k}$ for some k. Since M is compact, then $k = n$. *Q.E.D.*

3.4 Action–Angle variables

The construction of action–angle variables in a system with one degree of freedom makes essential use of the existence of a family of closed curves. The action variable is a particular parameterization of this family, which turns out to be particularly interesting because the conjugated variable φ is an angle.

Action–angle variables for systems with $n > 1$ degrees of freedom may be introduced via a deep extension of the argument used for $n = 1$.

Proposition 3.15: *Let $\Phi_1, \ldots, \Phi_n$ be a complete involution system, and assume that the level surface implicitely defined by $\Phi_1(p,q) = \ldots = \Phi_n(p,q) = 0$ contains a connected and compact component M_0. Then:*

(i) M_0 is a Lagrangian manifold diffeomorphic to a n–dimensional torus;

(ii) in an open neighbourhood $U(M_0)$ one can introduce action–angle variables $I \in \mathscr{G} \subset \mathbb{R}^n$ and $\vartheta \in \mathbb{T}^n$, where $\mathscr{G}$ is a neighbourhood of the origin, via a canonical diffeomorphism

$$\begin{array}{rcl} \mathscr{A} \ : \mathbb{T}^n \times \mathscr{G} & \to & U(M_0) \\ (\vartheta, I) & \mapsto & (q,p) = \mathscr{A}(\vartheta, I) \end{array}$$

such that $I_1, \ldots, I_n$ depend only on $\Phi_1, \ldots, \Phi_n$.

The proof requires two main steps. First remarking that the claim (i) is true in view of proposition 3.11 and corollary 3.12, we prove that in a neighbourhood of M_0 the level sets of $\Phi_1, \ldots, \Phi_n$ are a n–parameter family of n–dimensional tori parameterized by $\Phi_1, \ldots, \Phi_n$. Then we prove that the latter family can be parameterized by canonical variables conjugated to angular variables on a n–dimensional torus. These steps are worked out in separate subsections.

3.4.1 Periods in a neighbourhood of the torus

We extend now our considerations to a neighbourhood of the torus M_0 of point (i). Our aim is to prove that there is a smooth set of tori parameterized by $\Phi_1, \ldots, \Phi_n$ such that the flow ϕ^t on these tori admits periods $\tau_1(\Phi), \ldots, \tau_n(\Phi)$ close to the periods $e_1, \ldots, e_n$ on M_0.

Lemma 3.16: *Let the hypotheses of proposition 3.15 be fulfilled, and M_0 be the connected and compact manifold of point (i) of that proposition. Then the following statements hold true.*

(i) Let $P \in M_0$ be arbitrary but fixed, and let $\varrho > 0$ be large enough[7] so that $M_0 \subset \{\phi^t P\,,\, |t| \le \varrho\}$; then there is a neighbourhood U_ϱ of P such that for every $Q \in M_0$ the mapping

$$\phi^t \ : \ U_\varrho \to \phi^t(U_\varrho) \tag{3.35}$$

satisfying $\phi^t P = Q$ exists and is a canonical diffeomorphism.

(ii) Let P still be fixed. Then there is a neighbourhood $\mathscr{G}$ of the origin of $\mathbb{R}^n$ such that for all $\Phi \in \mathscr{G}$ the flow $\phi^t \chi(0,\Phi)$ is defined for all $t \in \mathbb{R}^n$ and possesses a stationary group of periods $G(\Phi)$; furthermore, there exist differentiable functions $W_1(\Phi), \ldots, W_n(\Phi)$ such that $G(\Phi)$ admits a basis

$$\tau_j(\Phi) = e_j + \partial_\Phi W_j \ , \quad j = 1, \ldots, n \ ,$$

where $\partial_\Phi W_j = \left(\frac{\partial W_j}{\partial \Phi_1}, \ldots, \frac{\partial W_j}{\partial \Phi_n}\right)$.

---

[7] For instance, set $\varrho = |e_1| + \ldots + |e_n|$, where $e_1, \ldots, e_n$ is the basis of the stationary group on M_0.

By (ii) we conclude that there is a neighbourhood of M_0 admitting a continuous foliation into a family of n–dimensional tori parameterized by $(\Phi_1, \ldots, \Phi_n) \in \mathscr{G}$.

Proof. (i) This is a consequence of known theorems on regularity of the solutions of a system of differential equations. Since the flow ϕ^t on M_0 is defined for all $t \in \mathbb{R}^n$, for every $\varrho > 0$ there is a neighbourhood U_ϱ of P such that for $|t| \le \varrho$ the mapping $\phi^t : U_\varrho \to \phi^t(U_\varrho)$ is a diffeomorphism. The diffeomorphism is canonical, being generated by a canonical flow; this follows from lemma 3.3. The condition above on ϱ ensures that for every $Q \in M_0$ there is a t such that $|t| \le \varrho$ and $\phi^t P = Q$. This implies in particular that the flow covers a neighbourhood of the torus.

