Francesco Fassò

Università di Padova
Dipartimento di Matematica Pura e Applicata
Via G. Belzoni 7, 35131 Padova, Italy
E–mail: fasso@math.unipd.it

Notes on Finite Dimensional Integrable Hamiltonian Systems

Preliminary Version
January 1999

Contents

Preface

These notes are based on graduate courses that I gave at the Universidade Federal de Pernambuco (Recife, Brazil) in 1993, at the University of California at Santa Cruz in 1996, and at the University of Padova in 1997. The goal was to provide an introduction to the classical theory of integrable Hamiltonian systems (with an eye to Hamiltonian perturbation theory, which was part of the courses but is not covered in these notes), that is, the Liouville–Arnol'd theorem, the construction of the action–angle coordinates, and the geometric structure of the fibration by the invariant tori.

The characteristic of the present approach is the emphasis on the case of 'degenerate', or 'superintegrable', systems, namely systems whose invariant tori are smaller than one half the dimension of the phase space. This is an important case, since it is met in many of the classical systems of mechanics, like Kepler, the free rigid body, and systems of resonant harmonic oscillators. In these degenerate cases, the primary object is the fibration by the invariant tori of the smallest dimension, which is *not* described by the Liouville–Arnol'd theorem. As it turns out, the correct approach is provided by the generalization of the Liouville–Arnol'd theory which is sometimes referred to as 'noncommutative integrability'. This is primarily a geometric theory, in which symplectic manifolds and Poisson manifolds naturally appear; instead than a single Lagrangian fibration, as in the Liouville–Arnol'd case, there is a 'dual pair' constituted by an isotropic fibration (the invariant tori) and by its polar coisotropic foliation. The consideration of both these foliations is essential in order to understand the structure of a degenerate system.

The notes are organized as follows.

Chapter 1 aims to give some familiarity with a number of examples of integrable systems. The idea is that the reader should construct 'by hands' and 'visualize' the invariant tori and the quasi–periodic motions in a number of examples before studying the general structure of the fibration by the invariant tori.

Chapter 2 is devoted to the Liouville–Arnol'd theorem, which is regarded as describing the local structure of a Lagrangian fibration. We provide a complete proof of the theorem, in which we have tried to enlighten the geometric aspects whenever this seems to help the comprehension.

Chapter 3 is devoted to the general properties of completely integrable systems, the relation with the Hamilton–Jacobi theory, and the construction of action–angle coordinates. We illustrate in details the construction of these coordinates for Kepler and for the free symmetric rigid body with a fixed point.

Chapter 4 deals with degenerate systems. In a way, the treatment in Chapter 4 makes the proof of the Liouville–Arnol'd theorem in Chapter 2 useless. Nevertheless, we have preferred to give first the simpler case of Liouville-Arnol'd, since it is probably more familiar.

The last Chapter contains a very short discussion of the relations between integrability and symmetry, a topic which has a remarkable foundational interest.

We assume that the the reader has a certain familiarity with Hamiltonian systems, differentiable manifolds, and the very bases of symplectic geometry. The theory of Poisson manifolds, which is central to the understanding of degenerate systems, is shortly reviewed. In the Appendix, we assume some knowledge of symplectic actions and momentum maps.

Acknowledgments: I wish to thank Hildeberto Eulalio Cabral for his invitation at the Departamento de Matemática of the Universidade Federal de Pernambuco in November and December 1993, where a first draft of these notes was written, and Tudor Ratiu and Debra Lewis for their hospitality at the Department of Mathematics of the University of California at Santa Cruz in the years 1995–1996, where a large part of the notes was rewritten and further developed. I'm very indebted to all of them and to Giancarlo Benettin for many stimulating and clarifying discussions.

My visit at UFdP was made possible by the CNpQ contract #452790/93-0(NV) and my visit at UCSC was granted by a fellowship of *Consiglio Nazionale delle Ricerche*.

Chapter 1

Examples of Integrable Systems

In this chapter we provide a few classical examples of 'completely integrable' Hamiltonian systems, with the purpose of putting into evidence some basic facts: these systems have a sufficiently large number of integrals of motion, the common level sets of such integrals, if compact, are diffeomorphic to tori of a certain dimension, and motions are linear on such tori, that is, they are 'quasi–periodic'.

The treatment in this chapter is deliberately not exhaustive: our main objective is to offer a collection of examples in which, by working out all of the details (which are often tacitly left as exercises) one can directly construct—and visualize—the invariant tori.

The (local) structure of a completely integrable system is described in a clear way by the introduction of the so–called *action–angle coordinates*, whose existence is assured by the Liouville–Arnol'd theorem, and by some generalizations of it, which are the object of the next chapters. The construction of these coordinates in specific cases is usually difficult (it is in a sense equivalent to the integration of the equations of motion, and in most cases can be done only up to quadratures), so we shall be able to exhibit them only in very few simple cases. But besides the explicit computations, we put the emphasis on some conceptual aspects of the construction of these coordinates which will enter the proof of the Liouville–Arnol'd theorem and of its generalizations.

1.1. The harmonic oscillator.

As a first example we consider the harmonic oscillator. This is an extremely simple case, whose consideration nevertheless allows one to introduce a few concept which will recurr in the sequel. The harmonic oscillator consists of a point constrained to a line and subject to a linear attracting force. The Hamiltonian is

$$h(p_x, x) \;=\; \frac{p_x^2}{2m} \;+\; k\frac{x^2}{2} \qquad (p_x, x) \in \mathbb{R}^2\,,$$

with positive constants m and k. The change of coordinates $(x, p_x) \mapsto (q = \sqrt{k}\,x, p = p_x/\sqrt{m})$ leads to the Hamiltonian

$$H(p, q) \;=\; \frac{\omega}{2}\,(p^2 \;+\; q^2)\,, \qquad\qquad (p, q) \in \mathbb{R}^2\,,$$

where $\omega = \sqrt{k/m} > 0$, to which we will refer in the sequel.

Since H is a first integral, the trajectories of the system lie on the level curves $H(p,q) = h$, which are a family of circles around the origin for $h > 0$ and the origin itself for $h = 0$, see figure 1.1.

The idea under the introduction of the action–angle coordinates is that of looking for canonical coordinates in which all periodic motions appear as linear motions on the circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$. Obviously, this is not possible for the equilibrium, so we restrict ourselves from the very beginning to the subset of the phase space filled with periodic orbits, namely the punctured plane

$$M = \mathbb{R}^2 \setminus \{(0,0)\} .$$

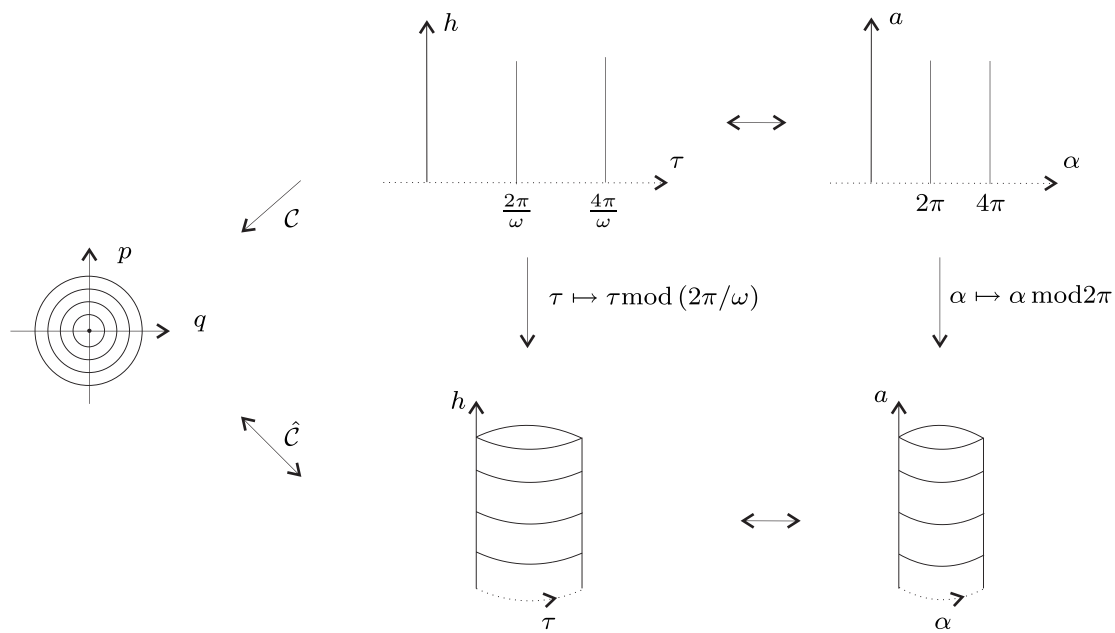


Figure 1.1

The energy-time coordinates. A quite natural idea is that of taking as coordinates the energy h, which provides a coordinate transversal to the orbits, and, as a coordinate on each orbit, the time τ along the flow of the system; since h is constant and τ runs with constant unit speed, the equations of motion are

$$\dot{h} = 0 , \qquad \dot{\tau} = 1 .$$

More precisely, in order to define the time coordinate τ, we need to choose preliminarily an 'origin' of time on each orbit, namely a point corresponding to the time $\tau = 0$. This means that we must choose a section $h \mapsto (p_o(h), q_o(h))$ of the foliation by the circles $H(p,q) = h$; smootheness of the coordinates of course requires smoothness of the section. We then define $\tau(p,q)$ as the time necessary for the system to reach the point (p,q) starting from the point $(q_o(h), p_o(h))$ on the same orbit.

Let us choose, for instance, $(p_o(h), q_o(h))$ as the point of the orbit of energy h which belongs to the semi–axis $p > 0$, namely

$$q_o(h) = 0\,, \qquad p_o(h) = \sqrt{2h/\omega}\,.$$

Correspondingly, since the flow of the system is the uniform rotation

$$p(t, p_0, q_0) = -q_0\, sin\,\omega t + p_o\, cos\,\omega t\,, \qquad q(t, p_0, q_0) = q_0\, cos\,\omega t + p_o\, sin\,\omega t\,,$$

the coordinates (p, q) are related to (h, τ) by

$$p(h, \tau) = \sqrt{2h/\omega}\; cos\,\omega\tau\,, \qquad q(h, \tau) = \sqrt{2h/\omega}\; sin\,\omega\tau\,.$$

Are the functions (h, τ) constructed in this way coordinates? Smoothness is obviously not an issue but, because of the periodicity of the flow, there is no bijective correspondence: all points $(h, \tau + 2\pi\nu/\omega)$, $\nu \in \mathbb{Z}$, correspond to the same point of the punctured plane M. In this obvious sense, (h, τ) are not global coordinates for M: they are instead global coordinates on its covering $\mathbb{R}_+ \times \mathbb{R}$.[†] In fact, what we have done above is to construct a (covering) map

$$\mathcal{C} : \mathbb{R}_+ \times \mathbb{R} \to M\,, \qquad (h, \tau) \mapsto \left(\sqrt{2h/\omega}\; cos\,\omega\tau,\; \sqrt{2h/\omega}\; sin\,\omega\tau\right)$$

which is obviously surjective but not injective. Its restriction to any subset on which it is bijective (that is, on which τ varies in intervals shorter than one period) is a diffeomorphism, so we can use its local inverses as (local, in the specified sense) coordinates on M.[‡] It is a remarkable fact that these coordinates are symplectic, namely $dh \wedge d\tau = dp \wedge dq$; hence, the Hamiltonian of the system is h.

An equivalent, but more global and geometric, point of view is obtained by identifying all points of the covering space $\mathbb{R}_+ \times \mathbb{R}$ which correspond to the same point of M, by means of the map $(h, \tau) \mapsto (h, \tau mod\,(2\pi/\omega))$. In this way, one obtains a symplectic diffeomorphism

$$\hat{\mathcal{C}} : \mathbb{R}_+ \times S^1_{2\pi/\omega} \to M\,, \qquad (h, \tau) \mapsto \left(\sqrt{2h/\omega}\; cos\,\omega\tau,\; \sqrt{2h/\omega}/, sin\,\omega\tau\right)$$

between M and the cylinder $\mathbb{R}^+ \times S^1_{2\pi/\omega}$, equipped with the symplectic two–form $dh \wedge d\tau$. (This map is a diifeomorphism because it is a bijection and a local diffeomorphism).

The action–angle coordinates. The coordinates (h, τ) just constructed differ from the action–angle coordinates we are looking for only in that the time coordinate τ is not a true angle: if the point (p, q) runs once on its orbit, $\tau(p, q)$ increases of the period $T = 2\pi/\omega$, not of 2π. We can thus 'normalize the period' with the canonical linear change of coordinates

$$(h, \tau) \mapsto (a, \alpha) = (h/\omega, \omega\tau)\,,$$

which is obviously a diffeomorphism. (The rescaling of the energy coordinate serves only to assure the symplecticity). Note that this transformation can be equivalent performed

[†] This is exactly the same situation which is encountered with the polar coordinates in the plane, or even just when coordinatizing the circle with the angle.

[‡] All together, these coordinates systems form an atlas for M; the transition functions between any two such local coordinate systems are of course $h' = h$, $\tau' = \tau + 2\nu\pi$, $\nu \in \mathbb{Z}$.

either on the cylinder $\mathbb{R}_+ \times S^1_{2\pi/\omega}$ or on its covering $\mathbb{R}_+ \times \mathbb{R}$, see figure 1.1, but in the latter case one has still to pass to the quotient.

So, the conclusion is that the set M is symplectically diffeomorphic to the standard cylinder $\mathbb{R} \times S^1$. This diffeomorphism, as well as the (local) coordinates (a,α) on M, are called action–angle coordinates. They are related to the orignal coordinates (p,q) by

$$p(a,\alpha) \;=\; \sqrt{2a}\,\cos\alpha\,, \qquad q(a,\alpha) \;=\; \sqrt{2a}\,\sin\alpha\,. \tag{1.1.1}$$

Expressed as a function of (a,α) the Hamiltonian is ωa, so the equations of motion are $\dot a = 0$, $\dot\alpha = \omega$.

1.2. The pendulum.

The pendulum is the system described by the Hamilton function $H(p,\theta) = \frac{p^2}{2} + \cos\theta$ on the cylinder $(p,\theta) \in \mathbb{R} \times S^1$. By drawing the level curves $H(p,\theta) = h$, $h \ge -1$, which can be represented as $p = \pm\sqrt{2(h-\cos\theta)}$, one obtains the phase portrait of the system, which is shown in figure 1.2, where are reported also the topological type and the period of the level curves as functions of the energy h. The phase portrait consists of two equilibria, of the two separatrices connecting the unstable one, and of two disconnectd subsets filled with periodic orbits, the librations (for $h < 1$) and the rotations (for $h > 1$).

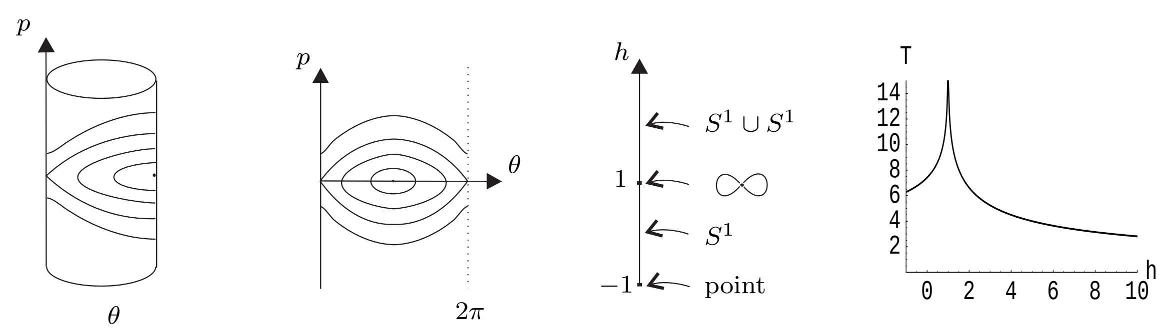


Figure 1.2

In principle, the construction of the action–angle coordinates for the pendulum can be performed in each subset of rotational and librational motions as in the case of the harmonic oscillator. There is however the difference that the periodic orbits have now different periods: the period $T(h)$ of the orbit of energy h is

$$T(h) \;=\; 2\int_{\theta_-(h)}^{\theta_+(h)} [2(h-\cos\theta)]^{-1/2} d\theta \tag{1.2.1}$$

where $\theta_\pm(h)$ are the two roots of $H(0,\theta) = h$ if $-1 < h < 1$ while $\theta_-(h) = 0$ and $\theta_+(h) = 2\pi$ if $h > 1$, see figure 1.2. Thus, the pendulum is an *anisochronous* system.

To be definite, we consider from now on only the librational motions, that is, we restrict our considrations to the anulus $M = \{(p,\theta) : -1 < H(p,\theta) < 1\}$.

The energy–time coordinates. We begin by introducing the energy–time coordinates (h,τ). As origin of time on each orbit we take the point $p = \sqrt{2(h+1)}$, $\theta = \pi$. This leads to the covering map $\mathcal{C} : (-1,1) \times \mathbb{R} \to M$ defined by

$$\mathcal{C}(h,\tau) \;=\; \Phi^H_\tau\left(\sqrt{2(h+1)}, \pi\right) ,$$

where $(\tau, (p,\theta)) \mapsto \Phi^H_\tau(p,\theta)$ is the pendulum's flow. Thus, (h,τ) are global coordinates on the covering $(-1,1)\times\mathbb{R}$ of M and we can use the local inverses of the map $\mathcal{C}$ as (local) coordinates on the anulus. As in the harmonic oscillator case, these coordinates are symplectic.

Note that, since the period is not constant, by identifying all points of the strip $(-1,1)\times\mathbb{R}$ which correspond to a same point of the anulus, one obtains a 'stretched' cylinder, as shown in figure 1.3. A little reflection will show that the map $(p,\theta) \mapsto (h, \tau \,mod\, T(h))$ is a diffeomorphism from M onto such a stretched cylinder.

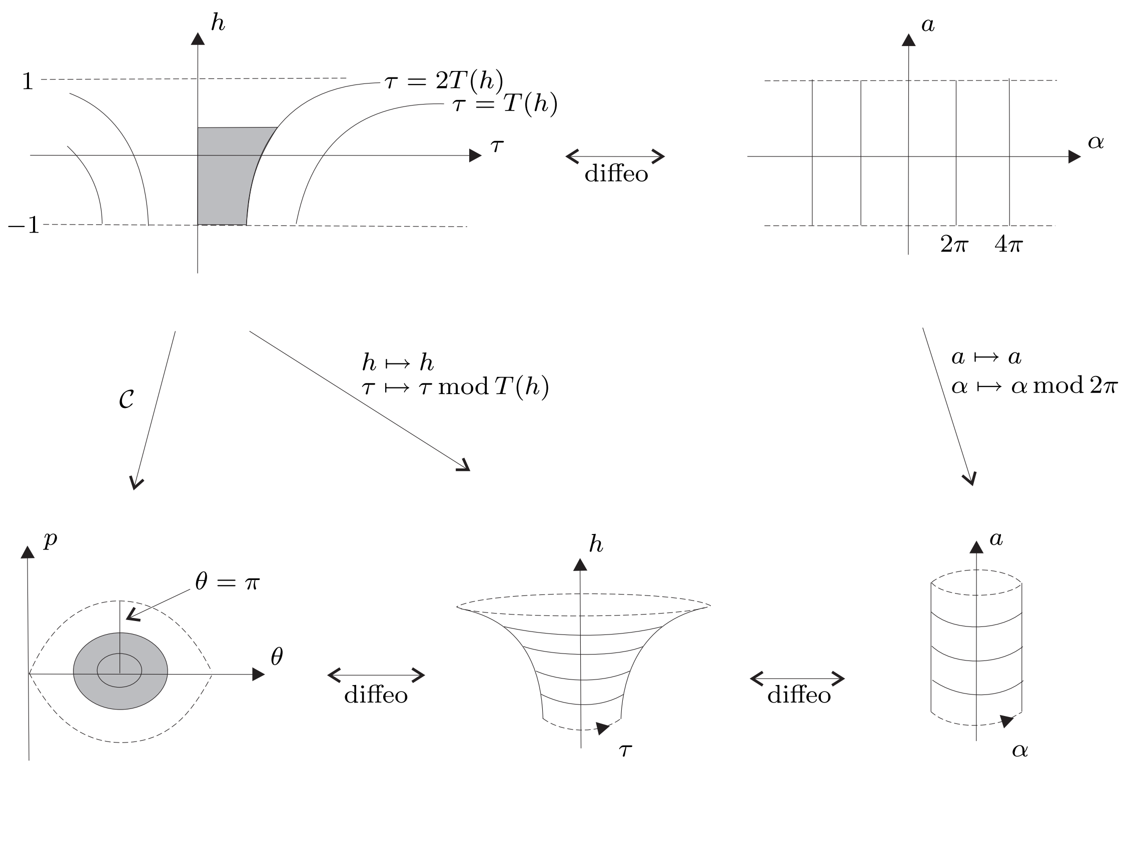


Figure 1.3: Construction of the action–angle coordinates of the pendulum.

Exercises: (i) Consider a Hamiltonian system with one degree of freedom whose all orbits, in a certain region of its phase space, are periodic. Outline the construction of its energy–time coordinates. Use the Hamilton–Jacoby method to verify that they are symplectic.

(ii) Show that $\mathcal{C}$ is a diffeomorphism from the set $\{(h,\tau) : |h| < 1, -T(h)/4 < \tau < 3T(h)/4\}$ onto the subset of the anulus M obtained by removing the segment $\{p = 0,\, 0 < \theta < \pi/2\}$. (Hint: the inverse $\mathcal{C}^{-1}$ is given by $h = H(p,\theta)$ and by

$$\begin{aligned}
\tau(p,\theta) &= \int_\pi^\theta \left[2(h-\cos\theta)\right]^{-1/2} d\theta && \text{if } p>0\\
&= \tfrac{1}{4}\,T(h) + \int_{\theta_+(h)}^\theta \left[2(h-\cos\theta)\right]^{-1/2} d\theta && \text{if } p<0\\
&= -\int_{\sqrt{2h}}^p \left[1-(h-p^2/2)^2\right]^{-1/2} dp && \text{if } \theta>\pi\,,
\end{aligned}$$

where $h = H(p,\theta)$ and $\theta_+(h)$ is as above; the last equation serves only to define τ on the segment $p = 0$, $\pi < \theta < 2\pi$).

The action–angle coordinates. We now rescale the coordinate τ so as to normalize its period to 2π; for definiteness, we work on the covering space. in order to keep the symplectic character of the coordinates, we need to correspondingly change the coordinate h. Thus, we take

$$\alpha(h,\tau) = \frac{2\pi}{T(h)}\,\tau \tag{1.2.2}$$

and we aim to construct a map $a : (-1,1) \to A$, $h \mapsto a(h)$, where A is some real interval, such that the map* $a\times\alpha$ is a symplectic diffeomorphism of the covering space $(-1,1)\times\mathbb{R}$ onto $A\times\mathbb{R}$; obviously, we want a to depend only on h so that the invariant circles are mapped into the curves $a = const.$ The map $a(h)$ is determined by observing that, on account of (1.2.2) and of $\frac{\partial a}{\partial\tau} = 0$, the symplecticity condition $da\wedge d\alpha = dh\wedge d\tau$ reduces to

$$\frac{da}{dh} = \frac{T(h)}{2\pi}\,. \tag{1.2.3}$$

Thus, we take

$$a(h) = \frac{1}{2\pi}\int_{-1}^h T(h')dh'\,. \tag{1.2.4}$$

Note that $da/dh = T(h) > 0$, so $a\times\alpha$ is actually a diffeomorphism.

In this way, combining the two maps $(a\times\alpha)^{-1}$ and $\mathcal{C}$, we get a covering map from $A\times S^1$ onto M, which provides M with ‘action–angle’ coordinates. The Hamilton function of the pendulum is $h(a)$ (i.e., the inverse of (1.2.4)) and the equations of motion read

$$\dot a = 0\,, \qquad \dot\alpha = \omega(a)\,, \tag{1.2.5}$$

where $\omega(a) = \frac{\partial h}{\partial a}(a) = 2\pi/T(h(a))$.

Finally, by identifying all points in the plane (a,α) whose coordinates α differ by a multiple of 2π, we get the standard cylinder $A\times S^1$. The diffeomorphism (a,α) between the covering spaces obviously induces a diffeomorphism between the two cylinders. The conclusion is that the anulus M is diffeomorphic, in a symplectic way, to the cylinder $A\times S^1$, and that the flow is conjugate to the linear flow (1.2.5) on it. Such a diffeomorphism, too, is called ‘action–angle coordinates’ of the system.

---

* We shall often use the following notation. If $f : A \to B$ and $g : A \to C$ are maps, we denote by $f\times g$ or else by (f,g) the map from A into $B\times C$ defined by $a \mapsto (f(a), g(a))$.

Let us add a few remarks. First of all, the integral appearing in (1.2.4) is the area in the plane (h,τ) shaded in figure 1.3. Since $\mathcal{C}$ (being symplectic) is area–preserving, this integral equals the area inside the closed curve $\gamma(h) = \{(p,\theta) : H(p,\theta) = h\}$ in the pendulum's phase space. Thus, one can write

$$a(h) \;=\; \frac{1}{2\pi}\int_{\gamma(h)} p d\theta\,,$$

which is the usual expression of the action as integral of the symplectic one–form along the orbit.

Second, equation (1.2.3) expresses the so called *period–energy* relation: the period of an orbit equals the derivative of the enclosed area with respect to the energy.

And finally, let us stress that the action–angle coordinates are coordinates 'adapted' to the foliation by the periodic orbits. They are manifestly not defined on the singularities of such a foliation, that is, on the equilibria and on the separatrices: on such orbits, an angle cannot be defined at all.

1.3 The rotator.

Consider a rigid body having a fixed axis (or equivalently, a disk in the plane with a fixed point), with no external forces acting on it. The configuration of the system is determined by an angle $\varphi \in S^1$, so the phase space is the cylinder $(p,\varphi) \in \mathbb{R}\times S^1$. The Hamilton function is

$$H(p,\varphi) \;=\; \frac{p^2}{2I}\,,$$

where I is the moment of inertia. The system is already given in action–angle coordinates. The equations of motion are

$$\dot p \;=\; 0\,, \qquad \dot\varphi \;=\; \omega(p)$$

where $\omega(p) = p/I$. So, all motions with $H \neq 0$ are periodic; the corresponding orbits are the circles $p = \mathit{const}$ on the cylinder. Note that the circle $p = 0$ is not a single orbit, but the union of all of the equilibria of the system. Since the frequency $\omega(p)$ is not constant, the system is anisochronous.

1.4 Two uncoupled harmonic oscillators.

We consider now systems with more than one degree of freedom. To begin with, we consider two uncoupled harmonic oscillators, whose Hamilton function is

$$H(p_1,p_2,q_1,q_2) \;=\; H_1(p_1,q_1) \;+\; H_2(p_2,q_2)\,, \qquad (p,q)\in\mathbb{R}^4\,,$$

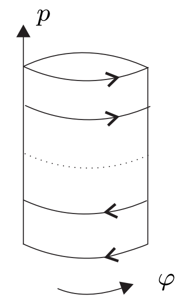


Figure 1.4

where $H_i(p_i, q_i) = \omega_i(p_i^2 + q_i^2)/2$, ω_1 and ω_2 being two positive constants. As is clear, the two functions H_1 and H_2 are independent integrals of motion.† The common level set

$$\begin{aligned} M_{h_1,h_2} &= \{(p_1, q_1, p_2, q_2) : H_1(p_1, q_1) = h_1, H_2(p_2, q_2) = h_2\} \\ &= \{(p_1, q_1) : p_1^2 + q_1^2 = 2h_1/\omega_1\} \times \{(p_2, q_2) : p_2^2 + q_2^2 = 2h_2/\omega_2\} \end{aligned}$$

is the product of two circles, i.e., a two–dimensional torus, if $h_1 > 0$ and $h_2 > 0$, a circle if either $h_1 = 0$ or $h_2 = 0$, and a point if $h_1 = h_2 = 0$; see figure 1.5.