(ii) Recall that the flow ϕ^t on M_0 admits a stationary group of periods G, with a basis $e_1, \ldots, e_n$. Recall also that we have $|e_1| \le \varrho, \ldots, |e_n| \le \varrho$ in view of our choice of ϱ in (ii). For $t = e_j$ we have $\phi^{e_j} P = P$, so that the mapping (3.35) is actually a canonical diffeomorphism between neighbourhoods of P. Possibly with a restriction of U_ϱ, we can always assume that $\phi^{e_j}(U_\varrho) \subset U(P)$, the neighbourhood of P on which a local diffeomorphism generated by the flow is defined. Consider now the sets $\chi^{-1}(U_\varrho)$ and $\chi^{-1}\bigl(\phi^{e_j}(U_\varrho)\bigr)$. `*** Il riferimento a` χ `rimanda al lemma 3.6 e proposizione 3.7. Sistemare ***` They are both neighbourhoods of the origin of $\mathbb{R}^n$, and we are allowed to define a mapping

$$\begin{array}{rcl} \psi \ : \chi^{-1}(U_\varrho) & \to & \chi^{-1}\bigl(\phi^{e_j}(U_\varrho)\bigr) \\ (\alpha, \Phi) & \mapsto & (\alpha', \Phi') = \chi^{-1} \circ \phi^{e_j} \chi(\alpha, \Phi) \ . \end{array}$$

This is a canonical diffeomorphism, being a composition of canonical diffeomorphisms, and satisfies

$$\psi(0,0) = (0,0) \ , \quad \psi(\alpha, \Phi) = \chi^{-1} \phi^{e_j + \alpha} \chi(0, \Phi) \ . \tag{3.36}$$

Furthermore, since the flow ϕ^t is just a translation on α, we have

$$\psi(\alpha, \Phi) = \bigl(\alpha + w(\Phi), \Phi\bigr) \tag{3.37}$$

with some differentiable vector function $w(\Phi)$ satisfying $w(0) = 0$. By canonicity, there exists a function $W_j(\Phi)$ such that[8]

$$w(\Phi) = -\partial_\Phi W_j \ , \quad \partial_\Phi W_j \Big|_{\Phi=0} = 0 \ . \tag{3.38}$$

We look now for a period $\tau_j(\Phi)$, i.e., for a solution of the equation

$$\phi^{\tau(\Phi)} \chi(0, \Phi) = \chi(0, \Phi) \ . \tag{3.39}$$

Remark again that the flow ϕ^t leaves Φ invariant, so that the period τ, if it exists, must depend only on Φ. Moreover, since $\tau_j(0) = e_j$, we are allowed to set $\tau_j(\Phi) =$

[8] Write the transformation as $\Phi' = \Phi \, , \, \alpha' = \alpha + w(\Phi)$. Since the first half of the transformation does not involve α, we can apply the arguments of example 2.10 (extended point transformation), concluding that there is a generating function $S(\Phi', \alpha) = \langle \Phi', \alpha \rangle + W(\Phi')$.

$e_j + \delta_j(\Phi)$, with $\delta(0) = 0$. Let for a moment $\delta_j(\Phi)$ be subjected only to the condition $\big(\delta_j(\Phi), \Phi\big) \in \chi^{-1}(U_\varrho)$. Then we have

$$\chi^{-1} \circ \phi^{\tau(\Phi)}\chi(0,\Phi) = \chi^{-1} \circ \phi^{e_j}\chi\big(\delta_j(\Phi),\Phi\big) = \psi(\delta_j(\Phi),\Phi) \ ,$$

On the other hand, by (3.37) and (3.38) we have

$$\psi(\delta_j(\Phi),\Phi) = \big(\delta_j(\Phi) - \partial_\Phi W_j, \Phi\big) \ .$$

The last two relations give

$$\chi^{-1} \circ \phi^{\tau(\Phi)}\chi(0,\Phi) = \big(\delta_j(\Phi) - \partial_\Phi W_j, \Phi\big) \ ,$$

and applying χ to both sides we see that (3.39) is satisfied with

$$\delta_j(\Phi) - \partial_\Phi W_j = 0 \ .$$

We conclude that for $j = 1, \ldots, n$ and for Φ in some neighbourhood of the origin of $\mathbb{R}^n$ there is a differentiable function $W_j(\Phi)$ such that $\tau_j(\Phi) = e_j + \partial_\Phi W_j$ is a period of ϕ^t. Therefore, in a neighbourhood $U(M_0)$ of M_0 the flow ϕ^t is defined for all $t \in \mathbb{R}^n$, and the subgroup of $\mathbb{R}^n$

$$G(\Phi) = \{m_1\tau_1(\Phi) + \ldots + m_n\tau_n(\Phi)\,,\,(m_1,\ldots,m_n) \in \mathbb{Z}^n\} \ ,$$

is a stationary group of periods of ϕ^t on M_Φ, a compact component of the manifold where the functions $\Phi_1(q,p), \ldots, \Phi_n(q,p)$ assume the constant value Φ. Remark that in view of (ii) and of the group property of the flow the periods do not depend on the point P, nor do they depend on the choice of the manifold Σ corresponding to $\alpha = 0$ as in the proof of proposition 3.6. We should prove that $G(\Phi)$ exhausts the set of periods, namely that if $\phi^t\chi(0,\Phi) = \chi(0,\Phi)$ for some $\Phi \in \mathscr{G}$ then $t \in G(\Phi)$. To this end, recall that $0 \in G(\Phi)$ for all $\Phi \in \mathscr{G}$, and that $G(\Phi)$ is a discrete group.[9] Let now t be a period on M_Φ. Then $t + G(\Phi)$ is a set of periods, and in this set there exists $t' = \mu_1\tau_1(\Phi) + \ldots + \mu_n\tau_n(\Phi)$ with $|\mu_j| =\leq 1/2$, $(j = 1, \ldots, n)$. By continuity, if Φ is sufficiently close to the origin, then t' must belong to an arbitrarily small neighbourhood of $0 \in G(0)$. Since $G(\Phi)$ is discrete, we conclude $t' = 0$, and so $t \in G(\Phi)$. *Q.E.D.*

3.4.2 Global coordinates and action–angle variables.

We come finally to the proof of point (ii) of proposition 3.15. Let us remark that the map χ defined by 3.6 can be extended to a global map $\chi : \mathbb{R}^n \times \mathscr{G} \to U(M_0)$ by setting $\chi(\alpha,\Phi) = \phi^\alpha\chi(0,\Phi)$ for $\alpha \in \mathbb{R}^n$ and $\Phi \in \mathscr{G}$, where $\mathscr{G} \subset V_\alpha$ is a neighbourhood of the origin of $\mathbb{R}^n$. As we have seen in the previous section χ is a differentiable map

---

[9] Suppose that $G(\Phi)$ has an accumulation point. Since it is a group, then the origin is an accumulation point, too. That is, there is $t \neq 0$ arbitrarily close to the origin such that $\phi^t\chi(0,\Phi) = \chi(0,\Phi)$, contradicting the fact that χ is a local diffeomorphism.

for every $\alpha \in \mathbb{R}^n$ and is a *local* canonical diffeomorphism.[10] However, it fails to be *globally* one–to–one, due to the existence of the groups of periods $G(\Phi)$. In order to obtain a global diffeomorphism we define a new mapping

$$\mathscr{C} \ : \ \mathbb{R}^n / G(\Phi) \times \mathscr{G} \to U(M_0) \tag{3.40}$$

as the restriction of χ to the quotient set of $\mathbb{R}^n$ with respect to the group $G(\Phi)$. This is a one-to-one mapping, and so it is a canonical diffeomorphism.

We are now ready to introduce action–angle variables ϑ, I via a canonical transformation $\psi \ : \ \mathbb{R}^n / G(\Phi) \times \mathscr{G} \ \to \ \mathbb{R}^n / G(\Phi) \times \mathscr{G}$ mapping (ϑ, I) to $(\alpha, \Phi) = \psi(\vartheta, I)$. To this end let us take the generating function[11]

$$S(\Phi, \vartheta) = \frac{1}{2\pi} \sum_{k=1}^{n} \vartheta_k \left[\langle e_k, \Phi \rangle + W_k(\Phi) \right] \ , \tag{3.41}$$

where $e_1, \ldots, e_n$ is the basis of the group of periods on M_0 and $W_1(\Phi), \ldots, W_k(\Phi)$ are the functions defining the periods of $G(\Phi)$ according to lemma 3.16. We show that the transformation is well defined, and that $\vartheta_1, \ldots, \vartheta_n$ are actually angles. The transformation is implicitly written as

$$\begin{aligned} I_j &= \frac{1}{2\pi} \langle e_j, \Phi \rangle + W_j(\Phi) \ , \\ \alpha_j &= \frac{1}{2\pi} \sum_k \vartheta_k \left(e_{k,j} + \frac{\partial W_k}{\partial \Phi_j} \right) \ , \quad j = 1, \ldots, n \end{aligned} \tag{3.42}$$

where we have denoted by $e_k = (e_{k,1}, \ldots, e_{k,n})$ the vectors $e_1, \ldots, e_n$ of the basis of $G = G(0)$. The functions $I_1(\Phi), \ldots, I_n(\Phi)$ are in involution, being functions of $\Phi_1, \ldots, \Phi_n$ only, and are independent, because the rows of the Jacobian matrix

$$\frac{\partial(I_1, \ldots, I_n)}{\partial(\Phi_1, \ldots, \Phi_n)}$$

are the vectors $e_k + \partial_\Phi W_k$, namely the basis vectors of the groups $G(\Phi)$. By the way, this also shows that the generating function $S(\Phi, \vartheta)$ fulfills the required invertibility condition.