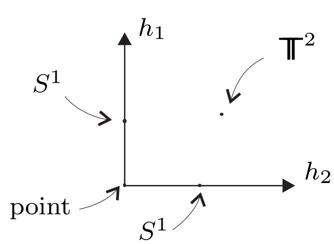


Figure 1.5

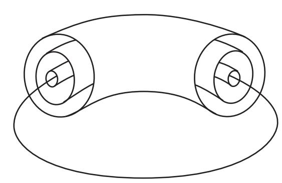


Figure 1.6

Let us restrict ourselves to the subset of the phase space where both H_1 and H_2 are positive, which is foliated by the invariant two–dimensional tori M_{h_1,h_2}. As coordinates 'adapted' to such a foliation one can take the functions H_1 and H_2 (which give coordinates transversal to the level sets) and, as coordinates on the level sets, the times τ_1 and τ_2 of the Hamiltonian flows of H_1 and, respectively, H_2; the latter choice is suggested by the fact that the torus $H_1 = h_1$, $H_2 = h_2$ is the product of two circles, each of which has the time τ_i as coordinate on it.

† By saying that k integrals of motion are independent, without any further specification, we mean that their differentials are linearly independent in some 'significant' open subset of the phase space, to which the consideration is tacitly restricted (and which should be determined case by case). Thus, such a set is foliated by the common level sets of the k integrals of motion, which are invariant submanifolds of codimension k.

In this way, after a normalization of the periods, one arrives at the action–angle coordinates $(a_1, a_2, \alpha_1, \alpha_2) \in \mathbb{R}^2 \times \mathbb{T}^2$, where each pair (a_i, α_i) is related to the corresponding coordinates (p_i, q_i) as in (1.1.1). These coordinates are defined in the whole subset of the phase space foliated by tori of dimension two, but not on the singularities of this foliation, which are located 'over' the lines $h_1 = 0$ and $h_2 = 0$. Locally, around a periodic orbit corresponding to the vanishing of just one of the two functions H_i, the foliation by the level sets of the map $H_1 \times H_2$ looks like in figure 1.6.

In action–angle coordinates, the Hamilton function is $\omega_1 a_1 + \omega_2 a_2$, so the equations of motion are

$$\dot{a}_1 = 0\,, \qquad \dot{a}_2 = 0\,, \qquad \dot{\alpha}_1 = \omega_1\,, \qquad \dot{\alpha}_2 = \omega_2\,,$$

and the system is isochronous. Note that motions are periodic if and only if the ratio ω_1/ω_2 is rational; otherwise, they are 'quasi–periodic' and fill densely every invariant torus $\mathbb{T}^2$ (Dirichlet's theorem, see chapter 3). The different between these two cases is very important: we shall come back on it in section 1.7.

Exercise: Prove that, if ω_1 and ω_2 are both positive (negative), then every energy surface $H(p,q) = h > 0$ is a three–dimensional sphere, and the level sets of H_1 and H_2 define a foliation of it by two–tori, whose only singularities are the two circles corresponding to the periodic orbits with, respectively, $H_1 = h$ and $H_2 = h$.

1.5 The spherical pendulum.

We now consider a point on the surface of a two–dimensional sphere, under the effect of gravity. The phase space is the cotangent bundle T^*S^2. Using spherical coordinates $(\varphi, \theta) \in S^1 \times (0, \pi)$ on the sphere, the Hamilton function is

$$H(p_\varphi, p_\theta, \varphi, \theta) \;=\; \frac{p_\theta^2}{2} \;+\; \frac{p_\varphi^2}{2\;sin^2\theta} \;+\; cos\,\theta\,,$$

(some care has of course to be posed to the presence of the singularities at $\theta = 0$ and $\theta = \pi$ of the coordinate system). This system has two independent integrals of motion: the Hamilton function H and the component J of the angular momentum along the vertical axis – that is, in the chosen coordinate system, p_φ. We want to determine the topology of the common level sets $H = h$, $J = j$, which (as far as $\theta \neq 0, \pi$) are given by

$$M_{h,j} \;=\; \big\{(\varphi, p_\varphi) : p_\varphi = j\big\} \;\times\; \left\{(\theta, p_\theta) : \frac{p_\theta^2}{2} + V_j(\theta) = h\right\} \tag{1.5.1}$$

where $V_j(\theta) = \frac{j^2}{2\;sin^2\theta} + cos\,\theta$ is the so–called effective (or amended) potential.

Let us first consider the case $j \neq 0$. The graph of $V_j(\theta)$, in the interval $(0, \pi)$, is shown in figure 1.7. It has a single minimum $\theta_0(j)$, which is always $> \pi/2$ (actually, it tends to $\pi/2$ if $j \to \infty$ and to π if $j \to 0$). Thus, the curve $p_\theta^2/2 + V_j(\theta) = h$ in the plane

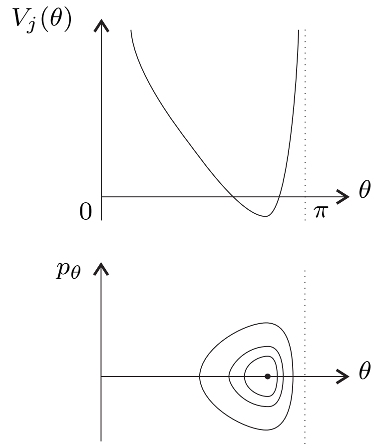


Figure 1.7

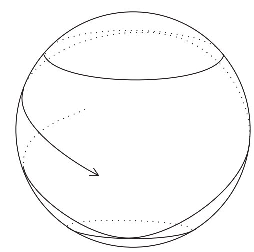


Figure 1.8

(θ, p_θ) is (diffeomorphic to[†]) a circle if $h > V_j(\theta_0(j))$ and is a point if $h = V_j(\theta_0(j))$; in both cases, the system does not pass through the singularities of the spherical coordinates, so the conclusions are consistent. We conclude that $M_{h,j}$ is the two–dimensional torus $\{\varphi \in S^1\} \times \{p_\theta^2/2 + V_j(\theta) = h\}$ or, if h equals the minimum of $V_j(\theta)$, it is the circle $\{\varphi \in S^1\}$. The latter (limit) case corresponds to the periodic motions which, for each value of the energy, take place in a horizontal plane. In all other motions, both coordinates vary; see figure 1.8. As it turns out, one can introduce suitable coordinates in which these motions appear as linear motions on the two–dimensional torus; this follows from the Liouville–Arnol'd theorem, but can also be verified directly, as explained below.

If $j = 0$, the pendulum moves in a vertical plane $\varphi =$ const, just as the pendulum of section 1.2; the only difference is that now the plane can have any orientation. So, the invariant set $M_{h,0}$ is still a two–dimensional torus if $h \neq \pm 1$ (i.e., if the pendulum either librates or rotates in a vertical plane), while it is a point if $h = -1$ (the lower, stable equilibrium) and, if $h = 1$, it is a non–regular two–dimensional surface, which consists of the unstable equilibrium and of all its homoclinic connections. (This surface is obtained by rotating an eight–shaped figure about its middle point– or equivalently, by identifying the north and south poles of a sphere). So, we arrive at the classification of the invariant surfaces shown in figure 1.9. There, the shaded region is the set of the possible values of H and J.

Exercise: [1] Prove that the range of the ("energy–momentum") map $H \times J : T^*S^2 \to \mathbb{R}^2$ is the region $h \geq \hat{h}(j)$, where $\hat{h}(j)$ is the curve described by the parametric equations

$$j(u) \;=\; u - \frac{1}{u^3}\,, \qquad h(u) \;=\; \frac{u^2}{2} - \frac{3}{u^2}\,, \qquad |u| > 1$$

(Sketch: The range of $J \times H$ is the region where $h \geq V_j(\theta_0(j))$. The points (j, h) of the curve $h = V_j(\theta_0(j))$ are the solutions of the system of equations $h = V_j(\theta)$, $dV_j(\theta) = 0$. Express them in terms of the parameter $u = \pm[1 - \cos\theta_0(j)]^{-1/2}$).

† In the following, we shall often call torus something which is only diffeomorphic to the torus.

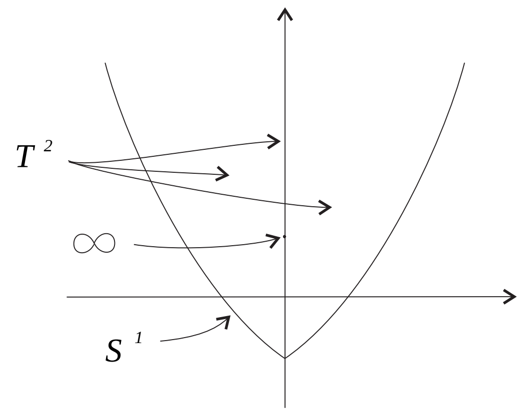


Figure 1.9

Quasi–periodicity of motions. We construct here coordinates in which motions (with $p_\varphi \neq 0$) are linear [2]. Let us use on every torus $M_{h,j}$ the coordinates (φ, τ_1), where $\tau_1 = \tau_1(h, j, \theta)$ is the time along the flow of the system defined by the Hamilton function $p^2/2 + V_j(\theta)$ in the plane (θ, p_θ), normalized to 2π. In these coordinates, the equations of motion are

$$\dot\tau_1 \;=\; \frac{2\pi}{T_1(h,j)}\,, \qquad \dot\varphi \;=\; \frac{j}{\sin^2\theta(\tau_1,h,j)}\,,$$

and the solution $t \mapsto \varphi(t)$ is not linear. Observe however that, denoting $g(\tau_1, h, j) = j/\sin^2\theta(\tau_1, h, j)$, one has (omitting the dependence on h and j)

$$\varphi(t) - \varphi(0) \;=\; \int_0^t g(\tau_1(t))dt \;=\; \frac{T_1}{2\pi}\int_0^{\tau_1(t)} g(\tau)d\tau \;=\; \frac{T_1}{2\pi}\,\overline{g}\,\tau_1(t) \;+\; \widetilde{g}(\tau_1(t))$$

where $\overline{g} = \overline{g}(h, j)$ is the average of $g(\tau_1, h, j)$ over $\tau_1 \in S^1$, and $\widetilde{g}(\tau_1) = \widetilde{g}(\tau_1, h, j)$ is a 2π–periodic function of τ_1 (this can be verified, for instance, by a Fourier expansion). Thus, taking as new coordinates on the torus $M_{h,j}$ the functions τ_1 and $\tau_2 = \varphi - \widetilde{g}(\tau_1, h, j)$ (on every torus, this is just a shift of the origin of the angle), the flow is given by

$$\dot\tau_1 \;=\; \frac{2\pi}{T_1(h,j)}\,, \qquad \dot\tau_2 \;=\; \overline{g}(h,j)\,,$$

and is linear (however, the coordinates (H, J, τ_1, τ_2) are not symplectic). Note that motions on the tori $M_{h,j}$ are periodic or quasi–periodic, depending on whether the ratio $2\pi/(T_1\overline{g})$ is rational or irrational.

1.6 Kepler and central forces. Degeneracy

The Kepler system is the mass point under the influence of a Newtonian force field. The Hamiltonian is

$$H(p,q) \;=\; \frac{\|p\|^2}{2} \;-\; \frac{1}{\|q\|}\,, \qquad q \in \mathbb{R}^3 \setminus \{0\}\,,\; p \in \mathbb{R}^3\,.$$

We first consider the planar Kepler problem, assuming that the point is constrained to the plane $q_3 = 0$. Using polar coordinates in the plane, the Hamilton function is

$$H(r, p_r, p_\varphi) \;=\; \frac{p_r^2}{2} + \frac{p_\varphi^2}{2\,r^2} - \frac{1}{r}\,, \qquad (r, p_r) \in \mathbb{R}_+ \times \mathbb{R}\,, \quad (\varphi, p_\varphi) \in S^1 \times \mathbb{R}\,.$$

The energy H and the component $J = p_\varphi$ of the angular momentum in the direction orthogonal to the plane are independent first integrals. The level set $H = h$, $J = j$ is

$$M_{h,j} \;=\; \{\varphi \in S^1\} \times \{p_r^2/2 + V_j(r) = h\}\,,$$

where $V_j(r) = j^2/2r^2 - 1/r$. Drawing the graph of the amended potential $V_j(r)$ one sees that such level sets are compact iff $h < 0$ and $j \neq 0$ and that they are (topologically) two–dimensional tori if $0 > h > -1/2j^2$ and circles if $h = 1/2j^2$; see figure 1.10. So, let us restrict to the subset M of the phase space where $h < 0$ and $j \neq 0$.

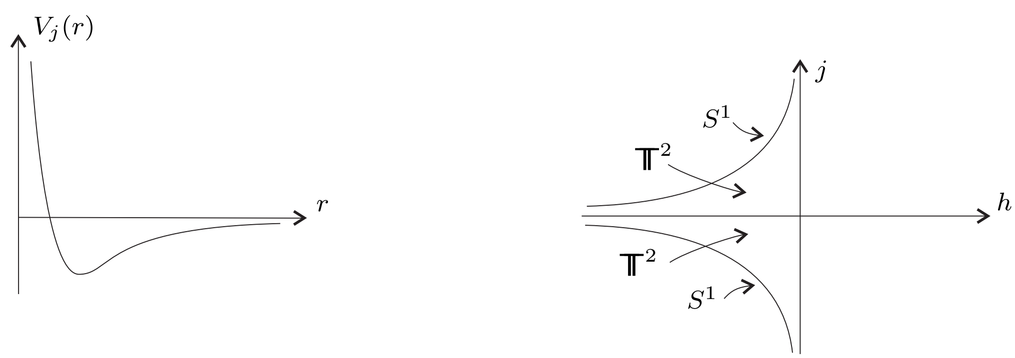


Figure 1.10

It is a well known fact that, in M, *all* orbits are ellipses, so all motions are periodic. This is due to the existence of an additional integral of motion, besides H and J, namely, the so–called *Laplace vector*†

$$A(p, q) \;=\; p \times (q \times p) - \frac{q}{\|q\|}\,. \tag{1.6.1}$$

As a consequence, there is a foliation by invariant circles of the subset M of the phase space, which is finer than that by two dimensional tori, in the sense that each two–torus is union of such circles.‡

Specifically, every ellipse in M has semimajor axis $1/(-2h)$ and eccentricity $\sqrt{1 + 2hj^2}$. The Laplace vector A belongs to the plane of the orbit ($A_3 = 0$), is directed towards the perihelion of the ellipse, and its lenght equals the eccentricity

$$A_1^2 + A_2^2 \;=\; 1 + 2hj^2\,. \tag{1.6.2}$$

† In the literature, it is also named after Runge and Lenz. For some historical notes see [3].

‡ Moreover, the circular orbits corresponding to $j = \pm 1/\sqrt{-2h}$ are regular leaves of the foliation by circles, but not of course of that by two–tori. So, the subset of the phase space foliated by circles is larger than that foliated by two–tori.

Systems like Kepler, whose motions are all quasi–periodic on tori of dimension less than one–half the dimension of the phase space, are called *degenerate* (or *superintegrable*). Various, and important, classical systems are degenerate.

A first example is the point in a central force field (not necessarily Keplerean) in the three–dimensional space. The phase space has dimension six, and the energy and the angular momentum vector are constants of motion. Their common level sets are submanifolds of dimension two (at most). If compact, they are tori (prove it, as an exercise), so motions are quasi–periodic with two–frequencies.

A theorem by Bertrand (see e.g. [4]) states that, among all central force fields, the only ones which possess an additional independent integrals of motion are Kepler and the elastic force: these systems are therefore "completely degenerate", and all their bounded orbits are circles. Correspondingly, among all central forces in the plane, the only degenerate ones are Kepler and the elastic force; we consider in detail the latter case in the next section.

1.7 Degenerate and not degenerate harmonic oscillators

As observed in section 1.4, all motions of the two uncoupled oscillators there considered are either periodic or quasi–periodic, depending on the ratio between the two frequencies. As there mentioned, these cases are deeply different:

- If ω_1/ω_2 is rational, then all motions are periodic. Therefore, every invariant torus M_{h_1,h_2} is union of periodic orbits, that is, it is in turn foliated by invariant circles. This fact reflects the existence of a third integral of motion, independent of H_1 and H_2, so that *the system is degenerate*. In fact, if $\omega_1/\omega_2 = j_1/j_2$ with relatively prime numbers j_1 and j_2, then any 2π–periodic function of $j_1\alpha_2 - j_2\alpha_1$ is an integral of motion independent of H_1 and H_2.†

For instance, if $\omega_1 = \omega_2$, three independent first integrals are

$$\begin{aligned} F_1(q,p) &= q_1q_2 + p_1p_2 = 2\sqrt{a_1a_2}\; cos(\alpha_2-\alpha_1) \\ F_2(q,p) &= p_1q_2 - p_2q_1 = 2\sqrt{a_1a_2}\; sin(\alpha_2-\alpha_1) \\ F_3(q,p) &= \frac{1}{2}(p_1^2+q_1^2-p_2^2-q_2^2) = a_1 - a_2 \end{aligned} \tag{1.7.1}$$

(and are globally defined in $\mathbb{R}^4$). Note that these integrals are not independent of the Hamiltonian, since

$$F_1^2 + F_2^2 + F_3^2 = H^2\,. \tag{1.7.2}$$

† This function is defined only in the subset of the phase space where the action–angle coordinates are defined; however, inthe special case $\omega_1 = \omega_2$ considered below, this function can be continuously extended to the whole subset of the phase space fibered by the periodic orbits.

- If ω_1/ω_2 is not rational, instead, every trajectory fills densely an invariant two–dimensional torus. Therefore, there is no (continuous) integral of motion independent of H_1 and H_2 (why?). Hence, *the system is not degenerate.*

Remark: If ω_1/ω_2 is not rational, one could expect that it is not possible to decompose the two–dimensional invariant tori so as to get a *finer* invariant foliation of the phase space. This is a subtle point, in which the exact (technical) definition of foliation is crucial. In fact, there do exists a finer invariant foliation of the phase space, namely, the foliation whose leaves are the periodic orbits. However, since the two–dimensional tori are closure of orbits, there is no finer invariant foliation whose leaves are level sets of a differentiable map. Hence, the invariant two–dimensional tori are the finest invariant *simple foliation* of the phase space (a simple foliation is a foliation whose leaves are the level sets of a submersion).

Complement: The fibration by the periodic orbits for the isotropic oscillator. We consider now the case of an isotropic oscillator, for which $\omega_1 = \omega_2$; this system is also referred to as $1:1$ *resonance.* We want to study the foliation by the periodic orbits of an energy surface $H(p,q) = h > 0$. Specifically, we show that the space of the leaves of such a foliation (obtained by identifying all points on a same orbit) is diffeomorphic to S^2, and that this foliation is in fact the Hopf fibration $S^3 \to S^2$.

First of all, identify $R^4 \ni (q,p)$ with $\mathbb{C}^2 \ni z = (z_1, z_2)$ via $z_j = q_j + ip_j$. Under this identification, the orbit through a point $(p,q) \neq (0,0)$ becomes the circle $\{(e^{it}z_1, e^{it}z_2) : t \in \mathbb{R}\}$, which lies on the sphere $S^3 = \{z' \in \mathbb{C}^2 : \|z'\| = \|z\|\}$, where $\|z\|^2 = |z_1|^2 + |z_2|^2$. Thus, denoting $[z]$ the orbit containing the point z, the orbits are the level sets of the map $z \mapsto [z]$.

Remember now that the complex projective space $\mathbb{C}P^1$ is the quotient of $\mathbb{C}^2 \setminus \{(0,0)\}$ under the equivalence relation

$$(z_1, z_2) \sim (\lambda z_1, \lambda z_2) \quad \forall \lambda \in \mathbb{C} \setminus \{0\}\,.$$

The restriction to S^3 of this equivalence relation is $(z_1, z_2) \sim (\lambda z_1, \lambda z_2)$ with $|\lambda| = 1$ and coincides with the equivalence relation of 'belonging to an orbit'. Hence, $z \mapsto [z]$ is a surjection from S^3 onto $\mathbb{C}P^1$. This map is indeed a fibration, and is called 'Hopf fibration'. Its fibers are obviously circles.

In order to conclude the proof, we now show that $\mathbb{C}P^1$ is diffeomorphic to S^2. To this end, we choose an atlas with two charts for $\mathbb{C}P^1$ and an atlas with two charts for S^2 and we show that there are diffeomorphisms among each pair of coordinate domains which 'agree' on the intersection of the domains, so they define a global diffeomorphism.

Specifically, $\mathbb{C}P^1$ has a (complex) atlas made of two charts, with coordinates $\zeta_+([(z_1,z_2)]) = z_1/z_2$ (defined everywhere but in $[(1,0)]$) and, respectively, $\zeta_-([(z_1,z_2)]) = z_2/z_1$ (defined everywhere but in $[(0,1)]$); both coordinate systems are onto $\mathbb{C}$.

On the other hand, $S^2 = \{\xi \in \mathbb{R}^3 : \xi_1^2 + \xi_2^2 + \xi_3^2 = 1\}$ has an atlas made of two charts with stereographic coordinates $(x_\pm, y_\pm) : S^2 \setminus \{(0,0,\pm 1)\} \to \mathbb{R}^2$, defined by

$$x_\pm = \frac{\xi_1}{1 \mp \xi_3}\,, \qquad y_\pm = \frac{\xi_2}{1 \mp \xi_3}\,;$$

both coordinate systems are onto $\mathbb{R}^2$, and their transition functions are such that $(x_+ + iy_+)(x_- + iy_-) = 1$.

Therefore, the two maps

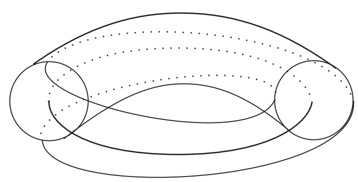
$$\zeta_+ = x_+ + iy_+, \qquad \zeta_- = x_- + iy_-$$

are the local representatives of a diffeomorphism of S^2 onto $\mathbb{C}P^1$. (This map can be described as follows: it maps the north pole $(0,0,1) \in S^2$ into the point $[(0,1)] \in \mathbb{C}P^1$ and, for any $(x,y) \in \mathbb{R}^2$, it maps the point of the sphere whose stereographic coordinates are $x_+ = x$, $y_+ = y$ into the point $[(x+iy, 1)]$).

Exercise: Generalize the above treatment to the case of a d–dimensional isotropic oscillator. (Answer: the fibration by the periodic orbit of every (positive) level set of the Hamiltonian is the Hopf fibration $S^{2d-1} \to \mathbb{C}P^{2d-2}$).

Complement: The resonance 2 : 1 The foliation by invariant tori (or periodic orbits) of an integrable system has often, but not always, the structure of a fibration. The simplest counterexample is provided by two harmonic oscillators with frequency ratio $\omega_1/\omega_2 = 2$. In this case, if one excludes the equilibrium, all orbits have the same period T but the special orbits in which only the oscillator of higher frequency moves, which have period $T/2$. In every submanifold of constant energy there is exactly one of these special orbits (which is based on the line $h_2 = 0$ of figure 1.5). This special orbit is surrounded by two–dimensional tori filled by periodic orbits which close themselves after two turns. Therefore, every surface transversal to such a special orbit intersects any other orbit in *two* points, see figure 1.11. This implies that, near the orbits of half period, the foliation by periodic orbits does not possess a local section. Hence, it is not a fibration (it has the structure of a 'Seifert foliation', see [5]). One gets however a fibration by restricting himself to the subset of the phase space where $H_2 > 0$.

Figure 1.11



1.8 Euler–Poinsot

One further example of degenerate system is the 'free' rigid body, or Euler–Poinsot system, that is, the rigid body with a fixed point and no external forces acting on it.† This system has three degrees of freedom, so its phase space has dimension six. As in the case of the point in a central force field, the kinetic energy H and the angular momentum vector in space m^s are independent constants of motion. We want to provide some evidence about the fact that their common level sets are (generically) two–dimensional tori.

Let us first recall some elementary facts about the Euler–Poinsot system (for a complete treatment, see [4,6,7]). Its phase space is the cotangent bundle $T^*SO(3)$, which can be identified with $SO(3) \times \mathbb{R}^3$. The identification can be realized with the aid of two orthogonal

† More precisely, the external forces must have null moment relative to the fixed point. This is the case for a heavy rigid body if the fixed point is the center of mass.

reference frames $\mathcal{B}_b = \{e_1, e_2, e_3\}$ and $\mathcal{B}_s^* = \{e_x^*, e_y^*, e_z^*\}$ attached, respectively, to the body and to the space. We assume that the axes of the body frame $\mathcal{B}_b$ are directed like the axes of inertia of the body and we denote by I_1, I_2, I_3 the corresponding moments of inertia. The spatial frame is instead completely arbitrary (in the absence of external forces there are no preferred directions in space). We denote by $u^b = (u_1, u_2, u_3)$ and $u^s = (u_x^*, u_y^*, u_z^*)$ the representatives of a vector u in the two bases $\mathcal{B}_b$ and $\mathcal{B}_s^*$, respectively. Thus, for instance, $e_z^s = e_3^s = (0, 0, 1)$.

The identification of $T^*SO(3)$ with $SO(3) \times \mathbb{R}^3$ can be done in such a way that the state of the system is represented by a point $(\mathcal{R}, m^b)$, where $\mathcal{R} \in SO(3)$ is the matrix which determines the mutual orientation of the two frames, defined by $(e_1)^s = \mathcal{R}(e_x^*)^s$, $(e_2)^s = \mathcal{R}(e_y^*)^s$, $(e_3)^s = \mathcal{R}(e_z^*)^s$, and $m^b = (m_1, m_2, m_3) \in \mathbb{R}^3$ is the representative of the angular momentum vector m in the body base $\mathcal{B}_b$. (This is the so–called body representation).

The kinetic energy H and the angular momentum vector *in space*, that is, the representative $m^s = (m_x^*, m_y^*, m_z^*)$ of the angular momentum vector in the frame $\mathcal{B}_s^*$, as functions on the phase space $SO(3) \times \mathbb{R}^3$, are

$$H(\mathcal{R}, m^b) \;=\; \sum_i \frac{m_i^2}{2I_i}\,, \qquad\qquad m^s(\mathcal{R}, m^b) \;=\; \mathcal{R} m^b$$

(H is actually independent of $\mathcal{R}$).

Proposition 1.1 *If $\hat{m}^s = (\hat{m}_x, \hat{m}_y, \hat{m}_z) \in \mathbb{R}^3$ and $h \in \mathbb{R}$ are such that $h \neq \|\hat{m}^s\|^2/2I_i$ for $i = 1, 2, 3$, then the level set $H = h$, $m^s = \hat{m}^s$ is diffeomorphic to a two–dimensional torus.*

Sketch of the proof. We verify only that the considered level set is in bijective correspondence with $\mathbb{T}^2$. Observe first of all that the vector m^b belongs to the intersection of the ellipsoid $H(m^b) = h$ with the sphere $\|m^b\| = \|\hat{m}^s\|$. As shown in figure 1.12, if $\|m^b\|$ is not equal to the length of one of the semi–axes of the ellipsoid (that is, if $h \neq \|\hat{m}^s\|^2/2I_i$), then sphere and ellipsoids intersect in a closed curve. The claim now follows from the fact that, for each fixed $m^b \neq 0$, the set of matrices $\{\mathcal{R} \in SO(3) : m^s(\mathcal{R}, m^b) = \hat{m}^s\}$ is diffeomorphic to a circle, too. Indeed, let $\mathcal{R}_0 \in SO(3)$ be the matrix corresponding to the rotation about the axis directed like $m^b \times \hat{m}^s$ which brings m^b onto $\hat{m}^s$ anticlockwise (we take $\mathcal{R}_0$ to be the identity if m^b is parallel to $\hat{m}^s$). Then, every matrix $\mathcal{R} \in SO(3)$ such that $\mathcal{R} m^b = \hat{m}^s$ can be written uniquely as $\mathcal{R} = \widehat{\mathcal{R}}_\varphi \mathcal{R}_0$, where $\widehat{\mathcal{R}}_\varphi$ is the rotation of an angle $\varphi \in S^1$ about the (fixed) direction of $\hat{m}^s$. ∎

If $h \neq 0$, the 'exceptional' level sets where $h = \|m^s\|/2I_i$ consist of the steady rotations, that is, the periodic motions in which m^b is parallel to one of the axes of inertia, and of the four 'separatrices' connecting the (unstable) steady rotations about the middle axis of inertia. For $h = 0$, one obviously has the equilibria. All these exceptional level sets are singular leaves of the foliation by the invariant two–dimensional tori.

Exercises: (i) Show that the level set $H = h$, $m^s = \hat{m}^s$ is diffeomorphic to:

- $SO(3)$, if $\hat{m}^s = 0$;
- S^1, if $\hat{m}^s = 0$ but $\hat{m}^s \times e_i = 0$ for some $i = 1, 2, 3$;
- The union of four connected non–compact manifolds diffeomorphic to $(-1, 1) \times S^1$, if $h = \|\hat{m}^s\|^2/2I_2$ but $\hat{m}^s$ is not parallel to e_2 (we are assuming here that the moments of inertia satisfy $I_1 < I_2 < I_3$ or $I_1 > I_2 > I_3$).