In coordinates α, Φ the canonical equations for the Hamiltonian $I_j(\Phi)$ are

$$\dot{\alpha}_k = \frac{1}{2\pi} \left(e_{j,k} + \frac{\partial W_j}{\partial \Phi_k} \right) \ , \quad \dot{\Phi}_k = 0 \ ,$$

---

[10] It is a diffeomorphism between $U(0) \times \mathscr{G}$ and $U(\alpha) \times \mathscr{G}$, where $U(0)$ and $U(\alpha)$ are open neighbourhoods of 0 and α, respectively. It is canonical because it is a composition of two canonical diffeomorphisms, namely $\chi(\alpha, \Phi)$ restricted to a neighbourhood of the origin and the canonical flow ϕ^α.

[11] I use the notation $\langle a, b \rangle = \sum_j a_j b_j$ for vectors or vector–valued functions a, b.

where we have denoted by $e_j = (e_{j,1}, \ldots, e_{j,n})$ the vectors of the basis of $G = G(0)$. Therefore we have

$$\alpha_k(t) = \alpha_k(0) + \frac{t}{2\pi}\left(e_{j,k} + \frac{\partial W_j}{\partial \Phi_k}\big(\Phi_k(0)\big)\right) \ , \quad \Phi_k(t) = \Phi_k(0) \ . \tag{3.43}$$

In particular, for $t = 2\pi$ we have $\alpha(t) - \alpha(0) = \tau_j\big(\Phi(0)\big)$, the period on $M_{\Phi(0)}$. On the other hand, in coordinates ϑ, I the canonical flow due to the Hamiltonian I_j is

$$\vartheta_k(t) = \vartheta_k(0) + \delta_{j,k} t \ , \quad I_k(t) = I_k(0) \ .$$

By comparison with (3.43) we conclude that $\vartheta_1, \ldots, \vartheta_n$ are angular coordinates with period 1.

The form of the canonical transformation $\mathscr{A}$ of proposition 3.15 is found by setting $\mathscr{A} = \mathscr{C} \circ \psi$ where $(\alpha, \Phi) = \psi(\vartheta, I)$ is the canonical transformation(3.42) after a suitable inversion, and $(q,p) = \mathscr{C}(\alpha, \Phi)$ is the canonical transformation (3.40). This concludes the proof of proposition 3.15.

3.4.3 Non uniqueness of the action–angle variables

The action–angle variables defined above are not unique. Indeed, three arbitrary choices have been made throughout the proof, namely: (i) the choice that M_0 corresponds to $\Phi_1 = \ldots = \Phi_n = 0$; (ii) the choice of the point P corresponding to the origin of $\mathbb{R}^n$ and of the manifold Σ; (iii) the choice of the basis $e_1, \ldots, e_n$ of periods on M_0.

Lemma 3.17: *Let ϑ, I be action angle variables. New action angle variables are constructed by composition of the following canonical transformations:*

(i) translation of the action variables

$$I_j = \overline{I}_j + c_j \ , \quad \vartheta_j = \overline{\vartheta}_j \ , \quad 1 \le j \le n \tag{3.44}$$

with $c \equiv (c_1, \ldots, c_n) \in \mathbb{R}^n$;

(ii) translation of the origin of the angles by a quantity depending on the torus, namely

$$\vartheta_j = \overline{\vartheta}_j + \frac{\partial S}{\partial \overline{I}_j}(\overline{I}) \ , \quad I_j = \overline{I}_j \ , \quad 1 \le j \le n \tag{3.45}$$

where $S(\overline{I})$ is an arbitrary differentiable functions of the action variables;

(iii) linear transformation of the angle variables by a unimodular matrix[12] A:

$$\overline{\vartheta} = \mathsf{A}\vartheta \ , \quad I = \mathsf{A}^\top \overline{I} \ , \tag{3.46}$$

Proof. The canonicity of the transformations above is easily checked. For instance, for the transformations (i) and (ii) it is enough to verify that the fundamental Poisson brackets are preserved. For the transformation (iii) just remark that $\overline{\vartheta} = \mathsf{A}\vartheta$ involves

---

[12] A matrix A is said to be unimodular if it has integer entries, and $\det \mathsf{A} = \pm 1$.

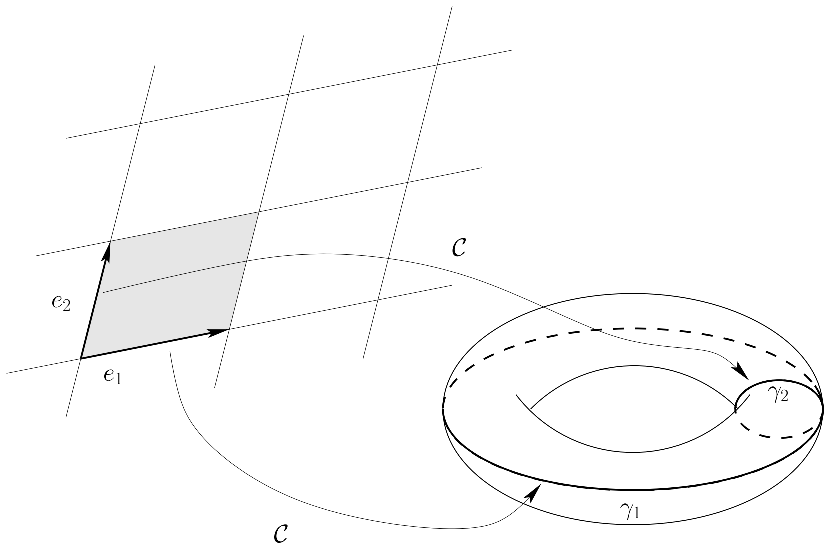

Figure 3.3. Illustrating the construction of cycles on a torus.

only the angles; therefore, is is legitimate to apply the method of the extended point transformation, thus writing the generating function as $S = \overline{I}^{\top} \mathsf{A}\vartheta$.
Only the condition that A be a unimodular matrix needs some justification. To this end, just remark that the angle structure of the torus is preserved if and only if the basis $(e_1, \ldots, e_n)$ of the group of periods is changed by $e'_j = \sum_k A_{jk} e_k$ (where $A_{j,k}$ are the entries of A) into a new basis of the same group. This leads to the condition that A should be unimodular. A detailed proof is given in appendix A, lemma A.3. *Q.E.D.*

3.4.4 Explicit construction of action–angle variables

The explicit algorithm for constructing action–angle variables is based on the construction of the family of periods on the manifolds M_Φ considered in the previous sections. For simplicity, denote by $e_1, \ldots, e_n$ the basis of the stationary group on any of the tori parameterized by a given value of Φ. The following properties are immediate:

(i) to each period e_j $(j = 1, \ldots, n)$ there corresponds on M_0 a differentiable closed curve γ_j, that we call a *cycle*, defined as (see fig. 3.3)
$$\gamma_j = \{\phi^{se_j} P \, , \, 0 \le s < 1\} \; ;$$

(ii) for $j \neq k$ the cycles γ_j , γ_k are independent, in the sense that γ_j can not be continuously deformed into γ_k;

(iii) in action–angle variables I, ϑ the continuous set of tori is parameterized by $I_1, \ldots, I_n$, and the cycle γ_j is represented by
$$\vartheta_j \in [0, 2\pi) \; , \quad \vartheta_k = \vartheta_{k,0} \text{ for } k \neq j \; ,$$
$\vartheta_{k,0}$ being constants.

Since the transformation to action–angle variables is canonical, we have

$$\oint_{\gamma_j} \sum_k p_k\, dq_k = \oint_{\gamma_j} \sum_k I_k\, d\vartheta_k\ .$$

On the other hand, by the characterization (iii) of the cycles all actions $I_1, \ldots, I_n$ and all angles ϑ_k with $k \neq j$ are constant on γ_j, so that the integral on the r.h.s. gives

$$\oint_{\gamma_j} \sum_k I_k\, d\vartheta_k = I_j \int_0^{2\pi} d\vartheta_j = 2\pi I_j\ .$$

We conclude

$$I_j = \frac{1}{2\pi} \oint_{\gamma_j} \sum_k p_k\, dq_k\ , \tag{3.47}$$

where $p_1(\Phi, q), \ldots, p_n(\Phi, q)$ are obtained by inversion of $\Phi_1 = \Phi_1(q, p), \ldots, \Phi_n = \Phi_n(q, p)$ with respect to $\Phi_1, \ldots, \Phi_n$. We emphasize that, by Stokes theorem, a continuous deformation of the cycle γ_j does not change the result, so that any determination of the cycles can be used in computing the integral (3.47). The resulting functions $I_1, \ldots, I_n$ depend only on $\Phi_1, \ldots, \Phi_n$, of course.

In view of this discussion, the algorithm for constructing action–angle variables for a given problem requires three steps.