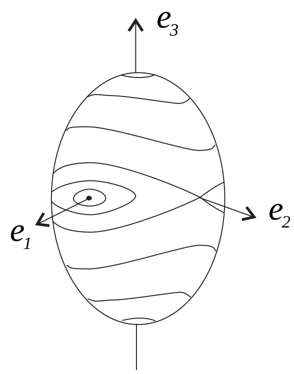


Figure 1.12

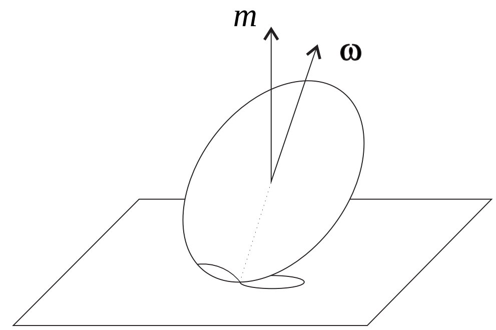


Figure 1.13

(ii) Assume that the body is (kinematically) symmetric around its axis e_3, that is, $I_1 = I_2$. Show that all the level sets $H = h$, $m^s = \hat{m}^s$ are diffeomorphic to $\mathbb{T}^2$, except when $\hat{m}^s = 0$ and when $\hat{m}^s \times e_3^s = 0$.

A deep insight into the Euler–Poinsot system is provided by Poinsot's classical geometric description. Poinsot's theorem can be stated as follows: think of the ellipsoid of inertia of the body relative to the fixed point as having its center in that point, and as being rigidly connected to the body; then, during the motion of the system, the ellipsoid of inertia rolls without sliding on a fixed plane, orthogonal to the angular momentum vector (figure 1.13). The proof of this theorem can be found in [4].

During the motion of the body, the point of contact between the ellipsoid of inertia and the plane draws a curve on the ellipsoid, called the *polhode*, and another curve on the plane, called the *erpolhode*. For all motions except those with $H = \|m^s\|/2I_i$, $i = 1, 2, 3$, the polhode is a closed curve (it is the locus of points of the ellipsoid which have a given distance from its center). The erpholode is contained within an anulus, and can be either open or closed.

Poinsot's theorem allows one to visualize the invariant tori. Given H and m^s, the position in space of the ellipsoid of inertia (which determines the configuration of the system) is determined by specifying the point P_E of the ellipsoid and the point P_P of the plane which are in contact. It is easy to see that, if $H \neq \|m^s\|/2I_i$, then the set of all possible (P_E, P_P) is a two–dimensional torus. In fact, P_E belongs to the polhode, which is a circle. On the other hand, the distance of P_P from the foot of m^s is uniquely determined by P_E, so that P_P is determined by its azimuth —i.e., by the position on the unit circle of the unit vector pointing from the foot of m^s towards P_P.

As it turns out, one can coordinatize the polhode and the unit circle in the fixed plane with coordinates which move with constant speed during the motion of the body. (The idea of this construction is the same as illustrated for the spherical pendulum). Moreover, the system turns out to be anisochronous, and there are both invariant tori supporting periodic motions and invariant tori supporting non–periodic, and in fact dense, motions.

Exercise: Show that, if a dense subset of the invariant two–dimensional tori are closure of trajectories, then there is no finer invariant 'simple' foliation of the phase space.

Motions can be regarded as composed by a 'proper rotation' of the body about an instantaneous axis which, in the general non–symmetric case, is not fixed in the body, and of a 'precession' (or 'nutation') of such an axis in space, about the fixed direction of the angular momentum vector m^s. Again, the system is degenerate: the third frequency would correspond to a precession of m in space, but is zero because of the invariance of the system under rotations in space, which reflects the absence of external forces.

1.9 The Lagrange top

As a last example, we consider a heavy rigid body with a fixed point, whose ellipsoid of inertia relative to the fixed point is symmetric (about the axis of inertia e_3, so that $I_1 = I_2$), and assume that the center of mass of the body lies on the symmetry axis e_3. This is the so called Lagrange top and, together with Euler–Poinsot, is one of the few known cases in which the equations of motion of a rigid body are integrable.†

By symmetry considerations, one knows that the energy H, the projection m_z of the angular momentum vector m on to the 'vertical' (the axis e_z fixed in space, parallel to the gravity) and the projection m_3 of m on to the symmetry axis of inertia e_3 are integrals of motion. Once again, their common level sets are (generically) three–dimensional tori.

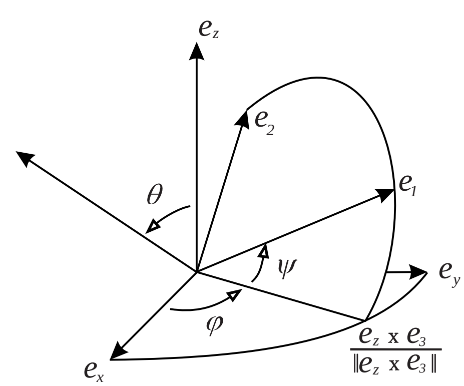


Figure 1.14: The Euler angles

† Actually, there is only one more known case in which they are integrable for all initial conditions, that is the Kowalevskaja case. There are however many cases in which the system is integrable in a submanifold of the phase space of positive codimension.

In order to verify this, let us introduce the Euler angles $(\varphi, \psi, \theta) \in S^1 \times S^1 \times (0, \pi)$, whose definition is shown in figure 1.14 (the axis e_x, e_y, and e_z are fixed in space, while e_1, e_2, and e_3 are the axes of inertia of the body relative to the fixed point O). We skip on all questions relative to the singularities of the Euler angles at $\theta = 0$ and $\theta = \pi$, which can be dealt without particular difficulties. The moments conjugate to the Euler angles are $p_\varphi = m_z$, $p_\psi = m_3$, and p_θ (which is equal to the projection of m along the nodal line parallel to $e_z \times e_3$). The Hamilton function is

$$H = \frac{p_\theta^2}{2I_1} + \frac{p_\psi^2}{2I_3} + \frac{(p_\varphi - p_\psi \cos\theta)^2}{2I_1 \sin^2\theta} + \cos\theta \, . \tag{1.9.1}$$

The level surface $H = h$, $m_z = \hat{m}_z$, $m_3 = \hat{m}_3$ is then

$$M_{h,\hat{m}_z,\hat{m}_3} = \{\varphi \in S^1\} \times \{\psi \in S^1\} \times \left\{ (\theta, p_\theta) \, : \, \frac{p_\theta^2}{2I_1} + V_{\hat{m}_z,\hat{m}_3}(\theta) = h - \frac{\hat{m}_3^2}{2I_3} \right\} ,$$

where

$$V_{\hat{m}_z,\hat{m}_3}(\theta) = \frac{(\hat{m}_z - \hat{m}_3 \cos\theta)^2}{2I_1 \sin^2\theta} + \cos\theta \, .$$

Drawing the graphic of the effective potential $V_{\hat{m}_z,\hat{m}_3}$, one sees that the coordinate θ performs (generically) a periodic motion, with period $T(h, \hat{m}_z, \hat{m}_3)$. Thus, one concludes that $M_{h,\hat{m}_z,\hat{m}_3}$ is (generically) a torus of dimension three. It could be seen that the motions are (generically) linear on such tori, with rationally independent frequencies. Thus, at variance from the Euler–Poinsot case, motions have now three frequencies: the introduction of the external potential destroys an integral of motion and removes the degeneracy.

Chapter 2

The Liouville–Arnol'd Theorem

The picture emerging from the examples of the previous chapter is that a subset of the phase space of an integrable Hamiltonian system with d degrees of freedom is fibered by invariant tori of a certain dimension $n \leq d$, which support quasi–periodic motions. There are, in general, also motions of different types (for instance unbounded motions, motions on separatrices, equilibria, etc.) but we shall not deal with them. The Liouville–Arnol'd theorem, and some generalizations of it, explain this situation by relating it to the existence of a sufficiently high number of integrals of motion with certain properties and assure the existence of 'action–angle' coordinates adapted to such a fibration.

In this chapter, we state and prove the Liouville–Arnol'd theorem. Since this theorem is at a large extent a statement in symplectic geometry, it is convenient to formulate it without any reference to a Hamiltonian system. The application to 'completely integrable' Hamiltonian systems, which is our main concern, will be considered in the next chapter.

2.1 Statement of the theorem

A. The theorem. We first present the usual formulation of the Liouville–Arnol'd theorem; a slightly more general formulation, which has a more geometric character and is indeed useful, will be presented in subsection B.

Let M_* be a $2d$–dimensional manifold and let $F = F_1 \times \ldots \times F_d : M_* \to \mathbb{R}^d$ be a smooth map. Let R_F be the open subset of M_* where $dF_1(x), dF_2(x), \ldots, dF_d(x)$ are linearly independent; thus, $F|_{R_F}$ is a submersion and the level sets of F, in R_F, are submanifolds of dimension d. We are interested only in the compact connected components of the level sets of F. Thus, for each point $x \in R_F$, we denote by N_x the connected component containing x of the level set $\{x' \in R_F : F(x') = F(x)\}$, and we restrict ourselves to the set

$$C_F = \{x \in R_F : N_x \textit{ is compact}\}.$$

It can be shown that C_F is open in M_* (see [1] or else [8], page 358). Thus, C_F is a $2d$–dimensional manifold foliated by compact connected submanifolds of dimension d. As a matter of fact, *such a foliation is a fibration.* This fact follows from a theorem by Ehresmann, and has an important role in the proof of the Liouville–Arnol'd theorem.†

Examples: i) (The pendulum) Consider the function $H(p,\theta) = p^2/2 + \cos\theta$ on the cylinder $\mathbb{R} \times S^1$. R_H is the whole cylinder but the points $(0,0)$ and $(0,\pi)$. To obtain M_H, we must still remove all what remains of the level set $H(p,\theta) = 1$ since, after the exclusion of $(0,0)$, it is not more compact. For each $(p,\theta) \in M_F$, $N_{p,\theta}$ is a circle (see figure 1.2).

ii) Consider now the planar Kepler system of section 1.6. The system of inequalities $dH \neq 0$, $dp_\varphi \neq 0$, $dH \wedge dp_\varphi \neq 0$ is satisfied everywhere, except where $p_\varphi = 0$ or where $p_r = 0$ (or equivalently, where $p_\varphi^2 = r$). Thus, R_{H,p_φ} consists of the whole phase space except the collisional motions ($p_\varphi = 0$) and the circular orbits ($p_r = 0$). Since the level sets of $H \times p_\varphi$ are compact iff $H < 0$ and $p_\varphi \neq 0$, and since $H = -p_\varphi^2/2$ on circular orbits, one concludes that C_{H,p_φ} is the region of the phase space where $H < 0$, $0 < |p_\varphi| < 1/\sqrt{-2H}$.

iii) For the spherical pendulum, $R_{J,H}$ is the region of the phase space which is projected by $J \times H$ onto the (open) region shown in figure 1.9 and defined in an exercise in section 1.5. (The verification is left as an exercise).

Theorem 2.1 (Liouville–Arnol'd) *Consider a symplectic manifold (M,Ω) of dimension $2d$, and assume that there exists a map $F : M \to \mathbb{R}^d$ whose components $F_1,\ldots,F_d$ are pairwise in involution, that is,*

$$\{F_i, F_j\} = 0\,, \qquad \textit{for all} \quad i,j = 1,\ldots,d\,.$$

Then, for each point $x \in C_F$:

i) *N_x is diffeomorphic to the d–dimensional torus $\mathbb{T}^d$.*

ii) *There exists a neighbourhood U of N_x in C_F and a diffeomorphism*

$$a \times \alpha : U \mapsto \hat{A} \times \mathbb{T}^d$$

where $\hat{A}$ is an open subset of $\mathbb{R}^d$ and $\mathbb{T}^d = \mathbb{R}^d/(2\pi\mathbb{Z})^d$, such that, writing $a = (a_1,\ldots,a_d)$ and $\alpha = (\alpha_1,\ldots,\alpha_d)$:

1) *the restriction of Ω to U is equal to $\sum_i da_i \wedge d\alpha_i$;*

2) *there exists a diffeomorphism $\hat{a}$ of $F(U)$ onto $\hat{A}$ such that $a = \hat{a} \circ F$ (that is, the a_i's are functions of the F_i's alone—see figure 2.1).*

The proof of the theorem is deferred to the next section.

Statement *ii)* of the theorem means that the set C_F is equipped with local systems of *action–angle coordinates* $a \times \alpha$, which are symplectic coordinates 'adapted' to the fibration $F : C_F \to F(C_F)$.† Note that the theorem has both global and local aspects. Global,

† We recall that a fibration is a surjective submersion $\pi : M \to B$ which has the following property: every point b in the base manifold B has a neighbourhood V with a *local trivialization*, that is, a diffeomorphism $g : V \times \pi^{-1}(b) \to \pi^{-1}(V)$. If B is connected, all the fibers $\pi^{-1}(b)$, $b \in B$, are diffeomorphic. Ehresmann's theorem states that *a submersion with compact fibers is a fibration.* For the proof of this theorem see for instance [9] (where the theorem is stated for a proper map — but one recognizes that all what is needed in the proof is the compactness of the level sets.) For the notion of foliation see for instance [10].

† If M is a manifold of dimension $2d$ and $\pi : M \to B$ is a fibration with fibers of dimension n, then a system of local coordinates $(u,v) = (u^1,\ldots,u^{2d-n},v^1,\ldots,v^n)$ in M is *adapted to the fibration* if one has $u(x) = u(x')$ if and only if $\pi(x) = \pi(x')$ (it is not required that $v^1,\ldots,v^k$ are global coordinates on the fibers). With a little abuse of notation, the u's can be considered as local coordinates on the base manifold.

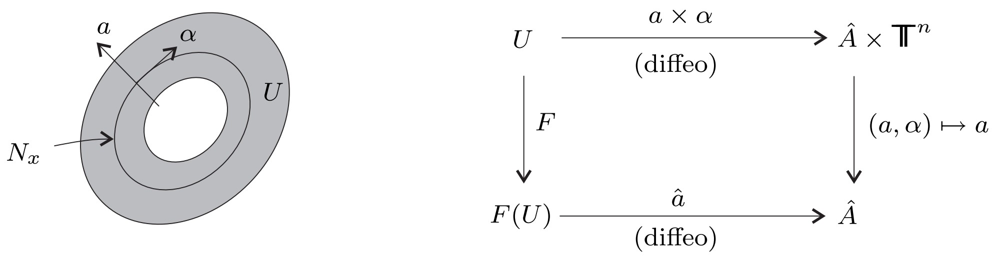


Figure 2.1

because the angles are global coordinates on the (covering of) the tori. Local, because the existence of the action–angle coordinates is assured only in a neighbourhood of each torus; and in fact, there are various important cases in which no single, global system of action–angle coordinates covers all of the set C_F. See e.g. [3].

Remarks: (i) If the level sets of F are not connected (example: the pendulum at energies $H > 1$) one should distinguish between the fibration $F : C_F \to F(C_F)$, whose fibers are unions of tori, and the fibration $\pi : M \to A$ whose fibers are the individual tori, that is, the connected components of the level sets of F (it is immediate to prove that this is indeed a fibration): the submersion π coincides with F in a neighbourhood of each torus, but the bases A and $F(C_F)$ of the two fibrations are different.

(ii) Arnold's original statement of theorem 2.1 made use of an additional hypothesis, namely, that the action coordinates be functionally independent of the functions $F_1, \ldots, F_d$ (see [11,12] and references therein). This hypothesis was subsequently removed by Jost [13] and, independently, by Nekhoroshev [14] and by Markus and Mayer [15]. An account of what was classically known before Arnold's work can be found in chapter 12 of [16]; we shall discuss Liouville's theorem in section 3.2.A. For some historical notes, see also [17].

B. A geometric approach. In the above approach, the emphasis is from the outset on the d functions in involution $F_1, \ldots, F_d$ which are used to describe the fibration by tori. However, the choice of such functions is never unique (any other set of functions $F'_i = F'_i(F_1, \ldots, F_d)$, $i = 1, \ldots, d$, describe the same fibration, and are in involution) and the set C_F depends on such a choice (think for instance of the effect of replacing a function with its square). Therefore, within this approach, the subset of the phase space fibered by the tori is affected by a certain arbitrariness.

Example: Consider the rotator of section 1.3. The set C_H corresponding to the Hamilton function $H = p^2/2$ is the whole cylinder but the circle $p = 0$, which consists of the equilibria. Nevertheless, the circle $p = 0$ is a regular fiber of the fibration by the invariant circles, and the action–angle coordinates (p, φ) do exist in its neighbourhood, too. In this case, this problem can be overcome by considering p, rather than $p^2/2$.

In order to avoid such arbitrariness, one should take as primary object the fibration by the tori. The key observation is that, under the hypotheses of the Liouville–Arnol'd theorem, the submanifolds $F = const$ are *Lagrangian* submanifolds of M.† This is seen by considering the Hamiltonian vector fields $X_{F_1}, \ldots, X_{F_d}$ of the functions $F_1, \ldots, F_d$ and by observing that, in C_F, they are everywhere linearly independent (because the dF_i's are linearly independent) and tangent to each submanifold N_x (since $L_{X_i} F_j = \{F_i, F_j\} = 0$ for all i, j). So, $X_{F_1}, \ldots, X_{F_d}$ form a basis for the tangent spaces of the manifolds N_x, $x \in C_F$. But then, since $\Omega(X_{F_i}, X_{F_j}) = \{F_i, F_j\} = 0$ for all i, j, $\Omega|_{N_x} = 0$.

There is a converse to this situation:

Proposition 2.2 *Let M be a symplectic manifold of dimension $2d$, and $\pi : M \to B$ a fibration with connected Lagrangian fibers. Then, locally, the fibers of π are the level sets of d functions $F_1, \ldots, F_d$ pairwise in involution.*‡

Proof. Let us consider a system of local coordinates $b = (b_1, \ldots, b_d) : V \to \mathbb{R}^d$ on the base B of the fibration, where $V \subset B$ is an open set. Then, the map $F = (F_1, \ldots, F_d) : \pi^{-1}(V) \to \mathbb{R}^d$ defined by

$$F_i \;=\; b_i \circ \pi\,, \qquad i = 1, \ldots, d\,,$$

is a submersion and its level sets coincide with the fibers of π. As we now show, the Lagrangian character of the fibers of π imply that the functions F_i are pairwise in involution. Indeed, if Y is any vector field tangent to the fibers of π, one has $L_Y F_i = 0$ for $i = 1, \ldots, d$ and thus

$$\Omega(X_{F_i}, Y) \;=\; 0\,, \qquad i = 1, \ldots, d\,. \tag{2.1.1}$$

Since this holds for all vector fields Y tangent to the fibers of π, this means that $X_{F_1}, \ldots, X_{F_d}$ are symplectically orthogonal to the fibers of π. Thus, since since the fibers are Lagrangian, $X_{F_1}, \ldots, X_{F_d}$ are tangent to them.⋆ But then, by (2.1.1), $\{F_i, F_j\} = \Omega(X_{F_i}, X_{F_j}) = 0$ for all i and j. ∎

So, the Liouville–Arnol'd theorem is a statement about the geometry of compact Lagrangian fibrations of symplectic manifolds.

C. Non–uniqueness of the action–angle coordinates. We now establish at which extent the action–angle coordinates are determined by the fibration by tori. To this end, we first need to give a precise definition of action–angle coordinates.

Definition 2.1 *Let (M, Ω) be a symplectic manifold of dimension $2d$, and let $\pi : M \to A$ be a fibration with compact, connected, Lagrangian fibers. A local system of* action–angle coordinates of the fibration *is a diffeomorphism*

$$a \times \alpha \;:\; \pi^{-1}(V) \to \hat{A} \times \mathbb{T}^d\,, \qquad \hat{A} \subset \mathbb{R}^d\,,$$

† A submanifold of a $2d$–dimensional symplectic manifold is called *isotropic* if the restriction to it of the symplectic two–form vanishes. Isotropic submanifolds have dimension $n \le d$. *Lagrangian* submanifolds are isotropic submanifolds of the maximum dimension d.

‡ Here, 'locally' means 'locally in the base': each point of A has a neighbourhood V and a map $F : \pi^{-1}(V) \to \mathbb{R}^d$ such that the statement holds.

⋆ Let M be a symplectic manifold, and E_x a subspace of the tangent space $T_x M$ at a point $x \in M$. The *symplectic orthogonal* of E_x is the largest subspace $E_x^\perp$ of $T_x M$ such that $\Omega(v, v') = 0$ for all $v \in E_x$ and all $v' \in E_x^\perp$. As it turns out, a submanifold N of M is isotropic iff $(T_x N)^\perp \subset T_x N$ at each point $x \in N$; in particular, it is Lagrangian iff $(T_x N)^\perp = T_x N$ for all $x \in N$.

where $V \subset A$ is an open set, which is such that

i) $\Omega|_U = \sum_i da_i \wedge d\alpha_i$;

ii) $a(m) = a(m')$ *if and only if the points* m *and* m' *belong to the same fiber of* π.

In the following, we shall denote by $SL_\pm(d, \mathbb{Z})$ the group of $d \times d$ matrices with integer coefficients and determinant ± 1, and by Z^{-T} the inverse of the transpose of a matrix Z.

Proposition 2.3 *Let* $a \times \alpha : U \to \hat{A} \times \mathbb{T}^d$ *and* $a' \times \alpha' : U' \to \hat{A}' \times \mathbb{T}^d$ *be two local systems of action–angle coordinates of a Lagrangian fibration* $\pi : M \to A$. *Assume that* $U \cap U'$ *is non–empty and connected.*† *Then, there exist a matrix* $Z \in SL_\pm(d, \mathbb{Z})$, *a vector* $z \in \mathbb{R}^d$, *and a map* $\mathcal{F} : a(U \cap U') \to \mathbb{R}^d$ *which satisfy*

$$\frac{\partial}{\partial a_i}(Z^{-T}\mathcal{F})_j \;=\; \frac{\partial}{\partial a_j}(Z^{-T}\mathcal{F})_i\,, \qquad i, j = 1, \ldots, d\,, \tag{2.1.2}$$

and are such that one has, in $U \cap U'$,

$$\begin{aligned} a' \;&=\; Za \,+\, z \\ \alpha' \;&=\; Z^{-T}\alpha \,+\, \mathcal{F}(a) \quad (mod 2\pi)\,. \end{aligned} \tag{2.1.3}$$

Conversely, given a local system of action–angle coordinates $a \times \alpha : U \to \hat{A} \times \mathbb{T}^d$, *a matrix* $Z \in SL_\pm(d, \mathbb{Z})$, *a vector* $z \in \mathbb{R}^d$ *and a map* $\mathcal{F} : a(U) \to \mathbb{R}^d$ *satisfying (2.1.2), the map* $a' \times \alpha'$ *defined by (2.1.3) is a local system of action–angle coordinates of the fibration in* U.‡

Proof. In order to prove the first statement, observe that since both a and a' are invertible functions of F, one has $a' = a'(a)$. Then, the equality $\sum da_i \wedge d\alpha_i = \sum da'_i \wedge d\alpha'_i$ reduces to

$$\sum_i \frac{\partial a'_i}{\partial a_j}(a)\frac{\partial \alpha'_i}{\partial \alpha_l}(a, \alpha) \;=\; \delta_{jl}\,, \qquad j, l = 1, \ldots, d\,, \tag{2.1.4a}$$

$$\sum_i \frac{\partial a'_i}{\partial a_j}(a)\frac{\partial \alpha'_i}{\partial a_l}(a, \alpha) \;=\; \sum_i \frac{\partial a'_i}{\partial a_l}(a)\frac{\partial \alpha'_i}{\partial a_j}(a, \alpha)\,, \qquad j, l = 1, \ldots, d\,. \tag{2.1.4b}$$

The first set of equations gives

$$\frac{\partial \alpha'}{\partial \alpha}(a, \alpha) \;=\; \left[\frac{\partial a'}{\partial a}(a)\right]^{-T}\,, \tag{2.1.5}$$

which implies that $\frac{\partial \alpha'}{\partial \alpha}$ is independent of α. Thus, there exist a matrix $C(a)$ and a vector $\mathcal{F}(a)$ which depend smoothly on a and are such that

$$\alpha' \;=\; C(a)\,\alpha \,+\, \mathcal{F}(a) \qquad (mod 2\pi)\,.$$

Since α and α' are both coordinates on the torus $a = const$, the matrix $C(a)$ belongs to $SL_\pm(d, \mathbb{Z})$. Consequently, it is constant in $U \cap U'$. From (2.1.5) it also follows $a' = Za + const$

† Otherwise, the statement applies to each connected component of $U \cap U'$, separately.

‡ On account of equations (2.1.2), one has locally $\mathcal{F} = Z^{-T}\frac{\partial \mathcal{V}}{\partial a}$ with some function $\mathcal{V}(a)$.

with $Z = C^{-T}$. Hence, $\partial a'_i/\partial a_j = Z_{ij}$, and equations (2.1.4b) imply (2.1.2). The proof of the converse is obvious; let us only note that (2.1.2) is needed to prove that $a' \times \alpha'$ defined by (2.1.3) is symplectic. ■

So, the action–angle coordinates are never unique, but all possible systems of them are related by transformations of the form (2.1.3). We shall come back again on this point after the proof of the Liouville–Arnol'd theorem, in section 2.2.D.

Remark: Equations (2.1.3) show that the manifold A has an affine structure (i.e., an atlas with affine transition functions). This is a general property of the base of a Lagrangian fibration, see [18].

2.2 Proof of the Liouville–Arnol'd theorem

We articulate the proof of the Liouville–Arnol'd theorem in three steps. First, we prove that the level sets of the map $\mathcal{F}$ are tori. Then we prove that the functions F_i and the times of their Hamiltonian vector fields provide local symplectic coordinates; this is a particular case of a theorem by Charatheodory. Finally, we globalize these 'energy–time' coordinates to a neighbourhood of a whole torus, and pass to the action–angle coordinates.

A. Proof of statement i. We prove here that each connected component N_x, $x \in C_F$, of a level set of the map $F = F_1 \times \ldots \times F_d$ is diffeomorphic to the torus $\mathbb{T}^d$. (The way the manifolds $F = const$ are identified with the standard torus $\mathbb{T}^d$ is important for the proof of statement ii) below, and also for the comprehension of the way one constructs the angles in practical problems).

The basic observation is that, at each point of N_x, the Hamiltonian vector fields $X_{F_1}, \ldots, X_{F_d}$ of the functions $F_1, \ldots, F_d$ are tangent to N_x, are linearly independent, and pairwise commute. The existence of vector fields with these properties, as a consequence of the properties of the map F, is the only point in the proof of statement i) in which the symplectic structure of M plays a role. Indeed, one has the following:

Proposition 2.4 *Let N be a compact connected manifold of dimension d. Assume that it has d tangent vector fields $X_1, \ldots, X_d$ which are everywhere linearly independent and pairwise commute. Then, N is diffeomorphic to $\mathbb{T}^d$.*

Proof.† Since N is compact, each vector field X_i is complete, that is, its flow $(\tau_i, x) \mapsto \Phi^{X_i}_{\tau_i}(x)$ is defined for all $\tau_i \in \mathbb{R}$ and $x \in N$. Moreover, since the vector fields commute, their flows commute, too: $\Phi^{X_i}_{\tau_i} \circ \Phi^{X_j}_{\tau_j} = \Phi^{X_j}_{\tau_j} \circ \Phi^{X_i}_{\tau_i}$ for all i, j. Thus, it is possible to define a map

$$\Phi \,:\, \mathbb{R}^d \times N \,\to\, N\,, \qquad (\tau, x) \mapsto \Phi_\tau(x)\,,$$

† We follow closely the proof in Arnold's book [4].

through $\Phi_{(\tau_1,\ldots,\tau_d)}(x) = \Phi^{X_1}_{\tau_1} \circ \Phi^{X_2}_{\tau_2} \circ \cdots \circ \Phi^{X_d}_{\tau_d}(x)$. Since

$$\Phi_0 \;=\; id_N \,,$$
$$\Phi_{\tau'} \circ \Phi_{\tau''} \;=\; \Phi_{\tau'+\tau''} \quad \textit{for all } \tau' \,,\; \tau'' \in \mathbb{R}^d \,,$$

Φ is an action of the commutative group $(\mathbb{R}^d, +)$ on N.