- (i) Find the cycles γ_j ($j = 1, \ldots, n$). This is expected to be the hardest part, because it requires in principle an integration of the system via Liouville's algorithm applied to the involution system $\Phi_1, \ldots, \Phi_n$. However, in the most commonly considered examples the first integrals have a nice form, so that the cycles are easily determined.
- (ii) Compute the action variables by quadrature, calculating the integrals (3.47). This can be possibly done by introducing some arbitrary angle variables on the cycles, and then integrating over them.
- (iii) Apply the algorithm of Liouville to the new involution system $I_1, \ldots, I_n$ in order to find the angle variables $\vartheta_1, \ldots, \vartheta_n$.
- (iv) If useful, and if there is any reason to do it, apply any of the transformations of lemma 3.17 in order to obtain better sets of action–angle variables, depending on the problem at hand.[13]

3.5 The theorem of Arnold–Jost

Having settled the problem of constructing action–angle variables, we turn now to the statement of the theorem of Arnold–Jost on integrable systems.

[13] The usefulness of the latter step is widely discussed in M. Born's treatise [15], in connection with the necessity of having a well definite set of action–angle variables to which quantization rules can be applied.

Theorem 3.18: *Let the Hamiltonian $H(q,p)$ on the phase space $\mathscr{F}$ possess an involution system $\Phi_1,\dots,\Phi_n$ of first integrals (so that it is integrable in Liouville's sense). Let $c=(c_1,\dots,c_n)\in\mathbb{R}^n$ be such that the level surface determined by the equations $\Phi_1(q,p)=c_1,\dots,\Phi_n(q,p)=c_n$ contains a compact and connected component M_c. Then in a neighbourhood U of M_c there are canonical action–angle coordinates I,ϑ mapping $\mathscr{G}\times\mathbb{T}^n$ to U, where $\mathscr{G}\in\mathbb{R}^n$ is an open set, such that the Hamiltonian depends only on $I_1,\dots,I_n$, and the corresponding flow is*

$$\vartheta_j(t)=\vartheta_{j,0}+t\omega_j(I_{1,0},\dots,I_{n,0})\ ,\quad I_j(t)=I_{j,0}\ ,\quad j=1,\dots,n\ ,$$

where $\vartheta_{j,0}$ and $I_{j,0}$ are the initial data, and $\omega_j=\frac{\partial H}{\partial I_j}$.

The proof is a straightforward application of proposition 3.15. Just proceed as in the proof of Liouville's theorem, using the actions $I_1,\dots,I_n$ as first integrals.

3.6 Delaunay variables for the Keplerian problem

A remarkable application of the theory of Arnold–Jost is the calculation of action–angle variables for the motion in a central field of force, with particular reference to the case of the Keplerian potential.[14] The latter problem is known to possess four independent first integrals (see examples 1.12 and 1.13). A complete involution system of first integrals has been constructed in example 3.1. Let us recall that in spherical coordinates r, ϑ, φ with the conjugated momenta p_r, p_ϑ, p_φ the functions are

$$J=p_\varphi\ ,\quad \Gamma^2=p_\vartheta^2+\frac{J^2}{\sin^2\vartheta}\ ,\quad E=\frac{1}{2m}\left(p_r^2+\frac{\Gamma^2}{r^2}\right)+V(r)\ , \tag{3.48}$$

where m is the mass of the point. Concerning the potential $V(r)$, in Kepler's case we put

$$V(r)=-\frac{k}{r}\ , \tag{3.49}$$

[14] The action angle variables that I'm going to calculate here were discovered by Delaunay. His aim was to replace the orbital elements of a Keplerian orbit, which were used since Lagrange's time in perturbation theory, with an appropriate set of canonical variables. A deduction of Delaunay variables using the method of Hamilton–Jacobi is found in Poincaré's treatises [71] and (more detailed) [72]. The calculation in these notes reflect the exposition in M. Born's book [15]. The introduction of action–angle variables in Born's treatise is connected with the search for *adiabatic invariants*, which were well known to physicists when dealing with one–dimensional oscillators depending on slow varying parameters. A clever generalization of these concepts to the case of separable systems with many degrees of freedom is made by Born. Essentially, it may be said that his book contains a complete treatment of action–angle variables for system which exhibit a clear separation into one–dimensional systems for which the existence of cycles is immediate. This in turn implies that the orbits lie on invariant tori. What is still missing with respect to the theorem of Arnold–Jost is that the existence of invariant tori is a general fact for integrable systems.