The following lemma collects a few basic properties of this action. We express some of them in terms of the maps

$$\Psi_x \,:\, \mathbb{R}^d \;\to\; N \,, \qquad \Psi_x(\tau) \;=\; \Phi_\tau(x) \,,$$

which are defined for each point $x \in N$. Let us recall that the *orbit* of a point $x \in N$ under the action Φ is the set $\{\Phi_\tau(x) \,:\, \tau \in \mathbb{R}^d\}$, that is, $\Psi_x(\mathbb{R}^d)$. The action is said to be *transitive* if the orbit of each point coincides with N, that is, if Ψ_x is surjective. Moreover, we recall that the *isotropy group* of a point x is $\mathcal{L}(x) = \{\tau \in \mathbb{R}^d \,:\, \Phi_\tau(x) = x\}$, that is, $\Psi_x^{-1}(\{x\})$.

Lemma 2.5 *For each point $x \in N$:*

i) Ψ_x is a local diffeomorphism: that is, for each $\tau \in \mathbb{R}^d$ there exist a neighbourhood $\mathcal{T}_{x,\tau}$ of τ and a neighbourhood $V_{x,\tau}$ of $\Psi_x(\tau) = \Phi_\tau(x)$ such that Ψ_x is a diffeomorphism of $\mathcal{T}_{x,\tau}$ onto $V_{x,\tau}$.

ii) Ψ_x is surjective.

iii) The isotropy group $\mathcal{L}(x)$ of x is a discrete subgroup of $\mathbb{R}^d$, and it is the same for all points $x \in N$.

Proof. One computes

$$\frac{\partial \Psi_x}{\partial \tau_1}\bigg|_\tau \;=\; \frac{\partial \Phi^{X_1}_{\tau_1}}{\partial \tau_1}\big(\Phi^{X_2}_{\tau_2} \circ \cdots \circ \Phi^{X_d}_{\tau_d}(x)\big)\bigg|_\tau \;=\; X_1\big(\Phi^{X_1}_{\tau_1} \circ \Phi^{X_2}_{\tau_2} \circ \cdots \circ \Phi^{X_d}_{\tau_d}(x)\big)\bigg|_\tau \;=\; X_1(\Phi_x(\tau)) \,.$$

Since the flows commute, the same computation gives $\frac{\partial \Psi_x}{\partial \tau_i}\big|_\tau = X_i(\Psi_x(\tau))$ for all $i = 1, \ldots, d$. Thus, the derivative $D\Psi_x|_\tau$ is an isomorphism, being represented by a matrix with linearly independent rows $X_1(\Psi_x(\tau)), \ldots, X_d(\Psi_x(\tau))$. Statement *i)* follows from the inverse function theorem.

In order to prove the second statement, consider any point $x' \in N$, and consider a curve joining it to x (it exists, since N is a connected manifold). For each point y of the curve, consider the neighbourhood $V_{y,0}$ as in i). Since a curve joining two points is compact, a finite number $V_{y_1,0}, \ldots, V_{y_k,0}$ of these open sets cover it. Take points $x_0 = x, x_1, \ldots, x_k = x'$ such that $x_i \in V_{y_{i-1},0} \cap V_{y_i,0}$. Since Ψ_{y_i} is a diffeomorphism onto $V_{y_i,0}$, there exists $\tau_i \in \mathbb{R}^d$ such that $x_i = \Phi_{\tau_i}(x_{i-1})$ (take $\tau_i = \tau_i'' - \tau_i'$, where $x_{i-1} = \Psi_{y_i}(\tau_i')$ and $x_i = \Psi_{y_i}(\tau_i'')$). Thus, $x' = \Psi_x(\tau)$ with $\tau = \sum_i \tau_i$, proving surjectivity.

Finally, since locally Ψ_x is one–to–one (by *i)*), the points $\tau \in \mathcal{L}(x) = \Psi_x^{-1}(\{x\})$ are isolated, so that the group $\mathcal{L}(x)$ is discrete. The fact that $\mathcal{L}(x) = \mathcal{L}(x')$ for all points x and x' of N follows from the transitivity and the commutativity of the action Φ: let $\tau \in \mathcal{L}(x)$ and let τ' be such that $x' = \Phi_{\tau'}(x)$. Then, $\Phi_\tau(x') = \Phi_{\tau+\tau'}(x) = \Phi_{\tau'}(x) = x'$. ∎

At this point, we make a short digression. It is a known fact that each discrete subgroup of $\mathbb{R}^d$ is a *lattice* of $\mathbb{R}^d$ of some dimension $0 \le k \le d$, namely, it is the set

$$\{\tau \in \mathbb{R}^d \,:\, \tau = \sum_{i=1}^{k} \nu_i u_i \,,\; \nu_i \in \mathbb{Z}\}$$

consisting of all the linear combinations with integer coefficients of k linearly independent vectors $u_1, \dots, u_k$ of $\mathbb{R}^d$ (the zero–dimensional lattice consists of the vector 0 alone). The proof of this fact can be found, for instance, in [4] or in [19].

The vectors $u_1, \dots, u_k$ as above are said to constitute a basis of the lattice. The basis of a lattice (of any dimension $n \geq 1$) is never unique: it is easy to prove that, if $\{u_1, \dots, u_k\}$ is a basis, then k vectors $v_1, \dots, v_k$ constitute another basis of the same lattice if and only if there exists a matrix $Z \in SL_{\pm}(d, \mathbb{Z})$ such that $v_i = Zu_i$ for all $i = 1, \dots, k$ (see [19]).

Every k–dimensional lattice of $\mathbb{R}^d$ is isomorphic to the standard lattice with basis $2\pi e_1, \dots, 2\pi e_k$, where $\{e_1, \dots, e_d\}$ is the canonical basis of $\mathbb{R}^d$. This is seen by completing a basis $\{u_1, \dots, u_k\}$ of the lattice to a basis $\{u_1, \dots, u_k, \dots, u_d\}$ of $\mathbb{R}^d$, by defining a $d \times d$ matrix L through

$$L_{ij} \;=\; u_j \cdot e_i \,, \tag{2.2.1}$$

and by performing the linear transformation $2\pi L^{-1} : \mathbb{R}^d \to \mathbb{R}^d$, which maps the vector u_i into the vector $2\pi e_i$, $i = 1, \dots, d$. This implies, in particular, that if $\mathcal{L}$ is a k–dimensional lattice, then the quotient space $\mathbb{R}^d/\mathcal{L}$ is diffeomorphic to the cylinder $\mathbb{R}^{d-k} \times \mathbb{T}^k$, where $\mathbb{T}^k$ is the standard k–dimensional torus $\mathbb{R}/(2\pi\mathbb{Z})^k$. Note that this identification depends on the choice of a basis of $\mathcal{L}$.

We can now conclude the proof of Proposition 2.4. Let $\mathcal{L}$ be the isotropy group of the points of N. Since $\Phi_\tau = id_N$ iff $\tau \in \mathcal{L}$, $\mathcal{L}$ and its elements will be called, respectively, the *period lattice* and the *periods* of the action Φ; similarly, the matrix L as in (2.2.1) will be called the *period matrix*.

Let us now fix a point $x \in N$. The surjective map $\Psi_x : \mathbb{R}^d \to N$ induces a bijective map $\widehat{\Psi}_x : \mathbb{R}^d/\mathcal{L} \to N$, which is indeed a diffeomorphism (since Ψ_x is a local diffeomorphism at every point). Thus, choosing a basis for $\mathcal{L}$, we find that N is diffeomorphic to a cylinder $\mathbb{R}^{d-k} \times \mathbb{T}^k$. Since N is compact, this implies $k = d$, and N is diffeomorphic to the standard torus $\mathbb{T}^d$. Note that the diffeomorphism is obtained by composing the map $\widehat{\Psi}_x$ with the isomorphism $\frac{1}{2\pi}\widehat{L} : \mathbb{T}^d \to (\mathbb{R}/\mathcal{L})^d$ induced by the isomorphism $\frac{1}{2\pi}L : \mathbb{R}^d \to \mathbb{R}^d$, as shown in figure 2.2. This concludes the proof of proposition 2.4. ∎

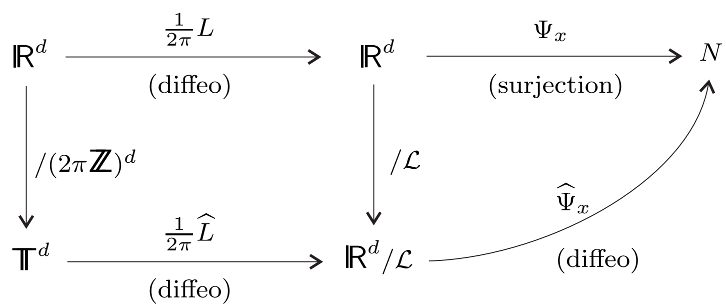


Figure 2.2

We add now a few remarks about the identification of N with the standard torus, which is provided by the angular coordinates

$$\alpha \;=\; 2\pi L^{-1}\tau\,. \tag{2.2.2}$$

This identification depends on the choice of a point x, which serves as origin of the angle, and on the choice of a basis $u_1,\dots,u_d$ of the period lattice $\mathcal{L}$, or equivalently of a period matrix L with entries $L_{ij} = (u_j)_i$. Specifically, the i–th coordinate circle $\{\alpha_i \in S^1,\, \alpha_j = 0\ for\ j \neq i\}$ on $\mathbb{T}^d$ coordinatizes the closed curve

$$\gamma_i : [0,2\pi] \to N\,, \qquad \lambda \mapsto \Phi_{\lambda u_i}(x) \tag{2.2.3}$$

on N.

Conversely, given any set of d independent cycles $\tilde\gamma_1,\dots,\tilde\gamma_d$ on N, based at the point x, there is a basis $\tilde u_1,\dots,\tilde u_d$ of $\mathcal{L}$ such that, for any $i=1,\dots,d$, the curve $\lambda \mapsto \Phi_{\lambda\tilde u_i}(x)$ is homologous to the cycle $\tilde\gamma_i$. (Indeed, there exists a matrix $Z \in SL_\pm(d,\mathbb{Z})$ such that $\tilde\gamma_i = Z\gamma_i$; hence, take $\tilde u_i = Zu_i$). The entries of the period matrix corresponding to such a choice are given by

$$\begin{aligned} L_{ji} \;&= \textit{the time that the flow of } X_{F_j} \textit{ acts while the action } \Phi \\ &\quad \textit{makes the point } x \textit{ to run once along a cycle on } N \textit{ homologous to } \gamma_i\,. \end{aligned} \tag{2.2.4}$$

So, we see that the choice of a basis of the period lattice amounts to the choice of a basis for the homology of N (that is, d independent cycles which serve as "coordinates circles" for the identification of N with the standard torus).

The remaining of the proof of the Liouville–Arnol'd theorem consists in showing that the angular coordinates constructed here on each level set $F = const$, if suitably chosen, can be completed to a system of canonical coordinates.

Exercise: Show that the orbits of the vector fields

$$Y_j \;=\; \frac{1}{2\pi}\sum_{i=1}^{d} L_{ij}X_{F_i}\,, \qquad j=1,\dots,d\,,$$

are periodic of period 2π and coincide with the cycles γ_j as in (2.2.3). Hence, the time along the flow of Y_j is the angle α_i on the standard torus. (Hints: Since the flows of $F_1,\dots,F_d$ commute, for each $\tau = (\tau_1,\dots,\tau_d) \in \mathbb{R}^d$ the map Φ_τ coincides with the time–one–map of the flow of the vector field $Y_\tau = \sum_i \tau_i X_{F_i}$).

B. The 'energy–time' coordinates. As a first step in the proof of statement ii. we show that, locally, the d functions $F_1,\dots,F_d$ and the 'times' $T_1,\dots,T_d$ along the flows of their Hamiltonian vector fields, measured from a suitably chosen origin, are Darboux coordinates. This fact is a particular case of a 'completion' theorem by Charatheodory (see [7]).

Proposition 2.6 *Let (M,Ω) be a symplectic manifold of dimension $2d$. Assume that there exist d functions $F_1,\dots,F_d$ which are pairwise in involution and have differentials linearly independent at a point $x_* \in M$. Then, there exists a neighbourhood P of x_* and functions $(T_1,\dots,T_d) : P \to \mathbb{R}^d$ such that $(F_1,\dots,F_d,T_1,\dots,T_d)$ are Darboux coordinates (i.e., $\Omega|_P = \sum_i dF_i \wedge dT_i$).*

Proof. We begin by choosing a d–dimensional submanifold of M which contains x_* and is transversal to the level set of F through x_*. Such a submanifold certainly exists and, in a sufficiently small neighbourhood P_1 of x_*, to which we restrict ourselves, it intersects in exactly one point every other level set of F and can be given by an immersion

$$\sigma \,:\, \mathcal{F}_1 \to P_1\,,$$

where $\mathcal{F}_1 = F(P_1)$ is a neighbourhood of $f_* = F(x_*)$. (This can be seen as follows. Since F is a submersion at x_*, it is a submersion at all points of a neighbourhood of x_*. Furthermore, there exists a neighbourhood P of x_* where is defined another submersion $Y : P \to \mathbb{R}^d$ such that (F, Y) are local coordinates. Let us then choose P_1 as the preimage under the map $F \times Y$ of a coordinate domain $\mathcal{F}_1 \times \mathcal{Y}_1$ which contains $(f_*, y_*) = (F(x_*), Y(x_*))$, and take $\sigma(f) = (F \times Y)^{-1}(f, y_*)$.) In the sequel, we call 'section' the submanifold $\sigma(\mathcal{F}_1)$.

We now coordinatize a (sufficiently small) subset of P_1 with the functions F_i and the corresponding times T_i, measured from the section $\sigma(\mathcal{F}_1)$. To this purpose, let us observe that there exists neighbourhoods $\mathcal{F}$ of f_*, $\mathcal{T}$ of $0 \in \mathbb{R}^d$ and P of x_* such that the map $\mathcal{C}$ defined by

$$\mathcal{C}(f, \tau) \;=\; \Phi_\tau(\sigma(f))\,, \tag{2.2.5}$$

where $\Phi_\tau = \Phi_{\tau_1}^{X_1} \circ \cdots \circ \Phi_{\tau_d}^{X_d}$, is a diffeomorphism from $\mathcal{F} \times \mathcal{T}$ onto P (see figure 2.3). This follows from the inverse function theorem: the derivative of $\mathcal{C}$ at $(f_*, 0)$ is given by the matrix with rows $X_{F_1}(x_*), \ldots, X_{F_d}(x_*)$, $\frac{\partial\sigma}{\partial f_1}(f_*), \ldots, \frac{\partial\sigma}{\partial f_d}(f_*)$ and so, since σ is transversal to $F^{-1}(f_*)$ in x_*, $D\mathcal{C}(f_*, 0)$ is nonsingular.

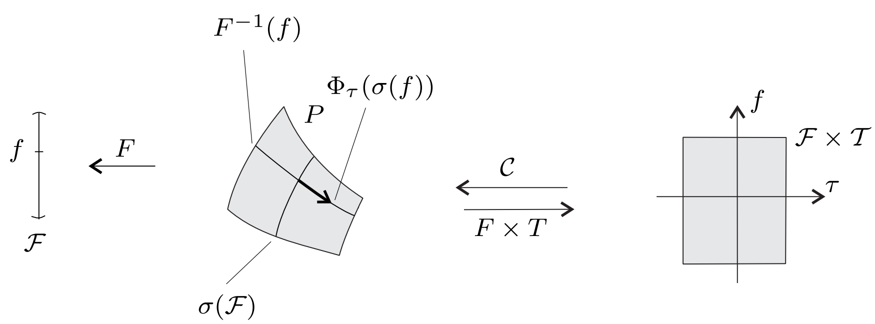


Figure 2.3

The inverse of the map $\mathcal{C}$, which we denote $F \times T : P \to \mathcal{F} \times \mathcal{T}$ provides local coordinates on P. The question now is whether these coordinates can be taken to be Darboux coordinates, that is, whether $\mathcal{C} : \mathcal{F} \times \mathcal{T} \to P$ is symplectic with respect to the standard two–form $\sum_i df_i \times d\tau_i$ on $\mathcal{F} \times \mathcal{T}$. As the next Lemma shows, this depends solely on the choice of the section σ:

Lemma 2.7 *The map $\mathcal{C} : \mathcal{F} \times \mathcal{T} \to P$ defined as above is symplectic (with respect to Ω and $\sum_i df_i \wedge d\tau_i$) if and only if the section σ is Lagrangian.*

Proof. If $\mathcal{C}$ is symplectic, then σ is Lagrangian, since it is the image under $\mathcal{C}$ of $\tau = 0$, which is Lagrangian for the considered symplectic form on $\mathcal{F} \times \mathcal{T}$. We thus prove the converse. We can write

$$\Omega|_P \;=\; \sum_{ij}\big(U_{ij}\,dF_i\wedge dT_j \;+\; \frac{1}{2}V_{ij}\,dF_i\wedge dF_j \;+\; \frac{1}{2}W_{ij}\,dT_i\wedge dT_j\big)\,,$$

with certain functions U_{ij}, $V_{ij}=-V_{ij}$, and $W_{ij}=-W_{ij}$. Since T_i is the time along the flow of X_{F_i}, one has $X_{F_i}=\frac{\partial}{\partial T_i}$ $(i=1,\dots,d)$. Thus,

$$W_{ij} \;=\; \Omega\Big(\frac{\partial}{\partial T_i},\frac{\partial}{\partial T_j}\Big) \;=\; \Omega(X_{F_i},X_{F_j}) \;=\; \{F_i,F_j\} \;=\; 0\,.$$

Moreover, using the relation between a Hamiltonian vector field and the differential of its Hamilton function, one computes

$$U_{ij} \;=\; \Omega\Big(\frac{\partial}{\partial F_i},\frac{\partial}{\partial T_j}\Big) \;=\; \Omega\Big(\frac{\partial}{\partial F_i},X_{F_j}\Big) \;=\; \langle dF_j,\frac{\partial}{\partial F_i}\rangle \;=\; \delta_{ij}\,.$$

Hence

$$\Omega|_P \;=\; \sum_i dF_i\wedge dT_i \;+\; \frac{1}{2}\sum_{ij}V_{ij}dF_i\wedge dF_j\,. \tag{2.2.6}$$

We now show that the functions V_{ij} are invariant under the map Φ_τ, i.e.,

$$V_{ij}\circ\Phi_\tau \;=\; V_{ij} \tag{2.2.7}$$

for all τ (small enough that one remains in the coordinate domain). To this end, observe that one has $\Phi_\tau^* dF_i = dF_i$ and $\Phi_\tau^* dT_i = dT_i$ for each fixed τ, as it follows from $F_i\circ\Phi_\tau=F_i$ and $T_i\circ\Phi_\tau=T_i+\tau_i$. Using also $\Phi_\tau^*\Omega=\Omega$, one then deduces from (2.2.6)

$$0 \;=\; \Phi_\tau^*\Omega-\Omega \;=\; \frac{1}{2}\sum_{ij}\big(V_{ij}\circ\Phi_\tau - V_{ij}\big)\,dF_i\wedge dF_j\,,$$

which proves (2.2.7).

We can now conclude the proof of the lemma. If $\sigma:\mathcal{F}\to P$ is Lagrangian, then Ω vanishes on it and so, on account of (2.2.6) and of the fact that $\frac{\partial}{\partial F_1},\dots,\frac{\partial}{\partial F_d}$ are a basis for the tangent spaces to $\sigma(\mathcal{F})$, one has $V_{ij}(\sigma)=0$ for all i and j. But then, by (2.2.7), the functions V_{ij} are everywhere zero in P, so that $\Omega=\sum_i dF_i\wedge dT_i$ and $\mathcal{C}$ is symplectic. ∎

On account of Lemma 2.7, the proof of Proposition 2.6 is concluded by the following

Lemma 2.8 *There exists a Lagrangian section $\sigma:\mathcal{F}\to M$, where $\mathcal{F}$ is a suitable neighbourhood of f_*.*

Proof. We construct a Lagrangian section by suitably modifying, and possibly restricting, the section $\sigma:\mathcal{F}\to P$ introduced at the beginning of the proof of Proposition 2.6. Let us here identify the map F with its restriction to the set P, so that we have a surjective submersion $F:P\to\mathcal{F}$.

As seen in the proof of the previous Lemma, in P, the symplectic form Ω has the expression (2.2.6), where the functions V_{ij} are constant on the fibers of F. The latter fact implies that there exist functions $\widehat{V}_{ij}:\mathcal{F}\to\mathbb{R}$ such that $V_{ij}=\widehat{V}_{ij}\circ F$. Therefore, the two–form $\beta = \frac{1}{2}\sum_{ij} V_{ij} dF_i \wedge dF_j$ on $\mathcal{F} \times \mathcal{T}$ is the pull–back $F^*\widehat{\beta}$ of the two–form $\widehat{\beta} = \frac{1}{2}\sum_{ij} \widehat{V}_{ij}(f) df_i \wedge df_j$ on $\mathcal{F}$.

Since $\beta = \Omega - \sum dF_i \wedge dT_i$ is closed, $\widehat{\beta}$ is closed, too. Hence, possibly restricting the neighbourhood $\mathcal{F}$ of f_*, there exists a one–form $\widehat{\eta} = \sum_i \widehat{\eta}_i(f) df_i$ on $\mathcal{F}$ such that $\widehat{\beta} = d\widehat{\eta}$. This implies $\beta|_{F^{-1}(\mathcal{F})} = d\eta$, with $\eta = F^*\widehat{\eta} = \sum_i (\widehat{\eta}_i \circ F) dF_i$. So

$$\Omega|_{F^{-1}(\mathcal{F})} = \sum_i dF_i \wedge d(T_i - \widehat{\eta}_i \circ F)\,,$$

which shows that the section defined by $T_i(x) = \widehat{\eta}_i(F(x))$, $i = 1, \ldots, d$, is Lagrangian. ■■

C. Proof of statement ii. We conclude now the proof of Theorem 2.1. We choose a local Lagrangian section σ of the fibration defined by F, through the considered point x, and we consider the corresponding "energy–time" Darboux coordinates (F, T). As shown in the previous subsection, these coordinates are defined in a certain neighbourhood $P = \mathcal{C}(\mathcal{F} \times \mathcal{T})$ of x, where $\mathcal{F} \subset F(M)$ is an open set in $\mathbb{R}^d$ and $\mathcal{T}$ is a neighbourhood of $0 \in \mathbb{R}^d$.

First of all, we extend these coordinates to the whole neighbourhood $F^{-1}(\mathcal{F})$ of the torus N_x. (If the level sets of F are not connected, we tacitly identify F with its restriction to the connected component of $F^{-1}(\mathcal{F})$ which contains the point x). To this end, observe that in fact equation (2.2.5) defines a map $\mathcal{C} : \mathcal{F} \times \mathbb{R}^d \to F^{-1}(\mathcal{F})$, which is obviously surjective and symplectic. (one uses here the compactness of the torus, which assures that the vector fields X_{F_i} are complete). Moreover, as we now show, this map is a local diffeomorphism at every point of $\mathcal{F} \times \mathbb{R}^d$. Specifically, we show that, for every $\overline{\tau} \in \mathbb{R}^d$, $\mathcal{C}$ is a diffeomorphism of $\mathcal{F} \times (\overline{\tau} + \mathcal{T})$ onto $\Phi_{\overline{\tau}}(\mathcal{C}(\mathcal{F} \times \mathcal{T}))$, where $\overline{\tau} + \mathcal{T} = \{\tau + \overline{\tau} \,:\, \tau \in \mathcal{T}\}$. In fact, if $f \in \mathcal{F}$ and $\tau \in \mathcal{T} + \overline{\tau}$, then $\mathcal{C}(f, \tau) = \Phi_{\overline{\tau}} \circ \Phi_{\tau - \overline{\tau}}(\sigma(f))$ is obtained by composing the translation $(f, \tau) \mapsto (f, \tau - \overline{\tau})$, the restriction of $\mathcal{C}$ to $\mathcal{F} \times \mathcal{T}$, and the 'time evolution' $\Phi_{\overline{\tau}}$, which are all diffeomorphisms; see figure 2.4.

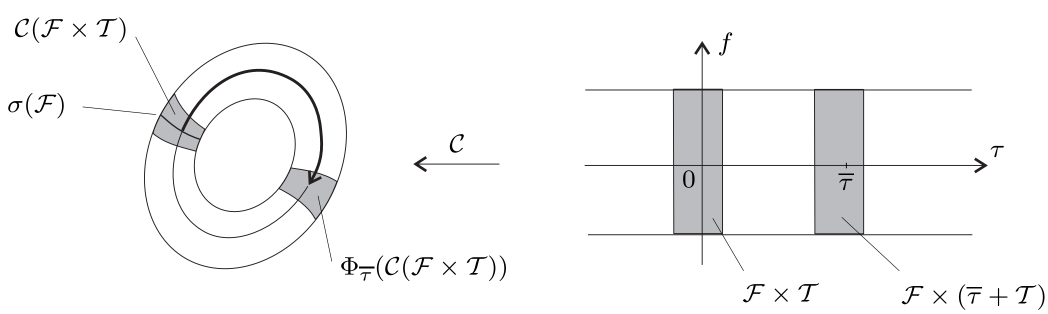


Figure 2.4

The map $\mathcal{C} : \mathcal{F} \times \mathbb{R}^d \to F^{-1}(\mathcal{F})$ is of course not injective. For every $f \in \mathcal{F}$, one can define the period lattice $\mathcal{L}(f)$ of the submanifold $F^{-1}(f)$ as in the subsection 2.A: a point $\tau \in \mathbb{R}^d$ belongs to $\mathcal{L}(f)$ if and only if

$$\Phi_\tau(\sigma(f)) = \sigma(f)\,.$$

By applying the implicit function theorem to this equation, one sees that the period lattice is a smooth function of f, in some neighbourhood $\mathcal{F}$ of f_*. Similarly, the period matrix $L(f)$ depends smoothly on f. (Roughly speaking, one chooses a basis for the period lattice of a given torus, and then smoothly transports this basis to the neighbour tori).

We now normalize the periods to 2π (see section 2.2.A):

Lemma 2.9 *There exists a diffeomorphism $\widehat{a} : \mathcal{F} \to \hat{A}$ from some neighbourhood $\mathcal{F}$ of f_* onto an open set $\hat{A} \subset \mathbb{R}^d$ such that the map*

$$\widehat{a} \times \widehat{\alpha} \,:\, \mathcal{F} \times \mathbb{R}^d \to \hat{A} \times \mathbb{R}^d\,, \tag{2.2.8}$$

where $\widehat{\alpha} : \mathcal{F} \times \mathbb{R}^d \to \mathbb{R}^d$ is given by

$$\widehat{\alpha}(f,\tau) \;=\; 2\pi\, L(f)^{-1}\tau\,, \tag{2.2.9}$$

is a symplectic diffeomorphism (with respect to the two–forms $\sum df_i \wedge d\tau_i$ and $\sum da_i \wedge d\alpha_i$).

Proof. Let $\widehat{\alpha}$ be defined as in (2.2.9). We construct a map $\widehat{a} = \widehat{a}(f)$ in such a way that $\widehat{a} \times \widehat{\alpha}$ is a symplectic diffeomorphism. Since

$$d\widehat{a}_i \wedge d\widehat{\alpha}_i \;=\; \sum_j \frac{\partial \widehat{a}_i}{\partial f_j} df_j \wedge d\widehat{\alpha}_i\,,$$

$$df_i \wedge d\tau_i \;=\; \frac{1}{2\pi}\, df_i \wedge \left(\sum_{jl} \frac{\partial L_{ij}}{\partial f_l} \widehat{\alpha}_j df_l \;+\; \sum_j L_{ij} d\widehat{\alpha}_j \right)\,,$$

the condition $\sum d\widehat{a}_i \wedge d\widehat{\alpha}_i = \sum df_i \wedge d\tau_i$ is satisfied if

$$\frac{\partial L_{ij}}{\partial f_l}(f) \;=\; \frac{\partial L_{lj}}{\partial f_i}(f)\,, \qquad\qquad i,j,l \;=\; 1,\ldots,d\,, \tag{2.2.10a}$$

$$\frac{\partial \widehat{a}_i}{\partial f_j}(f) \;=\; \frac{1}{2\pi}\, L_{ji}(f)\,, \qquad\qquad i,j \;=\; 1,\ldots,d\,, \tag{2.2.10b}$$

Equations (2.2.10a) are the closeness conditions which assure that equations (2.2.10b) can be integrated to get a diffeomorphism $f \mapsto \widehat{a}(f)$ from a neighbourhood $\mathcal{F}$ of f_* onto an open set $\hat{A}$;† the fact that such a map is a diffeomorphism, if $\mathcal{F}$ is small enough, is assured by $det(\partial\widehat{a}/\partial f) = det\, L/2\pi \neq 0$.