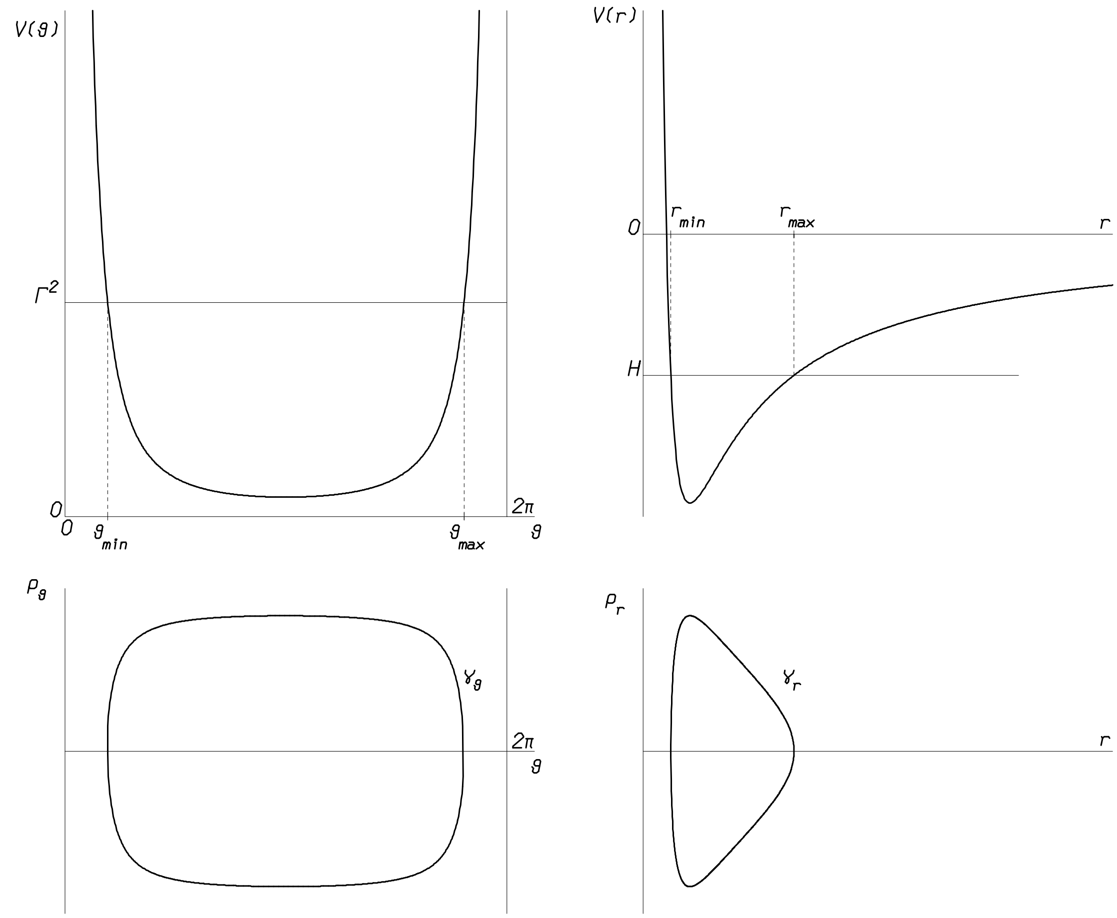

Figure 3.4. Illustrating the construction of the cycles γ_ϑ and γ_r for the problem of motion in a central field under a generic potential. The cycle γ_r is represented for the case of the Keplerian ptential.

where k is a positive constant. We also recall the expression of the Hamiltonian

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\vartheta}\right) + V(r)\ ; \tag{3.50}$$

this actually coincides with the third integral above when the explicit expressions of Γ^2 and J are substituted.

3.6.1 Determination of cycles

We consider the canonical flows generated by the three functions (3.48). The discussion here is quite plain, because each function involves only two conjugated variables. This considerably simplifies the task of constructing the cycles.

The function J is a trivially integrable Hamiltonian: the conjugate variable is actually an angle, and is a cyclic variable, so that the cycle γ_φ is parameterized by the angle φ itself.

The function Γ^2 can be considered as the Hamiltonian of a point with unit mass, moving on the segment $[0, \pi]$ under the action of the potential $V(\vartheta) = J^2/\sin^2\vartheta$. For $\Gamma^2 > \Gamma^2_{\rm min} = J^2$ the orbit in the phase plane ϑ, p_ϑ is a closed line, giving the second cycle γ_ϑ (see fig. 3.4). Remark that the construction of the cycles γ_φ and γ_ϑ does not depend on the form of the potential $V(r)$.

The third function can be considered as the Hamiltonian of a point moving on the half–line $r > 0$ under the action of the potential

$$V^*(r) = \frac{\Gamma^2}{2mr^2} + V(r)\ .$$

In the Keplerian case the motion on the half line r is bounded for $E_{\rm min} < E < 0$, with $E_{\rm min} = -mk^2/(2\Gamma^2)$, while for $E \geq 0$ it is unbounded. In the first case the orbit in the phase plane r, p_r is a closed curve, and this gives the third cycle γ_r (see fig. 3.4). Conversely, no cycle can be found for $E \geq 0$, and the invariant surface in phase space for the complete problem is actually the product $\mathbb{T}^2 \times \mathbb{R}$. In the latter case the angular variables can be introduced only for the cycles γ_φ and γ_ϑ.