Thus, it only remains to prove that the period matrix satisfies (2.2.10a). As we now show, this is a consequence of the Lagrangian character of the section $f \mapsto \sigma(f)$. Since $\mathcal{C}$ is symplectic, this implies that $f \mapsto \mathcal{C}^{-1}(\sigma(f))$ is a Lagrangian submanifold of $\mathcal{F} \times \mathbb{R}^d$. But

$$\mathcal{C}^{-1}(\sigma(f)) \;=\; \bigcup_{\tau \in \mathcal{L}(f)} (f\,,\,\tau) \;=\; \bigcup_{\nu \in \mathbb{Z}^d} (f\,,\,L(f)\nu) \tag{2.2.11}$$

since $\tau \in \mathcal{L}(f)$ iff there exists $\nu \in \mathbb{Z}^d$ such that $\tau = L(f)\nu$. Hence, for any $\nu \in \mathbb{Z}^d$, $f \mapsto (f, L(f)\nu)$ is Lagrangian. A basis of vector fields tangent to the submanifold $\tau = L(f)\nu$ is constituted by

$$Z_l \;=\; \frac{\partial}{\partial f_l} \;+\; \sum_{ij} \frac{\partial L_{ij}\nu_j}{\partial f_l} \frac{\partial}{\partial \tau_i}\,, \qquad l = 1,\ldots,d\,.$$

† Observe that (2.2.10b) is a generalization of the period–energy relation (1.2.3).

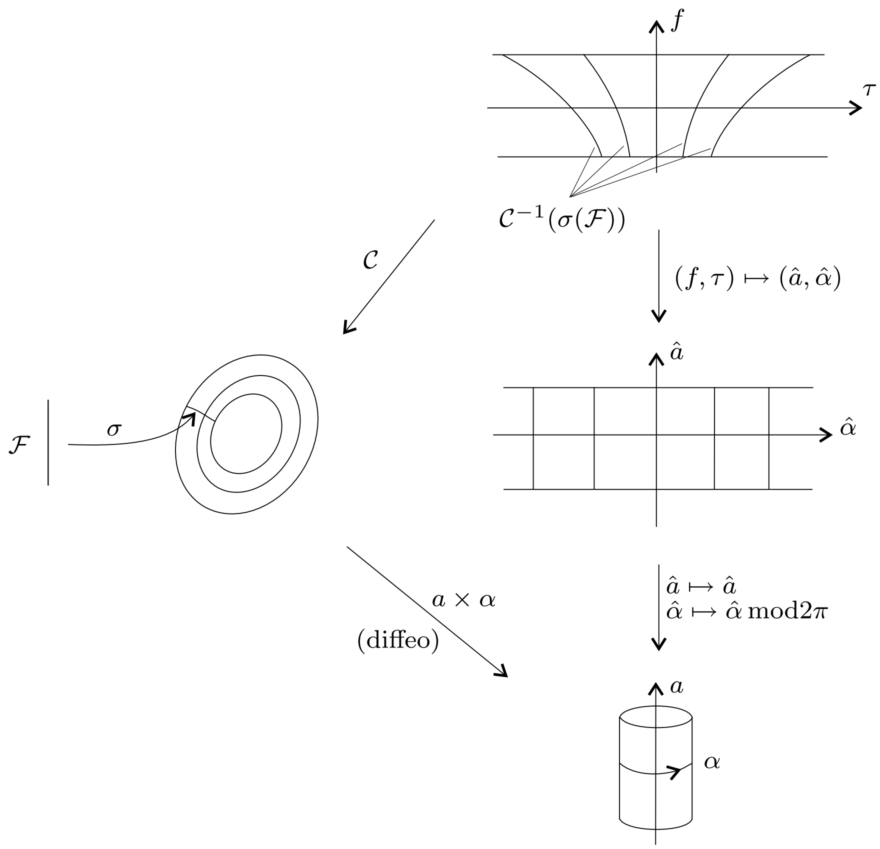


Figure 2.5

As one readily verifies, the conditions $\Omega(Z_i, Z_j) = 0$ give exactly (2.2.10)a. ■

At this point, the proof of the Liouville–Arnol'd theorem is completed just by considering the quotient manifold $\hat{A} \times (\mathbb{R}^d/(2\pi\mathbb{Z})^d)$. The 'action–angle' coordinates (a, α) are defined by $a = \widehat{a} \circ F$, $\alpha = \widehat{\alpha} \circ (F \times T)$ $(mod 2\pi)$. The whole construction is shown in figure 2.5.

Exercise: Show that each vector field $Y_j = \frac{1}{2\pi} \sum_j L_{ij}(F) X_{F_i}$ is Hamiltonian, with Hamilton function $\widehat{a}_j \circ F$. (Hint: use (2.2.10a)).

D. On the non–uniqueness of the action–angle coordinates. We relate here the non–uniqueness of the action–angle coordinates to the proof of the Liouville–Arnol'd theorem. Within the proof, the construction of the action–angle coordinates is affected by the following arbitrary choices:

i) the choice of a Lagrangian section as origin of the angles;
ii) the choice of the basis of the period lattice;
iii) an additive constant in the actions, coming from the integration of (2.2.10b).

Points i) and ii) reflect the choice of the identification of N_x with the standard torus. As we now discuss, these three facts are directly related, respectively, to the appearance of $\mathcal{F}$, Z, and z in (2.1.3).

In fact, it is obvious that a change of the section $\tau = 0$ produces a shift $\alpha \mapsto \alpha + \mathcal{F}(a)$ of the angles, with some map $\mathcal{F}$. It is not difficult to check that the shifted section is Lagrangian if and only if $\mathcal{F}$ satisfies

(2.1.2).

Concerning point ii), let us consider the period lattice $\mathcal{L}(f)$ of a given torus. By choosing a basis $\{u_1, \ldots, u_d\}$ of $\mathcal{L}(f)$, one constructs the period matrix $L(f)$ by (2.2.1) and the angles $\alpha = 2\pi L(f)^{-1}\tau$. By considering another basis $\{u'_1, \ldots, u'_d\}$ of $L(f)$, with a corresponding period matrix $L'(f)$, one arrives at the new angles $\alpha' = 2\pi L'(f)^{-1}\tau$ (we are assuming that the choice of the origin of the angles has not been changed). Hence $\alpha' = L'(f)^{-1}L(f)\alpha$ and by comparison with (2.1.3) one concludes

$$Z^{-T} \;=\; L'(f)^{-1}L(f)\,.$$

Chapter 3

Completely Integrable Systems

This chapter is devoted to completely integrable Hamiltonian systems. In section 3.1 we present some elementary properties of these systems. Next, we discuss the connection between complete integrability and the Hamilton–Jacobi theory. The remaining of the chapter deals with the determination of the action–angle coordinates and with examples.

3.1 Completely integrable systems

A. Definition and examples. We say that a Hamiltonian system on a symplectic manifold M_* of dimension $2d$ is *completely integrable* in a subset $M \subset M_*$ if it possesses d integrals of motion in involution which have differentials linearly independent at every point of M. Restricting M if necessary, we furthermore (tacitly) assume that, in M, all the level sets of the integrals of motion are compact.

Under these hypotheses, by the Liouville–Arnol'd theorem, the set M is fibered by d–dimensional tori, and action–angle coordinates $a \times \alpha : U \to \hat{A} \times \mathbb{T}^d$ can be introduced in a neighbourhood U of each of these tori. As is obvious, these tori are invariant under the flow of H (being level sets of integrals of motions), and the actions a are integrals of motion. Hence, as one sees by writing down Hamilton's equations, the local representative h of the Hamilton function H in the action–angle coordinates is independent of the angles. Written in action–angle coordinates, the equations of motion are

$$\dot{a}_i \;=\; 0\,, \qquad \dot{\alpha}_i \;=\; \omega_i(a)\,, \qquad i = 1, \ldots, m\,, \tag{3.1.1}$$

where the *frequency map* $\omega : \hat{A} \to \mathbb{R}^d$ is defined by

$$\omega_i \;=\; \frac{\partial h}{\partial a_i}\,, \qquad i = 1, \ldots, m\,.$$

So, motions are given by

$$a(t) \;=\; a(0)\,, \qquad \alpha(t) \;=\; \alpha(0) + \omega\big(a(0)\big)\,(mod 2\pi)\,,$$

and are linear on the tori $a = const$; these motions are called *quasi–periodic*.

We give now some examples of systems which are completely integrable (in some subset of their phase space):

i) All autonomous Hamiltonian systems with one degree of freedom (H is an integral of motion).

ii) Every autonomous Hamiltonian system with two degrees of freedom having an integral of motion independent of the Hamilton function. Examples are the spherical pendulum and the point in a plane under the influence of a central force field.

iii) The Lagrange top: H, m_3, and m_z are integrals of motion. In this case, $\{m_3, m_z\} = 0$ follows from the fact that $m_3 = p_\psi$ and $m_z = p_\varphi$ are the conjugate momenta of the coordinates φ and ψ.

iv) The point in a (spatial, now) central force field and the Euler–Poinsot systems. In both cases the Hamilton function and the angular momentum vector m are integrals of motion. One can form a set of three scalar integrals of motion in involution by taking H, the modulus $\|m\|$ of the angular momentum, and the component of m along any axis fixed in space.

v) Other examples are: $d \geq 2$ uncoupled harmonic oscillators, the Kowalevskaja top, the geodesic motion on an ellipsoid, the free particle, ...

Even though almost all the classical systems of mechanics are completely integrable, Hamiltonian systems of this type are in a sense exceptional (see for instance [15]). Nevertheless, their study is worthwhile for different reasons. One is that various interesting mechanical systems are small perturbations of completely integrable ones.

Remarks: (i) Often, one of the d integrals of motion is taken to be the Hamilton function, but this is not necessary. However, the Hamiltonian is never independent of any set of d independent integrals of motion in involution. Indeed, on a symplectic manifold of dimension $2d$ there are at most d functions in involution whose differentials are linearly independent at a point. This is proven by contradiction: If $F_1, \ldots, F_{d+1}$ are independent and in involution, then, by Frobenius theorem, the Hamiltonian vector fields X_{F_i} form a completely integrable distribution with $(d+1)$–dimensional leaves; but the equalities $L_{X_{F_i}} F_j = \{F_i, F_j\} = 0$ show that these leaves are contained in the intersection of the level sets $F_i =$ const, which have dimension $d-1$.

(ii) Here, we consider only the subset M of the phase space of a completely integrable system which is fibered by the invariant Lagrangian tori. In the complement of this subset, the system can have quite different properties. In particular, there are motions which are qualitatively different from the quasi–periodic motions on the Lagrangian tori, such as scattering states and separatrices. However, we are also excluding from the consideration all equilibria and certain families of quasi–periodic motions, which take place on low–dimensional tori which are singular leaves of the foliation by the invariant tori, like the periodic orbit shown in figure 1.6. In the neighbourhood of these lower–dimensional tori one cannot of course introduce action–angle coordinates. However, if the foliation by the invariant Lagrangian tori has an elliptic structure transversal to these 'limit' tori, then one can replace the transversal pairs of action–angle coordinates with cartesian coordinates. This possibility is made precise by a generalization of the Liouville–Arnol'd theorem, for which we refer to [20–24] and references therein.

Exercises: (i) Show that, if $F_1 = H, F_2, \ldots, F_d$ are independent in an open invariant set M and are in involution, and if the sets $\{x \in M : H(x) \leq h\}$ are compact, then the level sets of the map F in M are compact.

(ii) Complete integrability could also be defined as the existence of a fibration $\pi : M \to A$ of the phase space with compact, connected, Lagrangian fibers invariant under the flow of the system. (This is slightly more general than referring to a set of integrals of motion in involution, since the fibration π need not be described by a single submersion $F : M \to \mathbb{R}^d$; this happens, for instance, if the base manifold A is not diffeomorphic to an open set in $\mathbb{R}^d$). Show that this definition is equivalent to requiring that every point $x \in M$ has a neighbourhood U in which is defined a symplectic diffeomorphism $a \times \alpha : U \to \hat{A} \times \mathbb{T}^d$ which conjugates the flow to the flow (3.1.1).

B. Resonances. We now review some basic properties of the quasi–periodic motions. We thus consider the differential equation $\dot{\alpha} = \omega$ on $\mathbb{T}^d$, where $\omega \in \mathbb{R}^d$ is a fixed vector. We say that ω is *nonresonant* if

$$\omega \cdot \nu \neq 0 \qquad \textit{for all} \qquad \nu \in \mathbb{Z}^m \setminus \{0\}\,.$$

The following is a classical result, the proof of which can be found, for instance, in [25].

Theorem 3.1 (Dirichlet) *For any $\alpha_0 \in \mathbb{T}^d$, the curve $t \mapsto \alpha_0 + \omega t$ (mod2π) is dense on $\mathbb{T}^d$ if and only if ω is nonresonant.*

The vector $\omega \in \mathbb{R}^d$ is said to be *resonant* if $\omega \cdot \nu = 0$ for some non–zero vector $\nu \in \mathbb{Z}^d$. The set of all vectors $\nu \in \mathbb{Z}^d$ such that $\omega \cdot \nu = 0$ is a lattice of $\mathbb{R}^d$ of some dimension k, $0 \leq k \leq d$, called the *resonant lattice* of ω. The limiting cases $k = d$ and $k = 0$ correspond, respectively, to $\omega = 0$ and to nonresonant ω's. The sets of nonresonant vectors and that of resonant vectors are both dense in $\mathbb{R}^d$, but the latter has zero Lebesgue measure; for any $0 < k < d$, the set of all vectors which have a resonant lattice of dimension k is also dense in $\mathbb{R}^d$.

Proposition 3.2 *If the resonant lattice of a vector $\omega \in \mathbb{R}^d$ has dimension $k \leq d$, then the closure of the orbits of $\dot{\alpha} = \omega$, $\alpha \in \mathbb{T}^d$, is diffeomorphic to a torus of dimension $d - k$.*

Proof. We use the following:

Lemma 3.3 *If the resonant lattice of $\omega \in \mathbb{R}^d$ has dimension k, $0 \leq k < d$, then there exists a matrix $L \in SL_{\pm}(d, \mathbb{Z})$ such that $L\omega = (0, \omega'')$, where $0 \in \mathbb{R}^k$ and $\omega'' \in \mathbb{R}^{d-k}$ satisfies the nonresonance condition $\omega'' \cdot \nu'' \neq 0$ for all $\nu'' \in \mathbb{Z}^{d-k} \setminus \{0\}$.*

Proof of the lemma. Consider any basis $\{u_1, \ldots, u_k\}$ of the resonant lattice of ω. It is a known fact that, if $k < d$, one can complete such a basis to a basis $\{u_1, \ldots, u_d\}$ of $\mathbb{Z}^d$ (see [19]). The matrix L defined by $L_{ij} = (u_i)_j$ belongs to $SL_{\pm}(d, \mathbb{Z})$ and is such that $L^T e_i = u_i$, so $L\omega$ has the stated properties. ■

Proposition 3.2 is now proved by performing the linear transformation

$$L \;:\; \mathbb{T}^d \;\to\; \mathbb{T}^k \;\times\; \mathbb{T}^{d-k}\,, \qquad \alpha \;\mapsto\; L\alpha \;=\; (\alpha', \alpha'')\; mod 2\pi\,,$$

which conjugates the equation $\dot{\alpha} = \omega$ to $\dot{\alpha}' = 0$, $\dot{\alpha}'' = \omega''$, and invoking Dirichlet's theorem. ■

Exercise: Consider an invariant torus of a completely integrable system. Show that the dimension of the resonant lattice of its frequency vector does not depend on the choice of the system of action–angle coordinates.

C. Anisochrony and non–degeneracy. Consider a completely integrable Hamiltonian system, in the domain of a system of action–angle coordinates (a, α). If the Hamilton function h is linear in the action, $h = \omega \cdot a$ with constant $\omega \in \mathbb{R}^d$, then all motions have equal frequencies and the system is called *isochronous*. Harmonic oscillators are examples of isochronous systems (and conversely, every analytic* isochronous system written in action–angle coordinates looks like a system of harmonic oscillators). Systems for which the frequencies of motions change from torus to torus are instead called *anisochronous*. Since the resonance properties of motions change with the frequencies, anisochronous systems can have both resonant and nonresonant motions.

This fact can be made precise if the Hamilton function, written in any system of action–angle coordinates (a, α), satisfies the condition

$$rank\, \frac{\partial^2 h}{\partial a \partial a}(a_0) \;=\; d \qquad (3.1.2)$$

at a certain point a_0. (Note that this condition is independent of the choice of a specific system of action–angle coordinates.) Equation (3.1.2) means that the frequency map $a \mapsto \omega(a) = \partial h / \partial a(a) \in \mathbb{R}^d$ is a local diffeomorphism near a_0. Hence, certain properties of the frequency vectors are, so to say, inherited by the invariant tori in a neighbourhood of the torus $a = a_0$; specifically, both nonresonant and resonant tori are dense in a neighbourhood of the torus $a = a_0$.

A completely integrable system which satisfies condition (3.1.2) at any point is called *nondegenerate*, or else *strictly anisochronous*.

Remark: The fibration by invariant Lagrangian tori of a completely integrable system need not be unique. As an example, consider the system described by the Hamiltonian $H = a$ on the manifold $\mathbb{R} \times S^1 \times S^1 \times S^1 \ni (a, \alpha, p, q)$ with symplectic two–form $\Omega = da \wedge d\alpha + dp \wedge dq$. Both the equations

$$a \;=\; \text{const}\,, \quad p \;=\; \text{const}$$

and

$$a \;=\; \text{const}\,, \quad q \;=\; \text{const}$$

define fibrations with invariant Lagrangian fibers diffeomorphic to $\mathbb{T}^2$. As we shall see in the next chapter, this situation is not merely artificial. On the contrary, it is typical of 'degenerate' (or 'superintegrable') systems, and it is indeed met in important cases like Kepler and the free rigid body. However, the fibration by invariant Lagrangian tori of a *nondegenerate* system is uniquely defined, because a dense subset of the invariant tori, carrying nonresonant motions, are closure of trajectories. On the other hand, this is not peculiar of nondegenerate systems: any system of harmonic oscillators with nonresonant frequencies (for instance $H = a_1 + \sqrt{2}a_2 + \sqrt{3}a_3$ on $(a, \alpha) \in \mathbb{R}^3 \times \mathbb{T}^3$) is isochronous, but all of its invariant tori are closure of trajectories.

Exercise: Consider a completely integrable system in a subset M of its phase space fibered by invariant Lagrangian tori, and assume that a dense subset of these tori are closure of trajectories. Show that:

(i) The fibration by these tori is the finest invariant fibration of any open invariant subset of M.†

(ii) In no open invariant subset of M there exist more than d independent integrals of motion.

iii) In M, the local action–angle coordinates (that is, any symplectic diffeomorphism $(a, \alpha) : U \to \hat{A} \times \mathbb{T}^d$ in which the flow is given by (3.1.1)) are uniquely defined, up to transformations (2.1.3).

(These conditions are in general not equivalent; they are however equivalent if everything is real analytic).

---

* necessary ???

† Questions related to integrals of motion are meaningful only in invariant subsets: locally, near any nonsingular point, there always exist $2d - 1$ integrals of motion (rectification theorem), which can even be chosen in such a way that d of them are in involution (see [7], pg. ....).

3.2 The Hamilton–Jacobi equation

A. The Hamilton–Jacobi equation. In this Section we discuss the connection between complete integrability and the Hamilton–Jacobi equation.

Let us consider a Hamiltonian systems defined by a Hamilton function $H(p,q)$ in an open subset of $\mathbb{R}^{2d}$, equipped with canonical coordinates (p,q). The (time independent) Hamilton–Jacobi equation for the value E of the Hamiltonian is the partial differential equation

$$H\Big(\frac{\partial W}{\partial q}(q), q\Big) \;=\; E \tag{3.2.1}$$

for the unknown function W. In the following, we tacitly assume that E is a regular value of H.

A *local integral* of the Hamilton–Jacobi equation is a function $W : Q \to \mathbb{R}$ which is defined in some open set Q and satisfies equation (3.2.1).

A *local complete integral* of the Hamilton–Jacobi equation is constituted by a pair of functions $W : \tilde{P} \times Q \to \mathbb{R}$ and $E : \tilde{P} \to \mathbb{R}$, where Q and $\tilde{P}$ are open sets in $\mathbb{R}^d$, which are such that

$$det\Big(\frac{\partial^2 W}{\partial \tilde{p} \partial q}\Big)(\tilde{p}, q) \neq 0 \qquad for\ any\ (\tilde{p}, q) \in \tilde{P} \times Q \tag{3.2.2}$$

and satisfy (3.2.1) for any given $\tilde{p} \in \tilde{P}$. (In other words, a complete integral is a family of integrals depending 'in an essential way' on d parameters.) We say that the local complete integral as above *is defined near a point* (p_*, q_*) if Q is a neighbourhood of q_* and $H(p_*, q_*) = E(\tilde{p}_*)$ for some $\tilde{p}_* \in \tilde{P}$.†

The practical usefulness of the Hamilton–Jacobi equation rests on the following, well known

Proposition 3.4 (Hamilton–Jacobi) *If the Hamilton–Jacobi equation has a local complete integral $W(\tilde{p}, q)$, $E(\tilde{p})$ near a certain point, then the function W is the generating function of a canonical transformation $(p,q) \mapsto (\tilde{p}, \tilde{q})$ which is defined in a neighbourhood of that point, is implicitly given by the equations*

$$\tilde{q}_i \;=\; \frac{\partial W}{\partial \tilde{p}_i}(\tilde{p}, q)\,, \qquad p \;=\; \frac{\partial W}{\partial q_i}(\tilde{p}, q)\,, \qquad i = 1, \ldots, d\,, \tag{3.2.3}$$

and conjugates the system to the system with Hamiltonian $E(\tilde{p})$.

Proof. On account of condition (3.2.2), equations (3.2.3) can be inverted to obtain a local diffeomorphism $(p,q) = \mathcal{C}(\tilde{p}, \tilde{q})$. The equality $\sum_i d\tilde{p}_i \wedge d\tilde{q}_i = \sum_i dp_i \wedge dq_i$ is readily verified using (3.2.3). The new Hamiltonian is $H \circ \mathcal{C}(\tilde{p}, \tilde{q}) = H(\frac{\partial W}{\partial q}(\tilde{p}, q(\tilde{p}, \tilde{q}), q(\tilde{p}, \tilde{q})) = E(\tilde{p})$ on account of (3.2.1). ∎

Under the hypotheses of Proposition 3.4, all motions are linear when viewed in the new coordinates: $\dot{\tilde{p}} = 0$, $\dot{\tilde{q}} = \frac{\partial E}{\partial \tilde{p}}(\tilde{p})$. So, finding a complete integral of the Hamilton–Jacobi equation reduces to the quadratures (in this case, to the inversion of (3.2.3)) the problem

† The use of the expression 'local' and 'near a point' is not standard; nevertheless, it is convenient in view of Proposition 3.5 (ii) below.

of integrating the equations of motion of the system. As is clear, however, all of this is meaningful only if the set Q is sufficiently large, namely, is an invariant set; we shall come back later on this point, which has a crucial importance.

The next Proposition characterizes the existence of complete integrals of the Hamilton–Jacobi equation; its proof is important, because it provides a way of computing these integrals.

Proposition 3.5 *(i) Assume that the Hamilton–Jacobi equation (3.2.1) has a local complete integral near a point (q_*, p_*). Then, the system described by the Hamiltonian H possesses d independent integrals of motion in involution in a neighbourhood of (p_*, q_*).*

(ii) Assume that a Hamiltonian system on a symplectic manifold (M, Ω) of dimension $2d$ has d integrals of motion in involution with differentials linearly independent at a point $m_* \in M$. Then, there exist canonical coordinates (p, q) in a neighbourhood of m_* such that the Hamilton–Jacobi equation for the local representative of the Hamiltonian in these coordinates has a local complete integral near $(p(m_*), q(m_*))$.

Proof. Statement (i) is obvious, the new momenta $\tilde{p}$ being integrals of motion in involution.

In order to prove statement (ii), we construct a complete integral which is such that the new momenta $\tilde{p}$ are exactly the d integrals of motion, which we denote $F = (F_1, \dots, F_d)$. Since, as already explained, the Hamiltonian H is functionally dependent on $F_1, \dots, F_d$, there is a function E which is defined in a neighbourhood U of $F(m_*)$ in $\mathbb{R}^d$ and is such that $H|_U = E \circ F$. The function E is one–half of the complete integral we are looking for.

We now introduce a (convenient) system of local canonical coordinates (p, q) near m_*, so as to write down and solve the Hamilton–Jacobi equation. In order to fulfill conditions (3.2.2) and (3.2.3), we choose these coordinates in such a way that the momenta p are transversal in m_* to the level set of F. This is indeed always possible, locally (as the very existence of the 'time–energy' coordinates (F, T) of Proposition 2.6 shows; things are different globally, i.e. in a neighbourhood of the whole torus, as we shall discuss below, but here we are taking a strictly local point of view).

We now determine the function W by integrating the equation $p = \frac{\partial W}{\partial q}$. To this end, we consider any map $\hat{q} : \mathcal{F} \to \mathbb{R}^d$, where $\mathcal{F}$ is a neighbourhood of $f_* = F(m_*)$,[†] and define

$$W(f, q) \;=\; \int_{\hat{q}(f)}^{q} \sum_i p_i dq_i \big|_{F^{-1}(f)} \,, \tag{3.2.4}$$

where the integral is extended to any curve belonging to the level set $F^{-1}(f)$ and joining $\hat{q}(f)$ and q. It is easy to verify that equation (3.2.4) defines a function W in some neighbourhood of the point (f_*, q_*). Indeed, since the level sets of F are Lagrangian, the restriction of the 1–form $\sum_i p_i dq_i$ to them is closed; hence, the integral at the r.h.s. of (3.2.4) does not depend on the choice of the curve, at least if we restrict ourselves to a sufficiently small neighbourhood of m_* in which the level sets of F are simply connected.

The proof is now concluded by observing that

$$H\Big(\frac{\partial W}{\partial q}(q, f), q\Big) \;=\; H(p, q)\Big|_{F(p,q)=f} \;=\; E(f) \,,$$

[†] Observe that $f \mapsto (p_*, \hat{q}(f))$ is a Lagrangian submanifold transversal to the level sets of F.

so that (W, E) is a local complete integral of the Hamilton–Jacobi equation near the point $(p(m_*), q(m_*))$. ∎

So, the existence of a local complete integral of the Hamilton–Jacobi equation amounts to the 'local' complete integrability of the system. However, at this local level, this fact is essentially meaningless: as already observed, every Hamiltonian system is completely integrable in a neighbourhood of every non–singular point; complete integrability, to be meaningful, must hold in *invariant* sets. Proposition 3.5 makes clear that there is no hope of integrating a system by the Hamilton–Jacobi method unless the system is completely integrable. Notwithstanding this fact, the Hamilton–Jacobi method plays an important role in the study of small perturbations of integrable systems, because of the possibility of constructing approximate (but global) solutions of it.

The proof of Proposition 3.5 illustrates a method for constructing a local complete integral of the Hamilton–Jacobi equation. Indeed, the integral at the r.h.s. of equation (3.2.4) can be computed as follows. The transversality condition $det(\frac{\partial F}{\partial p}(m_*)) \neq 0$ assures that the equation $F(p,q) = f$ can be locally inverted to obtain a map

$$p \;=\; \hat{p}(f,q) \tag{3.2.5}$$

which is defined in some open set $Q \times \mathcal{F}$, where Q is a neighbourhood of $q_* = q(m_*)$ (which can be taken simply connected) and $\mathcal{F}$ is a neighbourhood of $f_* = F(m_*)$. Hence,

$$W(f,q) \;=\; \int_{\hat{q}(f)}^{q} \sum_i \hat{p}_i(f,q) dq_i \,. \tag{3.2.6}$$

Of course, in specific cases, one can as well determine the function W by directly integrating the Hamilton–Jacobi equation. The difficulties in the two methods are equivalent, and one can practically perform all the computations (at least up to some quadratures) in those cases in which the Hamilton–Jacobi equation separates in some coordinate system (see the exercises below).

However, in general terms, the construction of the functions $E(\tilde{p})$ and $W(\tilde{p}, q)$ described in the proof consists of a finite number of inversions, integrations, and derivations of known functions ('quadratures'). This proves Liouville's theorem: *the equations of motion of a system with the maximal number of independent integrals of motion in involution can be integrated by quadratures.*

Remark: The construction outlined in the proof of Proposition 3.5 is strictly local. Indeed, there is clearly no system of coordinates (q,p) which is defined in a neighbourhood of a whole torus and has the p everywhere transversal to the torus. However, as we shall see in section 3.3.D, this construction can always be globalized to a neighbourhood of a whole torus.

Exercises: (i) A Hamiltonian system is said to be *separable* if its Hamilton–Jacobi equation, in some coordinate system (p,q), has a local complete integral $W(\tilde{p}, q)$, $E(\tilde{p})$ with $W(\tilde{p}, q) = \sum_{i=1}^{d} W_i(\tilde{p}, q_i)$. Show that this is equivalent to the fact that the inverse $\hat{p}(f,q)$ of $F(p,q) = f$ has the expression $\hat{p}(f,q) = \sum_{i=1}^{d} \hat{p}_i(f, q_i)$.

(ii) Show that the function $W(f,q)$ defined by (3.2.4) solves the Hamilton–Jacobi equation for *any* Hamilton function which has $F_1, \ldots, F_d$ as integrals of motion. Why?

B. Geometric meaning of the Hamilton–Jacobi equation. The previous approach to the Hamilton–Jacobi equation is strictly local, but all the matter has a geometric content which should help to clarify its meaning. There are two main facts:

i. Away from the critical points of H, any local integral $W(q)$ of the Hamilton–Jacobi equation defines *a Lagrangian submanifold which is contained in a level set of* H and is invariant under the flow. Indeed, the submanifold described by the equation $p = \frac{\partial W}{\partial q}(q)$ has dimension d, is contained in a level set of H, and is Lagrangian, since the restriction to it of the 1–form $\sum_i p_i dq_i$ equals dW.

ii. *Every Lagrangian submanifold* N *contained in a level set* M_H *of* H *is invariant under the flow of* X_H. Indeed, $N \subset M_H$ and $N = N^\perp$ imply $N \supset M_H^\perp$. But $M_H^\perp$ is one–dimensional and, at every regular point of H, its tangent space is generated by X_H, as follows from the fact that one has $\Omega(X_H, Y) = L_Y H = 0$ for any vector field Y tangent to M_H.

Therefore, a local complete integral of the Hamilton–Jacobi equation represents a simple Lagrangian foliation of some open set of the phase space whose leaves are contained within the level sets of the Hamiltonian (and are therefore invariant under the flow). This explains in geometric terms the link between Hamilton–Jacobi theory and complete integrability.

Let us also observe that the Hamilton–Jacobi equation itself has a global formulation, and its 'global' integrals are Lagrangian submanifolds (see [7], Appendix 7).

Exercises: (i) Show that the leaves of a foliation, if Lagrangian and invariant under the flow of X_H, are contained within the level sets of H. (This not need be true for a single invariant Lagrangian submanifold).

(ii) Consider a Hamiltonian system which is completely integrable and strictly anisochronous in a manifold M. Show that, given any point $m \in M$, there is exactly one Lagrangian submanifold which passes through m, is defined in a neighbourhood of m, and is contained in a level set of the Hamiltonian.

(iii) Prove that two Hamiltonian systems on the same symplectic manifold whose Hamiltonians have the same level sets have the same trajectories (only the time parameterization of motions differ).

3.3 Computation of the action–angle coordinates

A. Computing the action–angle coordinates. With few exceptions, the determination of the action–angle coordinates of a completely integrable system is a rather difficult task, which can be realized only partially. There is not a universal method, and one can proceed in slightly different ways. However, the operations involved are essentially those contained in the proof of the Liouville–Arnol'd theorem, namely,

(1) one begins by making a choice of d independent integrals of motion in involution, say $F_1, \ldots, F_d$, then
(2) determines the times $\tau_1, \ldots, \tau_d$ along the flows of the Hamiltonian vector fields of these functions, measured from a Lagrangian section, and finally
(3) determines the actions and the angles, what requires computing the period matrix.

The Hamilton–Jacobi method is a way of realizing step 2. Indeed, the new coordinates $\tilde{q}$ are the times along the flows of the Hamiltonian vector fields of the new momenta $\tilde{p}$, measured from a certain Lagrangian section. In practice, one succeeds in computing a complete integral of the Hamilton–Jacobi equation when this equation 'separates' in some coordinate system.

Let us mention, however, that in certain cases one can conveniently determine the time coordinates along the flows of the integrals of motion in involution 'by hand', i.e.

by integrating the differential equations $\dot{\tau}_i = \frac{\partial}{\partial F_i}$, without passing through the solution of the Hamilton–Jacobi equation. In practice, this requires that one already knows how to integrate the system, but this is often the case with classical systems, like Kepler or Euler–Poinsot, for which the solution of the equations of motion is known from elementary treatments. In such cases, which we consider in sections 3.4 and 3.5, this method is computationally somewhat simpler than the standard procedure based on the use of the Hamilton–Jacobi equation.†

There are slightly different ways of proceeding for step 3, also depending on the order of the operations involved.

B. First computing the angles. A first possibility, once the coordinates (F, τ) of step 2 are known, is that of computing the angles by normalizing the periods of the times τ, exactly as in the proof of the Liouville–Arnol'd theorem.

For this, one needs to compute a period matrix $L(f)$, or equivalently a basis $u_1(f), \ldots, u_d(f)$, for the period lattice of the action $\Phi_\tau = \Phi_{\tau_1}^{X_{F_1}} \circ \cdots \circ \Phi_{\tau_d}^{X_{F_d}}$ on each torus. This can be done as explained at the end of section 2.2.A. Namely, one consider d independent cycles $\gamma_1(f), \ldots, \gamma_d(f)$ on each torus $F = f$ and compute $(u_i(f))_j = L_{ji}(f)$ as the time that the flow of X_{F_j} acts while the action Φ makes a complete turn along a closed curve which lies on $F^{-1}(f)$ and is homologous to $\gamma_i(f)$. Then, the corresponding angles are

$$\alpha \;=\; 2\pi\, L(f)^{-1} \tau\,, \tag{3.3.1}$$

while the actions can be computed by integrating equations (2.2.10b), i.e.,

$$\frac{\partial \widehat{a}}{\partial f}(f) \;=\; \frac{1}{2\pi}\, L(f)\,. \tag{3.3.2}$$

C. First computing the actions. It is also possible to reverse the procedure, and compute the actions before computing the period matrix. This possibility rests on a remarkable expression for the actions, that we now derive.

We restrict ourselves to a neighbourhood $F^{-1}(\mathcal{F})$ of a given torus, and we assume that we know d independent cycles $\gamma_1(f), \ldots, \gamma_d(f)$ on each torus in $F^{-1}(\mathcal{F})$. Let (a, α) be a system of action–angle coordinates of the fibration by the invariant tori, with the angles $\alpha_1, \ldots, \alpha_d$ running on circles homologous to $\gamma_1, \ldots, \gamma_d$, i.e.,

$$\hat{\gamma}_i \;=\; \big\{ m \in M \;:\; a(m) = const\,,\; \alpha_j(m) = const \;\, for \;\, j \neq i\,,\; \alpha_i(m) \in S^1 \big\}\,.$$

(These coordinates are still undetermined, but their existence is granted by proposition 2.3: they are obtained from any set of action–angle coordinates by a transformation of the form (2.1.3), with a suitable matrix Z).

Let us now observe that, in a neighbourhood of an invariant torus, the symplectic 2–form is always exact, $\Omega|_U = d\Theta$.† So, the 1–form $\sum a_i d\alpha_i - \Theta$ is closed and, for any

---

† To my knowledge, this method was first used by Francoise [QUOTATION ??]; it is also used in the (still unpublished) book on Hamiltonian Mechanics by J. Moser and E. Zehnder to determine the action–angle coordinates of the n–dimensional Kepler problem.

† This follows at once from the existence of the action–angle coordinates, since $\Omega|_U = d(\sum_i a_i d\alpha_i)$.

closed curve γ, one has

$$\int_\gamma \sum a_i d\alpha_i \;=\; \int_\gamma \Theta \;+\; c([\gamma]) \tag{3.3.3}$$

$c([\gamma])$ being a constant which depends only on the homology class of γ.

Since the restriction of Θ to every Lagrangian torus is closed, the integral over the homologous cycles γ_i and $\hat{\gamma}_i$ are equal. Hence, equation (3.3.3) gives

$$2\pi\, a_i(m) \;=\; \int_{\gamma_i(m)} \Theta \;+\; c([\gamma_i(m)])\,, \qquad i = 1,\ldots,d\,.$$

But for a given i, all curves $\gamma_i(m)$, $m \in U$, are homologous, since they are homeomorphic to the homologous circles $\{a = const, \alpha_j = const, \alpha_i \in S^1\}$ in $\hat{A} \times \mathbb{T}^d$. So, $c([\gamma_i(m)])$ is constant in U and one concludes

$$a_i(m) \;=\; \frac{1}{2\pi} \int_{\gamma_i(x)} \Theta \;+\; const\,, \qquad i = 1,\ldots,d\,. \tag{3.3.4}$$

This formula (which is sometimes regarded as the very definition of the actions) can be used to compute the actions a, if one knows d independent cycles on the invariant tori (which is just the same as knowing the tori), and if one knows the restriction of Θ to the tori.

Once the actions are known, the angles can be determined in two alternative ways. If the times τ of step 2 are already known, then one can determine the period matrix as in (3.3.2) and compute the angles with (2.2.10b). However, one can also try to perform step 2 with the actions as integrals of motion, for instance by the Hamilton–Jacobi method; this has the obvious advantage that the times determined in this way are the already the angles.

D. Global aspects of the Hamilton–Jacobi method. The procedure of computing the action–angle coordinates with the Hamilton–Jacobi method, as described so far, is local, in the sense that it gives the angles in a neighbourhood of a point. However, this procedure can always be globalized to a neighbourhood of a whole torus. For this, one just needs to cover the tori with different coordinate systems, and then repeat the construction of a complete integral (relative to the same choice of the new momenta $\tilde{p}$) in each of them. The details of the construction are outlined in the following exercises.

Exercises: (i) Show that, using the Hamilton–Jacobi method, it is always possible to construct an atlas for a neighbourhood of an invariant torus constituted by a finite number N of charts with coordinates $(\tilde{p}, \tilde{q}^{(i)})$ having transition functions $\tilde{q}^{(i)} = \tilde{q}^{(i-1)}$ for $i = 2,\ldots,N$ and $\tilde{q}^{(1)} = \tilde{q}^{(N)} + T(\tilde{p})$, where $T(\tilde{p})$ are the periods. (Hints: Since the torus is compact, one can construct a denumerable atlas with coordinates $(\tilde{p}, \tilde{q}^{(i)})$ $(i = 1,\ldots,N)$ in such a way that the 'zero sections' $\tilde{q}^{(i)} = 0$ are contained in the domain of the $(i-1)$–th coordinate system $(0 \equiv N)$ and are described by the equation $\tilde{q}^{(i-1)} = T^{(i-1)}(\tilde{p})$ with certain $T^{(i)}$. Show now that the coordinates $(\tilde{p}, \tau^{(i)})$ defined by $\tau^{(1)} = \tilde{q}^{(1)}$ and $\tau^{(i)} = \tilde{q}^{(i)} + T^{(i-1)}$ $(i = 2,\ldots,N)$ are canonical.)

More generally, Ω is always exact in a neighbourhood of every Lagrangian submanifold. Indeed, by Weinstein's Lagrangian embedding theorem, if L is a Lagrangian submanifold of a symplectic manifold M, there are a neighbourhood U of L in M, a neighbourhood V of the zero section of T^*L, and a symplectic diffeomorphism from U onto V which maps L onto the zero section of T^*L (see [18]).

(ii) With reference to the previous exercise, consider local canonical coordinates $(p^{(i)}, q^{(i)})$ whose domains contain those of $(\tilde{p}^{(i)}, \tilde{q}^{(i)})$, and let $\sigma^{(i)}(\tilde{p})$ be the section used as zero of the $\tilde{q}^{(i)}$. Show that the functions

$$\begin{aligned} W^{(1)}(\tilde{p}, q^{(1)}) &= \int_{q^{(1)}(\sigma^{(1)}(\tilde{p}))}^{q^{(1)}} \sum_j \hat{p}_j^{(1)}(\tilde{p}, q^{(1)}) dq_j \\ W^{(i)}(\tilde{p}, q^{(i)}) &= \int_{q^{(i)}(\sigma^{(i)}(\tilde{p}))}^{q^{(i)}} \sum_j \hat{p}_j^{(1)}(\tilde{p}, q^{(1)}) dq_j \ + \ W^{(i-1)}(\tilde{p}, T^{(i-1)}(\tilde{p})) \,, \quad i = 1, \ldots, N \end{aligned}$$

(where $\hat{p}^{(i)}$ invert $F(p^{(i)}, q^{(i)}) = \text{const}$), are local complete integrals of the Hamilton–Jacobi equation and generate the transformations from the (original) coordinates $(q^{(i)}, p^{(i)})$ to the coordinates $(\tilde{p}, \tau^{(i)})$. (Note that $W^{(1)}, \ldots, W^{(N)}$ are not the local representatives of a (single–valued) function W defined on the neighbourhood of a torus—rather, they are the local representatives of a *multi–valued* function, i.e., a function defined on the covering of such a neighbourhood).

E. An example. As an example, we consider here the system described by the Hamiltonian

$$H(J, p, \varphi, q) \ = \ \frac{p^2}{2} \ + \ V(J, q) \tag{3.3.5}$$

on $T^*(S^1 \times \mathbb{R})$; φ and p are coordinates in $S^1 \times \mathbb{R}$, and J and p are their conjugate momenta; the coordinate φ is cyclic. The planar point in a central force field and (locally) the spherical pendulum belong to this class.

The two functions J and H are integrals of motion in involution of the system, and are everywhere independent except at the points (J, p, φ, q) with $p = 0$ and $\frac{\partial V}{\partial q}(J, q) = 0$. We assume that, for h and j in a certain subset of $\mathbb{R}^2$, and for q in a certain real interval Q (to which we tacitly restrict ourselves), the function $V(j, \cdot)$ has at least one minimum in Q and $\frac{\partial V}{\partial q}(j, q_\pm(j, h)) \neq 0$, where $q_- < q_+$ are the two 'inversions point', where $V(j, q_\pm) = h$ (see figure 1.7). In such a situation, the curve $p^2/2 + V(j, q) = h$ in the plane (p, q) is closed, and the level set $H = h$, $J = j$ is a two–dimensional torus.

We first compute the actions. Two independent cycles on the torus $H = h$, $J = j$ are

$$\begin{aligned} \gamma_1(j, h) &= \{\varphi \in S^1 \, ; \ \textit{all other coordinates kept constant}\} \\ \gamma_2(j, h) &= \{(p, q) \, : \, p^2/2 + V(j, q) = h \, ; \ \textit{all other coordinates kept constant}\} \, . \end{aligned}$$

The equation $H(j, p, \varphi, q) = h$ is inverted (locally: either for $p > 0$ or for $p < 0$; the case $p = 0$ can be recovered by continuity, see exercise (i)) by $p = \widehat{p}_\pm(j, h, \varphi, q)$, where

$$p_\pm(j, h, \varphi, q) \ = \ \pm \sqrt{2(h - V(j, q))} \, . \tag{3.3.6}$$

Hence, by (3.3.4), the actions corresponding to the two cycles γ_1 and γ_2 are

$$\begin{aligned} a_1(j, h) &= \frac{1}{2\pi} \int_{\gamma_1(j,h)} J d\varphi + \widehat{p}(j, h, \varphi, q) dq \ = \ j \\ a_2(j, h) &= \frac{1}{2\pi} \int_{\gamma_2(j,h)} J d\varphi + \widehat{p}(j, h, \varphi, q) dq \ = \ \frac{1}{\pi} \int_{q_-(j,h)}^{q_+(j,h)} \sqrt{2(h - V(j, q))} dq \, . \end{aligned} \tag{3.3.7}$$

Note that $a_2(j,h)$ is the area enclosed by the curve γ_2 in the plane (q,p). Note also that, under the hypothesis made above,

$$\frac{\partial a_2}{\partial h}(j,h) \;=\; \frac{1}{\pi}\int_{q_-(j,h)}^{q_+(j,h)} \frac{dq}{\sqrt{2(h-V(j,q))}}$$

is positive (and finite) for all j and h in the considered domain. In fact, the integral at the r.h.s. of the above equation equals $T(j,h)/2\pi$, where $T(j,h)$ is the period of the motion on the curve $p^2/2+V(j,q)=h)$. Therefore, there is a function $E(a_1,a_2)$ such that $h=E(j,a_2)$ solves the equation $a_2(j,h)=const.$

We construct now the angles α_1 and α_2 conjugate to a_1 and a_2 with the Hamilton–Jacobi method, choosing as new momenta the two actions a_1 and a_2. We choose the Lagrangian section defined by $\varphi=0$ and $q=q_-(h,j)$ (hence, $J=j$ and $p=0$; this section lies at the boundary of the two chart domains, but a little reflection shows that this is is not a problem). The generating function $W_\pm$ is (see (3.2.4))

$$W_\pm(\varphi,q,a_1,a_2) \;=\; \int_{(0,q_)}^{(\varphi,q)} jd\varphi+pdq \;=\; j\varphi \;\pm\; \sqrt{2}\int_{q_-}^{q}\sqrt{h-V(a_1,q)}dq \tag{3.3.8}$$

where $\hat{p}_\pm$, q_- and E must be considered as functions of a_1 and a_2. (In all formulas is understood that the upper sign holds for $p>0$ while the lower one holds for $p<0$). Therefore, the angles are

$$\begin{aligned}\alpha_1 \;&=\; \frac{\partial W_\pm}{\partial a_1} \;=\; \varphi \;\pm\; \int_{q_-}^{q}\frac{\frac{\partial E}{\partial j}-\frac{\partial V}{\partial j}}{\sqrt{2(E-V)}}dq\\ \alpha_2 \;&=\; \frac{\partial W_\pm}{\partial a_2} \;=\; \pm\frac{\partial E}{\partial a_2}\int_{q_-}^{q}\frac{dq}{\sqrt{2(E-V)}}\;.\end{aligned} \tag{3.3.9}$$

Note that

$$\alpha_2 \;=\; \frac{2\pi}{T(j,h)}\,\tau_2$$

where $\tau_2=\pm\int_{q_-}^{q}[2(h-V)]^{-1/2}dq$ is the time necessary for the coordinate q to go, during the motion of the system, from the point q_- to the point q. (The point q_- obviously corresponds to the two times $\pm T/2$.) Correspondingly,

$$\alpha_1 \;=\; \varphi \;+\; \frac{\partial H}{\partial j}\,\tau_2 \;-\; \varphi_H(\tau_2)\,,$$

where $\varphi_H(\tau_2)$ is the angle of which φ rotates, during the motion of the system, in the time τ_2, that is

$$\varphi_H(\tau_2) \;=\; \pm\int_{q_*}^{q(\tau_2)}\frac{\frac{\partial V}{\partial J}(j,q)}{\sqrt{2(h-V(j,q))}}\,dq\,,$$

where $q(\tau_2)$ is the value of the coordinate q at time τ_2, with $q=q_-$ initially. (With the particular choices made here, the angles α_1 and α_2 take values between $-\pi$ and π).

Exercises: (i) Show that α_1 and α_2 defined by (3.3.9) are continuous (and differentiable) also for $q=q_-$. What about $q=q_+$? Show that each function W_+ and W_- as in (3.3.8) can be extended to a differentiable function which is defined in an open set which contains both the points q_- and q_+.

(ii) Determine the angles without the Hamilton–Jacobi method, by normalizing the periods of the times τ_1 and τ_2 along the flows of, respectively, X_J and X_H. (Hints: Measure the times from the section considered above. Thus,

$$\tau_1 \;=\; \varphi - \varphi_H(\tau_2)\,.$$

The periods corresponding to the cycles γ_1 and γ_2 are $u_1 = (2\pi, 0)$ and $u_2 = (2\pi\partial a_2/\partial j, 2\pi\partial a_2/\partial h) = (\varphi_H(T), T)$. Hence,

$$L(j,h) \;=\; 2\pi \begin{pmatrix} 1 & \frac{\partial a_2}{\partial j} \\ 0 & \frac{\partial a_2}{\partial h} \end{pmatrix}, \qquad L(j,h)^{-1} \;=\; \frac{1}{2\pi} \begin{pmatrix} 1 & -\varphi_H(T)/T \\ 0 & 2\pi/T \end{pmatrix},$$

with $\varphi_H(T)/T = \frac{\partial a_2}{\partial j} / \frac{\partial a_2}{\partial h} = -\frac{\partial E}{\partial a_1}$.)

(iii) Compute the action–angle coordinates of the harmonic oscillator with the Hamilton–Jacobi method. Compare with section 1.1.

(iv) Show that two of the three actions of the symmetric Euler–Poinsot system of section 1.8 (with $I_1 = I_2$) and of the Lagrange top of section 1.9 can be taken to be m_z and m_3. Find an expression for the third action in each case.

(v) Generalize the analysis of this subsection to the case of a Hamiltonian system with d degrees of freedom having $d-1$ cyclic angular coordinates.

3.4 Kepler

A. The planar case. In this Section we construct the action–angle coordinates for the Kepler problem. For the sake of clearness, we consider first the planar case, which is somewhat simpler.

We begin by recalling a few elementary facts about the system. Using polar coordinates (r, φ) in the plane, with conjugate momenta (p_r, p_φ), the Hamiltonian is

$$H(p_r, r, p_\varphi, \varphi) \;=\; \frac{p_r^2}{2\mu} \;+\; \frac{p_\varphi^2}{2\mu r^2} \;-\; \frac{k}{r}\,,$$

where μ is the mass of the point and k is a positive constant.

The Hamiltonian H and the component $G := p_\varphi$ of the angular momentum in the direction orthogonal to the plane are first integrals, and we restrict our considerations to the subset M of the phase space where they are independent and their level sets are compact. As seen in sections 1.6 and 2.1, M is defined by the conditions $H < 0$, $G \neq 0$, $p_r \neq 0$, namely, it contains all states with negative energy but the rectilinear and circular motions. Moreover, the 'energy–momentum' map $H \times G$ maps M onto the set

$$B \;=\; \bigl\{(H, G) \in \mathbb{R}^2 :\, H < 0,\, 0 < |G| < k\sqrt{-2H/\mu}\bigr\}\,,$$

and so defines a fibration $M \to B$ with fiber $\mathbb{T}^2$, see figure 1.10. The two curves $G = \pm\sqrt{-2H/\mu}$ on the boundary of B correspond to the circular orbits.

The fact that, in the set M, the level sets $M_{H,G}$ of $H \times G$ are two–dimensional tori is easily understood on the basis of a simple picture, which will be useful in the sequel. As is well known, all orbits with equal energy $H < 0$ and equal angular momentum $G \neq 0$ are cofocal ellipses with equal semi–major axis $b(H)$ and equal eccentricity $e(H, G)$. Specifically, one has

$$b(H) \;=\; \frac{k}{-2H}\,, \qquad e(H, G) \;=\; \sqrt{1 + \frac{2HG^2}{\mu k^2}}\,. \tag{3.4.1}$$

So, for $(H, G) \in B$, the level set $M_{H,G}$ is the collection of all cofocal ellipses with given semi–major axis $b(H) > 0$ and eccentricity $e(H, G) < 1$. This set is a 'circle of ellipses', namely, a two–dimensional torus, see figure 3.1. (This set degenerates into a circle if the eccentricity is one, but circular orbits are excluded from M). Let us also recall, from the elementary treatments, that all orbits on the torus $M_{H,G}$ have the same period $T(H)$, given by

$$T(H) \;=\; 2\pi\sqrt{\frac{\mu}{k}}\, a(H)^{3/2}\,. \tag{3.4.2}$$

We now pass to the construction of the action–angle coordinates. To begin with, we construct the 'times' along the flows of X_H and X_G. This requires the choice of a Lagrangian section $\sigma : B \to M$, that is, roughly speaking, of a point on every torus $M_{H,G}$. Now, since $M_{H,G}$ is a union of Keplerean ellipses, a point on it is determined by choosing a Keplerean ellipse and a point on this ellipse. So, we choose the point $\sigma(H, G) \in M_{H,G}$ as follows: we pick the ellipse whose perihelion lies on the positive x–axis (i.e., its coordinate φ is zero), and on this ellipse we pick the perihelion. A moment of reflection, and a glance at figure 3.1, will convince that this defines one and only one point $\sigma(H, G)$ on every torus $M_{H,G}$. In coordinates $(p_r, r, p_\varphi, \varphi)$ we have

$$\sigma(H, G) \;=\; \bigl(0, c(H, G), G, 0\bigr)\,, \tag{3.4.3}$$

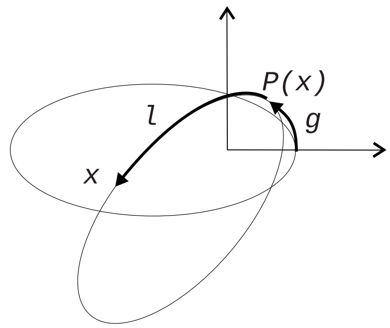


Figure 3.1

with $c(H,G) = [1-e(H,G)]b(H)$; this shows that $\sigma(B)$ is a Lagrangian smooth submanifold. We now compute the times $\tau_1(m)$ and $\tau_2(m)$ (modulo periods) needed to go from $\sigma(H,G)$ to any other point $x \in M_{H,G}$ along the flows of X_G and X_H. Since $X_G = \frac{\partial}{\partial\varphi}$, its flow is a shift of the coordinate φ, while the flow of X_H is obviously the Keplerean motion, which takes place along the ellipses. Thus, we reach the point x by acting with the flow of X_G for a time

$$\tau_1(x) \;=\; \varphi(P(x))\,, \qquad P(x) \;=\; \text{the perihelion of the ellipse containing } x\,,$$

so as to reach the perihelion of the Keplerean ellipse which contains x, and then acting with the flow of X_H for a time

$$\tau_2(x) \;=\; \text{the time necessary to go from } P(x) \text{ to } x \text{ in the Keplerean motion}$$

(see figure 3.1). A basis of the period lattice of the action $\Phi^{X_G} \circ \Phi^{X_H}$ is clearly constituted by the two vectors

$$u_1 \;=\; (2\pi, 0)\,, \qquad u_2 \;=\; (0, T(H))\,, \tag{3.4.4}$$

to which correspond the period matrix

$$L \;=\; \begin{pmatrix} 2\pi & 0 \\ 0 & T(H) \end{pmatrix}$$

and, on account of (3.3.1), the angles

$$\alpha_1 \;=\; \tau_1\,, \qquad \alpha_2 \;=\; \frac{2\pi}{T(H)}\tau_2\,.$$

In Celestial Mechanics, the angle $\alpha_1(x) = \varphi(P(x))$ is called the *argument of the perihelion* of x, and is often denoted by $g(x)$. The angle $\alpha_2(x)$ is the time needed to reach x from the perihelion $P(x)$ in a motion of the system, normalized to 2π; in Celestial Mechanics, such an angle is called the *mean anomaly* of the point x, and is often denoted $l(x)$.†

† Besides any formal argument, the fact that the two angles g and l are coordinates on each torus $M_{H,G}$ should be completely obvious in the light of the previous representation of $M_{H,G}$ as a circle of ellipses, i.e., of figure 3.1.

Finally, we compute the actions conjugate to the angles $\alpha_1 = g$ and $\alpha_2 = l$. Equation (3.3.2) gives

$$\frac{\partial a_1}{\partial G} = 1\,, \qquad \frac{\partial a_1}{\partial H} = 0\,, \qquad \frac{\partial a_2}{\partial G} = 0\,, \qquad \frac{\partial a_2}{\partial H} = \frac{T(H)}{2\pi}\,.$$

So, using the expression (3.4.2) of $T(H)$, we immediately compute

$$\begin{aligned} a_1 &= G \\ a_2 &= \frac{1}{2\pi}\int_0^H T(E)dE = \sqrt{\frac{\mu k^2}{-2H}} =: L(H)\,. \end{aligned} \tag{3.4.5}$$

Written in the action–angle coordinates (G, L, g, l), the Hamiltonian is

$$H(L) = -\frac{\mu k^2}{2L^2}\,,$$

so Hamilton's equations give $\dot{l} = \mu k^2/L^3$ (Kepler's third law) and $\dot{g} = \dot{L} = \dot{G} = 0$ (the Keplerean ellipses neither change shape nor rotate).

We mention here without proof that the coordinates (G, L, g, l) provide a global coordinate system for the set M; they take values in the set $\{(G, L) \in \mathbb{R}^2 : 0 < |G| < L\} \times \mathbb{T}^2$, as is seen by observing that $L(H)$ is the norm of the angular momentum of the circular orbit with energy H, that is, the supremum of $|p_\varphi|$ over a level set of H. However, the action–angle coordinates are obviously singular on the border of M, that is, in correspondence of the collisions and of circular orbits.