3.6.2 Construction of the action variables.

Here we restrict our consideration to the Keplerian potential, with the condition $E_{\rm min} < E < 0$. By inversion of (3.48) we get

$$\begin{aligned} p_r &= \left[2m(E - V(r)) - \frac{\Gamma^2}{r^2}\right]^{\frac{1}{2}} \\ p_\vartheta &= \left(\Gamma^2 - \frac{J^2}{\sin^2\vartheta}\right)^{\frac{1}{2}} \\ p_\varphi &= J\ . \end{aligned} \tag{3.51}$$

We have to integrate the differential form $p_r\,dr + p_\vartheta\,d\vartheta + p_\varphi\,d\varphi$ over the cycles γ_φ, γ_ϑ and γ_r. This gives

$$\begin{aligned} I_\varphi &= \frac{1}{2\pi}\oint_{\gamma_\varphi} p_\varphi {\rm d}\varphi = J \\ I_\vartheta &= \frac{1}{2\pi}\oint_{\gamma_\vartheta} p_\vartheta {\rm d}\vartheta = \Gamma - |J| \\ I_r &= \frac{1}{2\pi}\oint_{\gamma_r} p_r dr = -\Gamma + k\sqrt{-\frac{m}{2E}}\ . \end{aligned} \tag{3.52}$$

This gives the actions I_φ, I_ϑ and I_r as functions of J, Γ and E. The exlicit expression as a function of the canonical coordinate is readily found by replacing the expressions of J, Γ and E in (3.48).

3.6.3 Delaunay variables

By a straightforward inversion of the third of (3.52) we calculate the Hamiltonian as

$$H = -\frac{mk^2}{2\left(I_r + I_\vartheta + |I_\varphi|\right)^2}\ . \tag{3.53}$$

It is immediately seen that the Hamiltonian actually depends on the sum of the action variables. This implies that the three frequencies of the system coincide, which justifies the fact that in the Keplerian description of the planetary motion only one frequency does actually appear. A better set of action variables is constructed by introducing the variables of Delaunay L, G, Θ defined by the linear transformation

$$\begin{aligned} L &= I_r + I_\vartheta + |I_\varphi| \\ G &= I_\vartheta + |I_\varphi| \\ \Theta &= |I_\varphi| \ . \end{aligned} \tag{3.54}$$

It is immediate to notice that G and Θ coincide with Γ and J, respectively. Since the transformation is performed via a unimodular matrix, the corresponding transformation on the angles preserves the periods, as stated by lemma 3.17. The Hamiltonian in Delaunay's variables takes the well known form

$$H = -\frac{mk^2}{2L^2} \ . \tag{3.55}$$

Denoting by ℓ, g, h the angles conjugated to the actions G, G, Θ we can write Hamilton's equations as

$$\dot\ell = \frac{mk^2}{L^3} \ , \quad \dot g = \dot h == \dot L = \dot G = \dot\Theta = 0 \ .$$

Thus the motion is periodic with a single frequency

$$\omega(L) = \frac{mk^2}{L^3} \ .$$

3.6.4 Construction of the angle variables

The canonical transformation should now be completed by constructing the angle variables associated to the actions L, G and Θ. To this end we must first write the generating function

$$\begin{aligned} S &= \int (p_r dr + p_\vartheta d\vartheta + p_\varphi d\varphi) \\ &= \int \sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G^2}{r^2}}\, dr + \int \sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}\, d\vartheta + \int \Theta d\varphi \ . \end{aligned}$$

The angle variables are then given by

$$\begin{aligned} \ell &= \frac{\partial S}{\partial L} = \frac{m^2k^2}{L^3} \int \frac{dr}{\sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G}{r^2}}} \ , \\ g &= \frac{\partial S}{\partial G} = G \int \frac{d\vartheta}{\sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}} - G \int \frac{dr}{r^2\sqrt{-\frac{m^2k^2}{L^2} + \frac{2mk}{r} - \frac{G^2}{r^2}}} \ , \\ h &= \frac{\partial S}{\partial \Theta} = -\Theta \int \frac{d\vartheta}{\sin^2\vartheta \sqrt{G^2 - \frac{\Theta^2}{\sin^2\vartheta}}} + \int d\varphi \ . \end{aligned}$$

Thus, the calculation of the angle variables is reduced to a quadrature.

It may also be useful to recall the relation between the Delaunay actions and the so called *orbital elements*. I just report these relations

$$L = \sqrt{mka} \ , \quad G = L\sqrt{1 - e^2} \ , \quad \Theta = G \cos \iota \ , \tag{3.56}$$

where a is the semimajor axis, e is the eccentricity and ι is the inclination of the orbital plane.

The conjugated angles are also related to orbital elements. Indeed, ℓ is the so called *mean anomaly*, namely an angle that evolves uniformly, thus averaging in some sense the *true anomaly* which is the angle giving the actual position of the planet on the sky. The angle g and h are the *longitude of the perihelion* and the *longitude of the node*, respectively.