Remark: The singularity on the circular orbits deserves a special consideration, since it does not reflect any property of the dynamics of the system. In fact, *all* orbits are circles and the foliation by these circles has no singularity at all on the circular orbits. The situation is completely analogous to that met with the isotropic oscillator of section 1.7, that is, the foliation by the orbits is a fibration. The singularity of the action–angle coordinates on the circular orbits reflects a singularity of the foliation by two–dimensional tori, which are constructed by grouping together the periodic orbits. It is only this grouping (which has no dynamical meaning) which develops singularities: near the circular orbits, the foliation by the two–dimensional tori looks exactly as in figure 1.6. We shall meet other examples of this situation, which is in fact typical of degenerate systems; a general explanation is given in chapter 4.

Exercise: For comparison, compute the action–angle coordinates for the planar Kepler system with the Hamilton–Jacobi method. (Sketch: The cycle corresponding to the period u_1 is

$$\gamma_1 = \{\varphi \in S^1,\ \text{all other coordinates constant}\}\,,$$

while the cycle corresponding to u_2 is a Keplerean ellipse, in which φ runs once on γ_1 and at the same time the point (p_r, r) runs once over the curve

$$\gamma_r(H, G) = \left\{(p_r, r) \in \mathbb{R}^2 : \frac{p_r^2}{2} + V(G, r) = H\right\}$$

where $V(G, r) = G^2/(2\mu r^2) - k/r$. Denoting $\Theta = p_\varphi d\varphi + p_r dr$, we find

$$\begin{aligned} a_1(H, G) &= \frac{1}{2\pi}\int_{\gamma_1}\Theta = G \\ a_2(H, G) &= \frac{1}{2\pi}\int_{\gamma_1}\Theta + \frac{1}{2\pi}\int_{\gamma_r(H,G)}\Theta = G + \frac{1}{\pi}\int_{r_-(H,G)}^{r_+(H,G)}\sqrt{2\mu(H - V(G, r))}dr\,, \end{aligned}$$

where $r_\pm$ are the two roots of $p^2/2 + V(G,r) = H$. The computation of the last integral, whose value is $L - G$, requires some work. One can proceed in different ways, for instance by using the so called eccentric anomaly (see [26]), or else using the residue technique as in [27]. The determination of the angles $\alpha_1 = \frac{\partial W}{\partial G}$ and $\alpha_2 = \frac{\partial W}{\partial L}$, where $W(L,G,r,\varphi) = \int p_\varphi d\varphi + p_r dr$, requires the computation of similar integrals.)

B. The spatial case. We consider now the spatial Kepler system. The treatment goes along the same lines of the planar case, but it is just a little more cumbersome.

Let us fix a system of cartesian coordinates $q = (x,y,z)$ with the origin in the attracting force (but otherwise arbitrary). If $p = (p_x, p_y, p_z)$ are the conjugate momenta, the Hamiltonian is

$$H(q,p) \;=\; \frac{\|p\|^2}{2\mu} \;-\; \frac{k}{\|q\|}\,.$$

As already noticed in section 1.6, the angular momentum vector $m = q \times p$ is an integral of motion, and one can form sets of three integrals of motion in involution by taking H, the norm $G := \|m\|$ of m, and one component of m, for instance the component $J := m_z$ on the z-axis.

It can be shown† that the three functions H, G, J are everywhere independent and have compact common level sets in the subset of the phase space where

$$H \;<\; 0\,, \qquad 0 \;<\; G \;<\; k\sqrt{\frac{-2H}{\mu}}\,, \qquad |J| \;<\; G\,.$$

We shall denote by M_z this set, since it consists of all Keplerean motions but those on circular orbits ($G = k\sqrt{-2H/\mu}$) and those with 'zero inclination', that is, those whose angular momentum is aligned with the z–axis ($J = \pm G$).

In M_z, all level sets $M_{H,G,J}$ are diffeomorphic to $\mathbb{T}^3$. It is possible to directly verify this fact by introducing spherical coordinates (r,φ,θ) (relative to the cartesian coordinates (x,y,z) introduced above). The three functions H, G and J have the expressions

$$\begin{aligned} H \;&=\; \frac{p_r^2}{2\mu} \;+\; \frac{1}{2\mu\, r^2}\Big(p_\theta^2 \;+\; \frac{p_\varphi^2}{\sin^2\theta}\Big) \;-\; \frac{k}{r} \\ G \;&=\; \sqrt{p_\theta^2 \;+\; \frac{p_\varphi^2}{\sin^2\theta}} \\ J \;&=\; p_\varphi\,, \end{aligned} \tag{3.4.6}$$

(in order to verify the expression for G, use $G^2 = \|q\|^2\|p\|^2 - (q\cdot p)^2 = 2\mu\|q\|^2(H + k/r) - (rp_r)^2$) so that

$$\begin{aligned} M_{H,G,J} \;=& \Big\{(p_r,r) \in \mathbb{R}^2 : \frac{p_r^2}{2\mu} + V(G,r) = H\Big\} \times \Big\{(p_\varphi,\varphi) \in \mathbb{R} \times S^1 : p_\varphi = J\Big\} \\ &\times \Big\{(p_\theta,\theta) \in \mathbb{R}^2 : p_\theta^2 \;+\; \frac{J^2}{\sin^2\theta} = G^2\Big\}\,. \end{aligned}$$

where $V(G,r) = G^2/(2\mu\, r^2) - k/r$. If $H < 0$ and $0 < G < k\sqrt{-2H/\mu}$, the first and last factors are circles.

† But the computation is not obvious; the details can be found in the (still unpublished) book on Hamiltonian Mechanics by J. Moser and E. Zehnder.

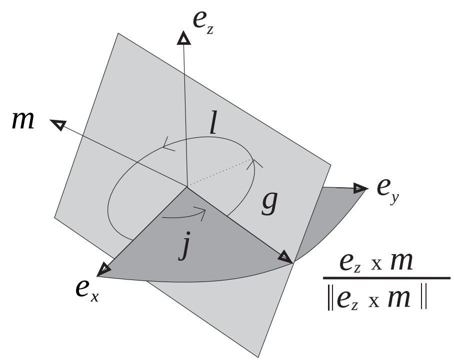


Figure 3.2

As in the planar case, we can give a pictorial illustration of this fact (see figure 3.2). Once the angular momentum vector and the energy have been fixed, one is left with all Keplerean ellipses, with given greater semiaxis and eccentricities, which lie in the plane orthogonal to m. As seen in the previous subsection, these ellipses form a two–dimensional torus. The set $M_{H,G,J}$ is thus obtained by considering all possible orientations of m, compatible with $m_z = J$. If m is not aligned with e_z, then there is a circle of vectors m such that $m_z = J$, so $M_{H,G,J}$ is a 'circle of two–dimensional tori'.

These three–dimensional tori collapse to two–dimensional tori when $|J| = G$ or $0 < G = k\sqrt{-2H/\mu}$, and to circles when both these conditions are satisfied. This is similar to what happens in the planar case, but the novelty is constituted by the orbits with zero inclination. If $|J| = G$ there is only one value of m such that $m_z = J$, so that $M_{H,G,J}$ is a torus of dimension two. It should be stressed that orbits with zero inclinations do not have any special dynamical meaning, because of the arbitrariness of the z–axis. What happens is just that it is not possible to globally group together the periodic orbits so as to form three-dimensional tori: in the process, some singularities are produced somewhere—in a direction which can be chosen arbitrarily, but cannot be avoided.

In Celestial Mechanics, one introduces three angular coordinates (j, g, l) on the tori $M_{H,G,J}$ which are defined in the following way (see figure 3.2). Given a point $x \in M_{H,G,J}$, consider the Keplerean ellipse containing it. The plane to which this ellipse belongs is fixed by the angle j between the vectors e_x and $e_z \times m$ (e_x and e_z being unit vectors in the positive directions of the axes x and z, respectively). The angle j is called *longitude of the ascending node*, since the line $e_z \times m$ intersects the orbit at the point at which the orbit crosses the plane xy from 'down' to 'up'. The position of the Keplerean ellipse on its plane is fixed by the *argument of the perihelion* g, measured from the ascending node. Finally, the position of the point x on its ellipse is determined by the mean anomaly l, measured from the perihelion. (The fact that these three functions are angular coordinates on the tori is quite obvious, as is seen on figure 3.2; here, we will show that they are indeed the angles we are looking for).

We now construct the action–angle coordinates for the system in the set M_z. We choose the Lagrangian section σ defined, in the spherical coordinates $(p_r, r, p_\varphi, \varphi, p_\theta, \theta)$, by

$$\sigma(H,G,J) \;=\; (0, c(H,G), J, 0, \sqrt{G^2-J^2}, \pi/2)$$

with c as in (3.4.3). Note that $\sigma(H,G,J)$ is the perihelion of that Keplerean ellipse on $M_{H,G,J}$ whose ascending node and perihelion both stay on the axis $x>0$.

Lemma 3.6 *Starting from $\sigma(H,G,J)$ one reaches any other point $x \in M_{H,G,J}$ by subsequently applying:*

- *The flow of X_J for a time $\tau_1(x) = j(x)$.*
- *The flow of X_G for a time $\tau_2(x) = g(x)$.*
- *The flow of X_H for a time $\tau_3(x) = 2\pi l(x)/T(H)$.*

Proof. Since $X_J = \frac{\partial}{\partial\varphi}$, the flow of X_J is a shift of the angle φ and so

$$\Phi_j^{X_J}(c(H,G),0,0,J,\pi/2,\sqrt{G^2-J^2}) \;=\; (c(H,G),0,j,J,\pi/2,\sqrt{G^2-J^2})\,.$$

The determination of the flow of X_G is a little more involved, but is simplified if one proceeds in the following way. Consider a new system of spherical coordinates (r', φ', θ'), relative to a cartesian frame (x', y', z') with the origin in the attracting center, the axis z' aligned with m, and the axis x' aligned with the ascending node (refer to figure 3.2). Since $p_{\varphi'} = m_{z'} = \|m\|$ and, on account of (3.4.6), $p_{\theta'}^2 = G^2 - p_{\varphi'}^2/\sin^2\theta' = 0$, the point $\Phi_j^{X_J}(\sigma(H,G,J))$ has coordinates

$$(p_{r'}, r', p_{\varphi'}, \varphi', p_{\theta'}, \theta') \;=\; (0, c(H,G), G, 0, 0, \pi/2)\,.$$

On the other hand, since $G^2 = p_{\theta'}^2 + p_{\varphi'}^2/\sin^2\theta'$, the flow of $X_{G^2/2}$ is defined by the differential equations

$$\begin{aligned} \dot r' &= 0\,, & \dot p_{r'} &= 0 \\ \dot\varphi' &= \frac{p_{\varphi'}}{\sin^2\theta'}\,, & \dot p_{\varphi'} &= 0 \\ \dot\theta' &= p_{\theta'}\,, & \dot p_{\theta'} &= -\frac{p_{\varphi'}\cos\theta'}{\sin^3\theta'}\,. \end{aligned}$$

This flow leaves the planes $\{r' = const,\, p_{r'} = 0,\, \theta' = \pi/2,\, p_{\theta'} = 0\}$ invariant, and its restriction to these planes is defined by the differential equations

$$\dot\varphi' \;=\; G\,, \qquad \dot p_{\varphi'} \;=\; 0\,.$$

Therefore, the flow of X_G leaves these planes invariant, too, and its restriction to them is defined by

$$\dot\varphi' \;=\; 1\,, \qquad \dot p_{\varphi'} \;=\; 0\,,$$

i.e., it is a shift of the coordinate φ'. Therefore, the point $\Phi_{g(x)}^{X_G}(\Phi_{j(x)}^{X_J}(\sigma(H,G,J)))$ is the perihelion of the ellipse containing x. Acting on this point with the flow of X_H for the time $l(x)$, one reaches x. ∎

The construction of the action–angle coordinates is now completed by observing that three independent periods of the action $\Phi^{X_G} \circ \Phi^{X_J} \circ \Phi^{X_H}$ are

$$u_1 \;=\; (2\pi, 0, 0)\,, \qquad u_2 \;=\; (0, 2\pi, 0)\,, \qquad u_3 \;=\; (0, 0, T(H))\,.$$

One immediately verifies that the corresponding action–angle coordinates are $a_1 = J$, $a_2 = G$, $a_3 = L$, $\alpha_1 = j$, $\alpha_2 = g$, $\alpha_3 = l$, where L is defined by (3.4.5). The action–angle coordinates (J, G, L, j, g, l) are known as the *Delaunay elements*. As it turns out, they are globally defined in the subset M_z of the phase space where $H < 0$, $|J| < G$, $G < L$.

3.5 The symmetric Euler–Poinsot system

A. Generalities. As a second example, we consider the Euler–Poinsot system of Section 1.8 (we freely use the notations introduced there). For simplicity, we restrict ourselves to the case of a symmetric body, which has two moments of inertia equal, say $I_1 = I_2$, so e_3 is the axis of symmetry of the ellipsoid of inertia. In the symmetric case one completely determines the action–angle coordinates, while in the general case of a tri–axial body this is done only up to quadratures (see the exercises).

As seen in Section 1.8, all common level sets of H and m^s in the set

$$M \;=\; \{(\mathcal{R}, m^b) \,:\, m^b \times e_3^b \neq 0\} \;=\; SO(3) \;\times\; (\mathbb{R}^3 \setminus \mathbb{R}e_3^b)$$

are diffeomorphic to the two–dimensional torus. This set contains all the states of the system, except the equilibria ($m^b = 0$) and the steady rotations about the symmetry axis of inertia (where m is parallel to e_3). Here, we will compute action–angle coordinates in a subset of the set M.

In order to construct a set of three integrals involution, let us observe that the constancy in space of the angular momentum vector implies that its norm and its projection along *any* direction e_z fixed in space are integrals of motion. In the body representation, these two functions are

$$G(\mathcal{R}, m^b) \;=\; \|m^b\|\,,$$
$$J(\mathcal{R}, m^b) \;=\; \mathcal{R}m^b \cdot e_z^s\,.$$

Using for instance the Euler angles (see equation (3.5.2) below), one verifies that the two functions G and J are in involution. Hence, together with the Hamiltonian H, they form a set of three integrals in involution. However, in the symmetric case, also the projection

$$L(\mathcal{R}, m^b) \;=\; m_3$$

of the angular momentum vector on the symmetry axis e_3 is an integral of motion. This function is not independent of H and m^s, since

$$H \;=\; \frac{1}{2I_1}(m_1^2 + m_2^2) \;+\; \frac{1}{2I_3}m_3^2 \;=\; \frac{1}{2I_1}\|m^s\|^2 \;+\; \left(\frac{1}{2I_3} - \frac{1}{2I_1}\right) m_3^2\,. \tag{3.5.1}$$

For the construction of the action–angle coordinates, it is convenient to consider from the outset G, L and J, rather than H, G and J, since, as is immediately seen, dH and dG are not independent on the stationary rotations about the axes orthogonal to e_3 (and moreover L, not H, is one of the actions).

One can verify (with a somewhat tedious computation) that G, L, and J are independent in the set

$$M_z \;=\; \left\{(\mathcal{R}, m^b) \in SO(3) \times \mathbb{R}^3 \,:\, m^b \times e_z^b \neq 0\,,\; m^b \times e_3^b \neq 0\right\}\,,$$

which is obtained by removing from M all states with m parallel to the chosen axis e_z. (As is obvious, because of the arbitrariness of this axis, these states do not have any dynamical meaning). The map (G, L, J) maps M_z onto the set

$$B \;=\; \{(G, L, J) \in \mathbb{R}^3 \,:\, |L| < G \;\; |J| < G\,\}$$

and, as it turns out, its level sets $M_{G,L,J}$ are diffeomorphic to $\mathbb{T}^3$. Indeed, $M_{G,L,J}$ is the union of all two–dimensional invariant tori with $H = H(G, L)$ and m^s such that $m_z = J$; thus, if $J < G$, the angular momentum vector in space lies on a circle, and so $M_{G,L,J}$ is a 'circle of two–dimensional tori'.

Proceeding more formally, one can verify that $M_{G,L,J} \approx \mathbb{T}^3$ by using the Euler angles. To this end, we introduce a reference frame e_x, e_y, e_z fixed in space, with the axis e_z coinciding with the axis used for the definition of J, and we denote by (φ, ψ, θ) the Euler angles relative to this spatial frame and to the body frame $\{e_1, e_2, e_3\}$, defined as in figure 1.13. Denoting $(p_\varphi, p_\psi, p_\theta)$ the conjugate momenta, it is not difficult to verify that the functions G, L, J have the expressions

$$G \; = \; \sqrt{p_\psi^2 \; + \; p_\theta^2 \; + \; \frac{(p_\varphi - p_\psi \cos\theta)^2}{\sin^2\theta}} \,, \qquad L \; = \; p_\psi \,, \qquad J = p_\varphi \,. \tag{3.5.2}$$

(The expression for G can be obtained by observing that $H = \frac{m_1^2+m_2^2}{2I_1} + \frac{m_3^2}{2I_3}$, so $G^2 = 2I_1 H - (I_1 - I_3)L^2/I_3$, and by replacing H with its expression in terms of the Euler coordinates, which is that of the kinetic part of the Hamiltonian of the Lagrange top given in (1.9.1)). Correspondingly, the invariant tori have the parameterization

$$\begin{aligned} M_{G,L,J} \; &= \big\{(p_\varphi, \varphi) \in \mathbb{R} \times S^1 : p_\varphi = J\big\} \; \times \; \big\{(p_\psi, \psi) \in \mathbb{R} \times S^1 : p_\psi = L\big\} \\ &\times \; \big\{(p_\theta, \theta) \in \mathbb{R}^2 : p_\theta^2 + V(\theta, L, J) = G^2 - L^2\big\} \end{aligned} \tag{3.5.3}$$

where

$$V(\theta, L, J) \; = \; \frac{(J - L\cos\theta)^2}{\sin^2\theta} \,.$$

For $(G, L, J) \in B$, the last factor of $M_{G,L,J}$ is a circle, as illustrated in figure 3.3.

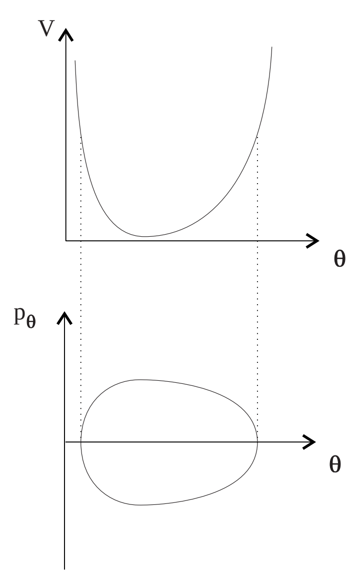


Figure 3.3

B. The action–angle coordinates. We construct now a set of action–angle coordinates of the fibration

$$(G,L,J)\ :\ M_z\ \to\ B\,.$$

We begin by choosing a Lagrangian section $\sigma : B \to M_z$ of this fibration, which will serve as origin of the angles. For greater clarity, we shall represent here the points $(\mathcal{R}, m^b)$ of M by giving the spatial representatives of the three axes of inertia e_1, e_2, e_3 and of the angular momentum m. Furthermore, we shall identify all vectors with their representatives in the chosen spatial basis $\{e_x, e_y, e_z\}$.

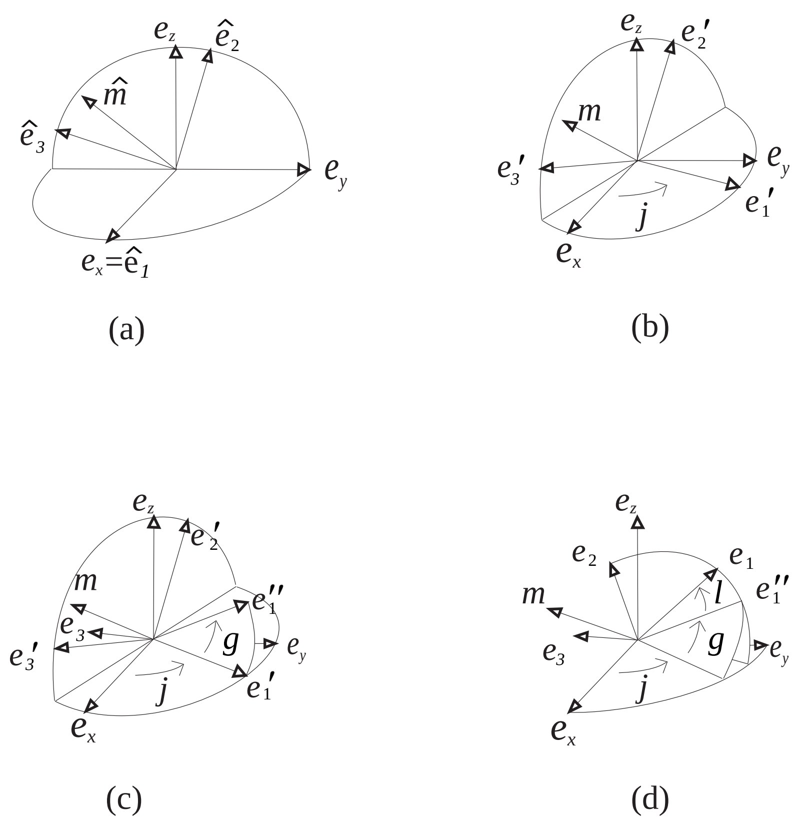


Figure 3.4

Lemma 3.7 *On the torus $M_{G,L,J}$, take the point $\sigma(G,L,J) = (\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{m})$ defined by the following conditions:*

- *$\hat{m}$ belongs to the half plane $x \in \mathbb{R}$, $y \leq 0$ (i.e, $e_z \times \hat{m}$ is oriented as e_x)*
- *$\hat{e}_1 = e_x$*

- *$\hat{m} \times \hat{e}_3$ is oriented like e_x*

(see figure 3.4.a). Then, $\sigma : B \to M_z$ is a Lagrangian section of $(G, L, J) : M_z \to B$.

Proof. A moment of reflection will show that on every torus there is exactly one point which satisfies the three conditions defining $\sigma(G, L, J)$. Its Euler coordinates are

$$(p_\varphi, \varphi, p_\psi, \psi, p_\theta, \theta) \; = \; (J, 0, L, 0, 0, \theta_+(G, J, L))\,, \tag{3.5.4}$$

where $\theta_- < \theta_+$ are the two roots of the equation $p_\theta^2/2 + V(\theta, J, L) = G^2 - L^2$, see figure 3.3. (Note that, if e_3, m and e_z lie in the same plane, the angle θ is either at its maximum θ_+ or at its minimum θ_-; requiring that $e_z \times m$ is directed like $m \times e_3$ forces $\theta = \theta_+$. It is also easy to verify that, if $|J| < G$ and $|L| < G$, then $0 < \theta_- < \theta_+ < \pi$, so the use of the Euler angles is justified). Equation (3.5.4) shows that σ is Lagrangian. Smoothness follows from the implicit function theorem. ∎

Lemma 3.8 *Any point $(e_1, e_2, e_3, m) \in M_{G,L,J}$ is reached from $(\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{m})$ by applying*
- *the flow of X_J for a time equal to the angle j between e_x and $e_z \times m$;*
- *the flow of X_G for a time equal to the angle g between $e_z \times m$ and $m \times e_3$;*
- *the flow of X_J for a time equal to the angle l between $m \times e_3$ and e_1.*

Proof. Since $J = p_\varphi$, the flow of X_J is a shift of the angle φ. Hence, the point $\Phi_j^{X_J}(\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{m})$ has Euler coordinates $(J, j, L, 0, 0, \theta_+)$. Note that this point is obtained by rotating the body *and* the angular momentum vector of an angle j around the axis e_z, so that, in the new position, the angular momentum has reached its value m, see figure 3.4.b. (The fact that the angular momentum is rotated in space by a shift of the coordinate φ —all other Euler coordinates staying constant— can be seen by recalling that the components of the angular velocity in the body base $\{e_1, e_2, e_3\}$ are independent of the angle φ; one has, indeed,

$$\omega_1 \; = \; \dot{\varphi} \, sin\,\theta \; sin\,\psi + \dot{\theta} \; cos\,\psi\,, \quad \omega_2 \; = \; \dot{\varphi} \; sin\,\theta \; cos\,\psi - \dot{\theta} \; sin\,\psi\,, \quad \omega_3 \; = \; \dot{\varphi} \; cos\,\theta + \dot{\psi}\,.$$

Hence, during a shift of the coordinate φ, the angular velocity remains fixed with respect to the body, and so rotates in space. The same happens to the angular momentum, because the inertia matrix in the body base has constant entries.) We denote $\Phi_j^{X_J}(\hat{e}_1, \hat{e}_2, \hat{e}_3, \hat{m}) = (e_1', e_2', e_3', m)$.

In order to study the flow of X_G we choose a suitable system of Euler angles $(\varphi', \psi', \theta')$. Using (3.5.2), we see that the flow of $X_{G^2/2}$ is determined by the differential equations

$$\begin{aligned}
\dot{p}_{\varphi'} \; &= \; 0\,, & \dot{\varphi}' \; &= \; \frac{p_{\varphi'} - p_{\psi'} \, cos\,\theta'}{sin^2\,\theta'} \\
\dot{p}_{\psi'} \; &= \; 0\,, & \dot{\psi}' \; &= \; p_{\psi'} \; - \; \frac{(p_{\varphi'} - p_{\psi'} \, cos\,\theta')\, cos\,\theta'}{sin^2\,\theta'} \\
\dot{p}_{\theta'} \; &= \; - \frac{(p_{\varphi'} - p_{\psi'} \, cos\,\theta') p_{\psi'}}{sin\,\theta'} \; + \; \frac{(p_{\varphi'} - p_{\psi'} \, cos\,\theta')^2 \, cos\,\theta'}{sin^3\,\theta'}\,, & \dot{\theta}' \; &= \; p_{\theta'}\,.
\end{aligned}$$

Therefore, for given G, L, J, the subspace

$$p_{\varphi'} \; = \; G\,, \quad p_{\psi'} \; = \; L\,, \quad p_{\theta'} \; = \; 0\,, \quad cos\,\theta' \; = \; L/G\,, \tag{3.5.5}$$

is invariant under the flow of $X_{G^2/2}$, and the restriction of $X_{G^2/2}$ to this subspace is $G\frac{\partial}{\partial \varphi'}$. Hence, X_G leaves the subspace (3.5.5) invariant, too, and its restriction is $\frac{\partial}{\partial \varphi'}$. Thus, if we choose the Euler angles $(\varphi', \psi', \theta')$ with reference to the spatial frame with axes $e_{z'} = m/\|m\|$, $e_{x'} = e'_1$, $e_{y'} = e_{z'} \times e_{x'}$ (so that the point (e'_1, e'_2, e'_3, m) has coordinates $\varphi' = 0$, $\psi' = 0$ and $p_{\varphi'} = 0$) we see that $\Phi^{X_G}_g$ is a rotation of the body of an angle g around m. The effect of this rotation is shown in figure 3.4.c: it leaves m fixed and carries e'_3 in its final position e_3, while the axes e'_1 and e'_2 are carried into vectors e''_1 and e''_2 which are rotated of an angle $-l$ with respect to their final positions e_1 and e_2. Therefore, we arrive at the considered state by applying the flow of $X_L = \frac{\partial}{\partial \psi}$ for a time l (figure 3.4.d). ∎

We now conclude the construction by observing that the action $\Phi^{X_J} \circ \Phi^{X_G} \circ \Phi^{X_L}$ has the three independent periods $u_1 = (2\pi, 0, 0)$, $u_2 = (0, 2\pi, 0)$, and $u_3 = (0, 0, 2\pi)$, so that G, L, J, g, l, j are a set of action–angle coordinates of the considered fibration; they are illustrated in figure 3.4.d.

Written in the action–angle coordinates, the Hamiltonian H of the the symmetric Euler–Poinsot system is, on account of (3.5.1),

$$H = \frac{1}{2I_1} G^2 + \left(\frac{1}{2I_3} - \frac{1}{2I_1} \right) L^2 .$$

From Hamilton's equations one thus sees that G, L, J and j stay constant, while

$$\dot{g} = \frac{1}{I_1} G , \qquad \dot{l} = \frac{\eta}{I_1} L .$$

Remark: The construction of the action–angle coordinates based on solving the Hamilton–Jacobi equation (which requires computing some difficult integrals) can be found in [28]. An elementary direct proof of the canonical character of these coordinates, based on considerations of spherical trigonometry, is given in [29] (but this method does not prove the fact that the constructed functions are actually coordinates). Historically, these coordinates appear in a complete way in Andoyer's book [30] (but see also [31]).

C. Comments. It is worth noticing that the coordinates (G, L, J, g, l, j) have the following meaning:

- G, J and j are coordinates for the angular momentum vector in space.
- G, L and l are coordinates for the angular momentum vector in the body.
- the angles g and l are angular coordinates on, respectively, the erpholode and on the polhode of Poinsot's description.

It is not difficult to see that these coordinates are globally defined in the set M_z. However, they are singular on the border of M_z, since the two nodal lines $e_z \times m$ and $m \times e_3$ are not defined when m is zero or when it is parallel to one of the two axes e_3 or e_z. The latter two singularities have different origins:

- The singularity at $m \times e_3 = 0$ corresponds to the stationary rotations about the inertia symmetry axis e_3, which as observed in section 1.8 are singular one–dimensional leaves of the foliation by invariant tori. Correspondingly, the three–dimensional tori $(G, L, J) =$ *const* collapse there into two–dimensional tori, and no action–angle coordinates cannot be defined there.
- The singularity at $m \times e_z = 0$ is instead created by the process of grouping the two–tori so as to form three–tori, and has no intrinsic meaning (the direction e_z is indeed completely arbitrary).

Exercise: : Let (G, L, J, g, l, j) and (G', L', J', g', l', j') be two systems of action–angle coordinates relative to two spatial frames, with non–parallel axes z and z'. Show that these two coordinate system constitute an atlas for the set M, with transition functions of the form

$$\begin{aligned} G' &= G\,, & L' &= L\,, & J' &= G\,\hat{J}(J/G, j) \\ g' &= g + \hat{g}(J/G, j)\,, & l' &= l\,, & j' &= \hat{j}(J/G, j) \end{aligned} \tag{3.5.6}$$

where $\hat{J}$, $\hat{j}$ and $\hat{g}$ are smooth functions of their arguments. (*Remark:* This atlas is *not* adapted to a fibration by three–dimensional tori, since the transformation equation for the 'action' J contains also j, that is, one of the 'angles'. It is adapted, instead, to the fibration by the invariant tori of dimension two; these coordinates are an example of the 'generalized', or 'partial', action–angle coordinates of the next chapter).

Chapter 4

Degenerate Systems: Noncommutative Integrability

In this chapter we study Hamiltonian systems which have more than d integrals of motion (with suitable properties), if d is the number of degrees of freedom, and are thus 'degenerate' or 'superintegrable'. As we have seen, various classical systems of mechanics belong to this class, and their phase space is foliated by invariant tori of a certain dimension $n < d$. This foliation is the primary object to consider in order to obtain a thorough understanding of these systems. Specifically, it is important to understand the symplectic geometry of the foliation by the invariant tori.

4.1. Noncommutative integrability (local description)

A. Noncommutative integrability. We begin by stating the following Theorem 4.1, which is a generalization of the Liouville–Arnol'd theorem to degenerate situations. In its original formulation, this theorem is due to Mischenko and Fomenko [32], and is often referred to as 'noncommutative' integrability:

Theorem 4.1 *Let M be an open subset of a symplectic manifold M_* of dimension $2d$. Assume that there exists a submersion $F = (F_1, \dots, F_{2d-n}) : M \to \mathbb{R}^{2d-n}$, $n \le d$, which has compact and connected level sets and has the property that there exist functions $P_{ij} : F(M) \to \mathbb{R}$ such that*

$$\{F_i, F_j\} \; = \; P_{ij} \circ F \,, \qquad i, j = 1, \dots, 2d - n \tag{4.1.1a}$$

$$rank\,(P(F(x)) \; = \; 2d - 2n \quad \forall x \in M \,, \tag{4.1.1b}$$

P being the matrix with entries P_{ij}. Then, in M, every level set of the map F is diffeomorphic to the torus $\mathbb{T}^n$ and has a neighbourhood U endowed with a diffeomorphism

$$b \times \alpha \; : \; U \to \mathcal{B} \times \mathbb{T}^n \tag{4.1.2}$$

where $\mathcal{B} = b(U)$ is an open subset of $\mathbb{R}^{2d-n}$, with the following properties:

(i) in U, the level sets of F coincide with the sets $b = const$;

(ii) writing $b = (p_1, \dots, p_{d-n}, q_1, \dots, q_{d-n}, a_1, \dots, a_n,)$*, the restriction to* U *of the symplectic 2–form of* M *is* $\sum_{s=1}^{d-n} dp_s \wedge dq_s + \sum_{i=1}^{n} da_i \wedge d\alpha_i$.

We shall prove this theorem later, in Section 4.5.

The local coordinates a and α as in Theorem 4.1 will be called, respectively, actions and angles. Since neither the p's nor the q's need be angles, the diffeomorphisms (4.1.2) are called *generalized* (or *partial*) *action–angle coordinates* [14,33].

A Hamiltonian system which possesses $2d-n$ integrals of motion $F_1, \dots, F_{2d-n}$ as in Theorem 4.1 is said to be *integrable in the noncommutative sense* (in the set M); for shortness, we shall often say that it is *integrable.* If $n < d$, we say that the system is *degenerate.*

The dynamics of an integrable system is easily described. Since the tori $F = const$ are invariant under the flow, the local representative h of the Hamiltonian H of the system in every local system of generalized action–angle coordinates depends on the actions a alone, $h = h(a)$. Therefore, Hamilton's equations read

$$\dot{a}_i = 0\,, \qquad \dot{p}_s = 0\,, \qquad \dot{q}_s = 0\,, \qquad \dot{\alpha}_i = \omega_i(a)\,,$$

where $\omega = \partial h / \partial a$. So, motions are quasi–periodic on the tori $b = const$, with frequencies which depend on the actions alone.

Condition (4.1.1a) means that the functions $F_1, \dots, F_{2d-n}$ form a Lie algebra (of infinite dimension) with the Poisson bracket as commutator. Since the Poisson bracket of two integrals of motion is still an integral of motion, condition (4.1.1a) is automatically satisfied if $F_1, \dots, F_{2d-n}$ form a maximal set of integrals of motion, in the sense that there is no other integral of motion which is independent of them. The role and the (geometric) meaning of the rank condition (4.1.1b) will be clarified later.

Remarks: (i) In [32], Mischenko and Fomenko considered only the special case of Theorem 4.1 in which the functions $F_1, \dots, F_{2d-n}$ form a Lie algebra $\mathcal{G}$ of dimension $2d-n$, i.e., there exist constants c_{ijl} such that $\{F_i, F_j\} = \sum_l c_{ijl} F_l$. In such a case, the rank condition (4.1.1b) can be restated as

$$\dim \mathcal{G} \;+\; \text{rank}\ \mathcal{G} \;=\; \dim M \tag{4.1.3}$$

(the rank of a Lie algebra is the codimension of its adjoint orbits). The expression 'noncommutative' integrability comes obviously from the fact that, at variance from complete integrability, the algebra $\mathcal{G}$ of the integrals is not commutative.

(ii) Before Mischenko and Fomenko, a special case of Theorem 4.1 had been considered by Nekhoroshev in [14]. Specifically, Nekhoroshev considered the case of a submersion $F = (F_1, \dots, F_{2d-n})$ from a symplectic manifold M into $\mathbb{R}^{2d-n}$ with the property that n of the functions F_i are mutually in involution with all other functions, that is,

$$\{F_i, F_s\} \;=\; 0\,, \qquad i = 1, \dots, n, \quad s = 1, \dots, 2d-n\,.$$

He showed that, if the level sets of F are connected and compact, then they are tori, and local generalized action–angle coordinates (4.1.2) do exists. In Proposition 4.11, we shall see that this result is a special case of Theorem 4.1.

B. Examples. We provide now a few examples of systems which are integrable in the noncommutative sense.

Completely degenerate systems, the isotropic oscillator, and Kepler. To begin with, we consider a Hamiltonian system which, in an open subset M of its $2d$–dimensional phase space, has $2d-1$ independent integrals motions $F_1,\dots,F_{2d-1}$. We assume, furthermore, that the system does not have equilibria in the set M. Hence, no integral of motion of the system is independent of $F_1,\dots,F_{2d-1}$ and one concludes that $F_1,\dots,F_{2d-1}$ satisfy the first of the two conditions (4.1.1) entering Theorem 4.1. As it turns out, these functions satisfy also condition (4.1.1b), so that the system is integrable in the noncommutative sense with $n=1$; however, we shall prove this fact only later, in Section 4.4, once we have at our disposition some geometric instruments which will make the proof obvious. Here, we limit ourselves to verify that condition (4.1.1b) is satisfied in a few special cases.

First, we consider the 2–dimensional isotropic oscillator of Section 1.7, that is, the Hamiltonian

$$H(p_1,p_2,q_1,q_2) \;=\; \frac{1}{2}\,(p_1^2+q_1^2+p_2^2+q_2^2) \tag{4.1.4}$$

on $\mathbb{R}^4$. The system has the three integrals of motion F_1, F_2, F_3 as in (1.7.1), that is,

$$F_1(q,p) \;=\; q_1q_2+p_1p_2\,, \qquad F_2(q,p) \;=\; p_1q_2-p_2q_1\,, \qquad F_3(q,p) \;=\; \frac{1}{2}(p_1^2+q_1^2-p_2^2-q_2^2)\,. \tag{4.1.5}$$

The matrix P of the Poisson brackets of these functions is

$$P \;=\; \begin{pmatrix} 0 & 2F_3 & -2F_2 \\ -2F_3 & 0 & 2F_1 \\ 2F_2 & -2F_1 & 0 \end{pmatrix}.$$

Since P is the matrix of the vector product with $-2F$, it has rank $2=2d-2n$ (with $d=2$ and $n=1$) everywhere except where $F=0$, i.e, everywhere in $\mathbb{R}^4\setminus\{0\}$.

As a second example, we consider the Kepler system in the plane. As seen in Section 1.6, the two components in the plane of the Laplace vector $A=p\times(q\times p)-q/\|q\|$, say A_1 and A_2, and the component J of the angular momentum in the direction orthogonal to the plane are integrals of motion. Instead of the Laplace vector, it is convenient to consider the rescaled vector

$$\tilde{A} \;=\; \frac{A}{\sqrt{-2H}}\,,$$

which is still an integral of motion and is correctly defined for negative energies (to which we are interested). Taking $F_1=\tilde{A}_1$, $F_2=\tilde{A}_2$ and $F_3=J$ we obtain

$$P \;=\; \begin{pmatrix} 0 & -F_3 & F_2 \\ -F_3 & 0 & -F_1 \\ -F_2 & F_1 & 0 \end{pmatrix}$$

and as above $\mathit{rank}\,P=2$ wherever $F\neq0$, i.e. in the whole subset of the phase space where $H<0$ and $J\neq0$ to which we are interested (as one sees by observing that $\|F\|^2=-\frac{1}{2H}$).

The point in a central force field. The Hamiltonian H and the three components m_x, m_y, m_z of the angular momentum vector satisfy the two conditions (4.1.1) with $d=3$ and $n=2$. Indeed, from the well known commutation relations among the components of the angular momentum, one immediately computes

$$P \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & m_z & -m_y \\ 0 & -m_z & 0 & m_x \\ 0 & m_y & -m_x & 0 \end{pmatrix}$$

The Euler–Poinsot system. This is another system which is integrable in the noncommutative sense, with $d = 3$ and $n = 2$. Indeed, in this case, too, one can take as integrals of motion the Hamiltonian and the three components of the angular momentum in space.

C. The relation with complete integrability. When $d = n$ and $P = 0$, Theorem 4.1 reduces to the Liouville–Arnol'd theorem. In the general case, however, the exact relation between complete and noncommutative integrability is somehow subtle.

On the one hand, every noncommutatively integrable system is obviously completely integrable within the domain of every local system of generalized action–angle coordinates, since a set of d integrals of motion in involution are provided by the n actions a and, for instance, by the $d-n$ coordinates p; moreover, if the $d-n$ coordinates q (for instance) can be taken to be angles, then such a domain is fibered by Liouville tori.

On the other hand, as we have pointed out in various occasions in the previous chapters, when $n < d$ it happens frequently that the fibrations by d–dimensional tori constructed in this way is not defined globally in the whole set M fibered by the tori of dimension n, and moreover such fibrations are not unique; this means that, in a sense, noncommutative integrability is locally, but not globally, equivalent to complete integrability. Examples of this situation will be met in the examples of noncommutatively integrable systems in Section 4.4.

4.2. Bifoliations

Theorem 4.1 describes the *local* geometry of the fibration of M by the level sets of the map F. Our main purpose in this chapter is to describe the *global* structure of this fibration, which is basic for a comprehension of degenerate systems. To this end, we need to introduce the notion of *bifoliation*, or *dual pair.*[†]

A. Definitions and examples. First of all, we recall a few notions from symplectic geometry and from the theory of Poisson manifolds (see for instance [7,9]).

Let M be a symplectic manifold of dimension $2d$, with symplectic 2–form Ω. If N is a submanifold of M, then the *symplectic complement* $(T_xN)^\perp$ of the tangent space T_xN $(x \in N)$ is the subspace of T_xM constituted by all vectors which are symplectically orthogonal to T_xN, that is, $(T_xN)^\perp = \{u \in T_xM : \Omega(u,v) = 0 \ \forall v \in T_xN\}$. If N has dimension n, then the symplectic complements of its tangent spaces have dimension $2d - n$.

A submanifold of M is said to be *isotropic* (resp. *coisotropic*) if its tangent spaces are contained in (resp. contain) their own symplectic complements; isotropic submanifolds have dimension $\leq d$, and coisotropic ones have dimension $\geq d$. Lagrangian submanifolds are both isotropic and coisotropic. A foliation is called isotropic (or coisotropic, or Lagrangian) if it has isotropic (or coisotropic, or Lagrangian) leaves.

Let $\mathcal{F}$ be a foliation of M. The *polar* foliation of $\mathcal{F}$, if it exists, is the unique foliation $\mathcal{F}^\perp$ of M with the property that the tangent spaces of its leaves are the symplectic complements

[†] The present treatment follows and expand [34].

of the tangent spaces of the leaves of $\mathcal{F}$. A foliation $\mathcal{F}$ which admits a polar foliation $\mathcal{F}^\perp$ is called *symplectically complete* [33], or else a *bifoliation* [35].

If the leaves of $\mathcal{F}$ have dimension n, then the leaves of $\mathcal{F}^\perp$ have dimension $2d-n$ and $(\mathcal{F}^\perp)^\perp = \mathcal{F}$. If $\mathcal{F}$ is isotropic (coisotropic), then $\mathcal{F}^\perp$, if it exists, is coisotropic (isotropic). If $\mathcal{F}$ is isotropic and $\mathcal{F}^\perp$ exists, then the leaves of $\mathcal{F}^\perp$ are coisotropic and properly contain every leaf of $\mathcal{F}$ that they intersect.

Here are a few examples of bifoliations:

- A trivial example is offered by any Lagrangian foliation, which coincides with its own polar.
- The orbits of a Hamiltonian vector field X_H are the leaves of a foliation which has a polar foliation: the leaves of the polar are the (connected components of the) level sets of the Hamiltonian H; indeed, if Y is any vector field tangent to the level set $H = const$ one has $\Omega(Y, X_H) = L_Y H = 0$.
- Every coisotropic foliation has a polar foliation: the fact that the distribution of the symplectic complements is involutive is verified by using the closure of Ω and Cartan's formula (or else, property B1 below).
- The orbits of a Hamiltonian action and the level sets of its momentum map are polar to each other (see chapter 5).
- Not every foliation with isotropic leaves has a polar foliation. Examples of isotropic foliations with no polar are constructed by considering any non–integrable distribution with leaves of codimension one: the symplectically orthogonal distribution is thus integrable and isotropic, being one–dimensional, but obviously does not admit a polar.

In the special case in which both $\mathcal{F}$ and $\mathcal{F}^\perp$ are described by two surjective submersions $\pi_1 : M \to B_1$ and $\pi_2 : M \to B_2$, then the pair $(\mathcal{F}, \mathcal{F}^\perp)$, also denoted $B_1 \overset{\pi_2}{\leftarrow} M \overset{\pi_1}{\rightarrow} B_2$, is called a *dual pair* [36]. If both submersions $\pi_i : M \to B_i$ are fibrations, then in [35] the name *bifibration* is used for both of them; in the sequel, we shall use this name for the pair $B_1 \overset{\pi_2}{\leftarrow} M \overset{\pi_1}{\rightarrow} B_2$.

Bifoliations and dual pairs, under slightly different hypotheses and different names, have been extensively studied in symplectic geometry (see [36, 33, 7, 35] and references therein). In the next subsection, we recall a few simple technical properties of bifoliations (see [7], chapter 3, section 9).

B. Properties of bifoliations. In the following, we shall (tacitly) use the following notation: if $\mathcal{F}$ is a foliation of a symplectic manifold, we denote by F the completely integrable distribution associated to $\mathcal{F}$ and by $F^\perp$ the distribution which, at every point $x \in M$, consists of the symplectic complement of $F(x)$. Thus, $\mathcal{F}$ has a polar foliation if and only if $F^\perp$ is completely integrable.

A *first integral* of a foliation is a function (possibly defined only in an open subset) which is constant on the leaves of the foliation. If the foliation is defined by a submersion, then its first integrals are the lifts of functions defined on the base. The following is a simple observation which should be kept in mind:

Lemma 4.2 *A function f is a first integral of a foliation $\mathcal{F}$ if and only if its Hamiltonian vector field $X_f \in F^\perp$.*

Proof. The statement follows from the fact that a function f is a first integral of $\mathcal{F}$ if and only if $L_Y f = 0$ for every vector field $Y \in F$, and from the identity $L_Y f = \Omega(Y, X_f)$. ■

We can now state the following

Proposition 4.3 *Let $\mathcal{F}$ be a foliation of a symplectic manifold (M, Ω). Then:*

(B1) $\mathcal{F}$ possesses a polar foliation if and only if the Poisson brackets of every two first integrals of $\mathcal{F}$ is a first integral of $\mathcal{F}$.

(B2) Assume that $\mathcal{F}^\perp$ exists. Then:

(i) the leaves of $\mathcal{F}$ are generated by the flows of the Hamiltonian vector fields of the first integrals of $\mathcal{F}^\perp$.

(ii) A function is a first integral of $\mathcal{F}$ if and only if it is in involution with every first integral of $\mathcal{F}^\perp$.

Proof. First we prove property B1. If the polar foliation of $\mathcal{F}$ exists, then $F^\perp$ is completely integrable. Let f and g be first integrals of $\mathcal{F}$. Hence, since X_f and X_g both belong to $F^\perp$, by Frobenius theorem one has $[X_f, X_g] \in F^\perp$. But then $X_{\{f,g\}} = [X_f, X_g]$ also belongs to $F^\perp$, and so $\{f, g\}$ is a first integral of $\mathcal{F}$.

In order to prove the converse, we observe that *the distribution $F^\perp$ has local bases made of Hamiltonian vector fields of first integrals of $\mathcal{F}$.* Indeed, any point $x \in M$ has a neighbourhood U in which are defined $c = codim\,(\mathcal{F})$ functions $z^1, \ldots, z^c$ which have differentials everywhere linearly independent and are first integrals of $\mathcal{F}$. (Just take local coordinates transversal to the leaves of $\mathcal{F}$). By the Lemma, the Hamiltonian vector fields $X_{z^1}, \ldots, X_{z^c}$ are tangent to $F^\perp$; being linearly independent, and in the right number, they form a local base. Now, by the hypothesis, $\{z^i, z^j\}$ is a first integral of $\mathcal{F}$, for all i, j. Hence, $X_{\{z^i,z^j\}} \in F^\perp$. Since $X_{\{z^i,z^j\}} = [X_{z^i}, X_{z^j}]$, this shows that $F^\perp$ is completely integrable so that $\mathcal{F}^\perp$ exists.

Statement B2.i follows from the fact just observed, that the tangent spaces to $\mathcal{F}^\perp$ have local bases made of Hamiltonian vector fields of first integrals of $\mathcal{F}$, and from the uniqueness of the leaves of a foliation.

Finally, B2.ii follows from the fact that, on account of B2.i, f is a first integral of $\mathcal{F}$ if and only if $0 = L_{X_g} f = \{g, f\}$ for every first integral g of $\mathcal{F}^\perp$. ■

C. Bifoliations and Poisson manifolds. A *Poisson manifold* is a manifold P equipped with Poisson brackets $\{\ ,\ \}_P$ for functions. By means of the Poisson brackets one can associate to every function $f : P \to \mathbb{R}$ a *Hamiltonian vector field* X_f, which is defined by the identity $L_{X_f} g = \{f, g\}_P$ for every function $g : P \to \mathbb{R}$.

The set of values of all Hamiltonian vector fields form a subspace S_x of every tangent space $T_x P$, $x \in P$. It is a basic fact that the collection of all the subspaces S_x, $x \in P$, is completely integrable, so that it defines a foliation of P, and moreover, that the leaves of this foliation carry a symplectic structure; for this reason, they are called the *symplectic leaves* of P.

The dimension of the symplectic leaf through a point $x \in P$ is called the *rank* of P at the point x; if $(y^1, \ldots, y^m)$ are local coordinates around x ($m = dim\,P$), the rank of P at x equals the rank of the matrix $\{y^i, y^j\}_P(x)$.

The (local) *Casimirs* of a Poisson manifold are the functions (defined in open sets) which are constant on the symplectic leaves, or equivalently, they are the functions which

Poisson–commute with every other function on P.

The simplest example of a Poisson manifold is a connected symplectic manifold; in this case, there is only one symplectic leaf, which coincides with the whole manifold. In the next Section, we will see that Poisson manifolds emerge naturally with integrable systems, since the base manifold of the fibration by the invariant tori has such a structure.

Proposition 4.4 *Consider a surjective submersion $\pi_1 : M \to B$ from a symplectic manifold (M, Ω) onto a manifold B, and assume that its fibers are connected. Denote by $\mathcal{F}$ the foliation of M whose leaves are the fibers of π_1.*

(B3) $\mathcal{F}$ has a polar foliation $\mathcal{F}^\perp$ if and only if there exists a Poisson structure on B with respect to which π_1 is a Poisson morphism (i.e., $\{f,g\}_B \circ \pi_1 = \{f\circ\pi_1, g\circ\pi_1\}_M$ for all functions f, g defined in open sets of B, if $\{\ ,\ \}_B$ denotes the Poisson bracket of this structure and $\{\ ,\ \}_M$ denotes those of (M, Ω)); such a Poisson structure is unique.

(B4) Assume that $\mathcal{F}^\perp$ exists. Then, the following four conditions are equivalent:

- *(i) The leaves of $\mathcal{F}$ are isotropic.*
- *(ii) The rank of the induced Poisson structure of B is everywhere equal to $2\dim B - \dim M$ (i.e., the number of independent local Casimirs of B equals everywhere the dimension of the fibers of π_1).*
- *(iii) The leaves of $\mathcal{F}$ are generated by the Hamiltonian vector fields of the lifts to M of the Casimirs of B.*
- *(iv) The first integrals of $\mathcal{F}^\perp$ are exactly the lifts to M of the Casimirs of B.*

Proof. First we prove B3. Assume that $\mathcal{F}^\perp$ exists. If U is an open subset of B and $f, g \in \mathcal{C}^\infty(U)$, then $f\circ\pi_1$ and $g\circ\pi_1$ are first integrals of $\mathcal{F}$ in $\pi_1^{-1}(U)$ and such is, by property B1, $\{f\circ\pi_1, g\circ\pi_1\}_M$. Therefore, since the level sets of of π_1 are connected, there exists a function $\mathcal{P}_{f,g}$ defined in U such that $\mathcal{P}_{f,g}\circ\pi_1 = \{f\circ\pi_1, g\circ\pi_1\}_M$. We thus define $\{f,g\}_B = \mathcal{P}_{f,g}$.

Conversely, assume that there exists a Poisson bracket $\{\ ,\ \}_B$ such that $\{f,g\}_B\circ\pi_1 = \{f\circ\pi_1, g\circ\pi_1\}_M$. Since every first integral of $\mathcal{F}$ is of the form $f\circ\pi_1$ for some function f defined on an open subset of B, this implies that the Poisson bracket of every two first integrals of $\mathcal{F}$ is a first integral of $\mathcal{F}$, which is therefore a bifoliation.

The uniqueness of the Poisson structure on B which makes π_1 a Poisson morphism is obvious.

We prove now B4. First we prove that statements (i) and (ii) are equivalent. On account of B3 and of the fact that the first integrals of $\mathcal{F}$ are the lifts of functions on B, the number n_x of independent Casimirs of B at a point $\pi_1(x)$ equals the number of independent germs of functions which are simultaneously first integrals of $\mathcal{F}$ and $\mathcal{F}^\perp$. Hence, since the Hamiltonian vector fields of these functions are simultaneously tangent to the leaves of $\mathcal{F}$ and to those of $\mathcal{F}^\perp$, one has $dim\,(F_x \cap F_x^\perp) = n_x$. Hence,

$$dim\,(F_x) \;\geq\; dim\,(F_x \cap F_x^\perp) \;=\; n_x$$

showing that $n_x = dim\,F_x$ (i.e., $rank\{\ ,\ \}_B = 2dim\,B - dim\,M$) if and only if $dim\,(F_x) = dim\,(F_x \cap F_x^\perp)$, i.e., if and only if $F_x = F_x \cap F_x^\perp$ (given that $F_x \cap F_x^\perp$ is a subset of F_x), i.e., if and only if F_x is isotropic.

Conditions (iii) and (iv) are just restatements of (ii). ∎

We conclude this study of bifoliations with the following

Proposition 4.5 *Consider a surjective submersion $\pi_1 : M \to B$ from a symplectic manifold (M, Ω) onto a manifold B whose fibers are connected and isotropic. Assume that the foliation $\mathcal{F}$ whose leaves are the fibers of π_1 has a polar foliation $\mathcal{F}^\perp$, whose leaves are the connected components of the fibers of a surjective submersion $\pi_2 : M \to A$, where A is a manifold. Then, the symplectic leaves of B are the connected components of the fibers of the unique map $\pi_3 : B \to A$ such that $\pi_3 \circ \pi_1 = \pi_2$, which is a surjective submersion (see figure 4.1.a).*

Proof. The existence of the map $\pi_3 : B \to A$ follows from the fact that the fibers of π_2, being coisotropic, are union of fibers of π_1. The map π_3 is a surjective submersion since both $\pi_1 : M \to B$ and $\pi_2 : M \to A$ are surjective submersions. The fact that the level sets of π_3 are the symplectic leaves of B follows from B4.iv. ∎

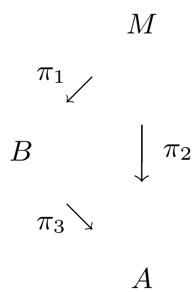


Figure 4.1.a

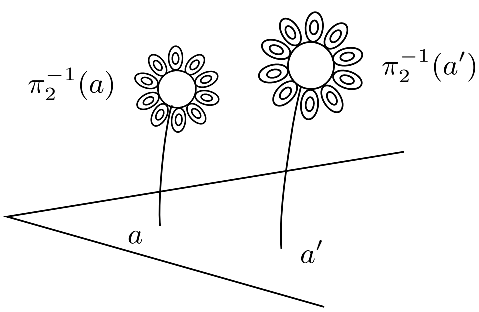


Figure 4.1.b

4.3. Noncommutative integrability—geometric aspects

A. The geometric meaning of Theorem 4.1. The following Theorem 4.6, which in various forms and at various levels of generality appears in [14, 32, 33, 35], is a basic result about the structure of an isotropic bifoliation:

Theorem 4.6 *Let M be a $2d$–dimensional symplectic manifold and let $\pi_1 : M \to B$ be a fibration with compact connected isotropic fibers of dimension $n \le d$. Assume that this fibration has a polar foliation $\mathcal{F}^\perp$. Then:*

(i) The fibers of π_1 are diffeomorphic to $\mathbb{T}^n$.

(ii) Every fiber of π_1 has a neighbourhood U endowed with generalized action–angle coordinates, that is, a diffeomorphism $(b \times \alpha) : U \to \mathcal{B} \times \mathbb{T}^n$ which satisfies properties (i)–(iii) of Theorem 4.1 (with F replaced by π_1, of course).

(iii) If, moreover, the leaves of $\mathcal{F}^\perp$ are the fibers of a fibration $\pi_2 : M \to A$, then A has an integer affine structure.

Statement (i) follows from the fact that, on account of property B4.iii, the leaves of $\mathcal{F}$ possess n linearly independent commuting tangent vector fields, i.e., the Hamiltonian vector fields of the lifts of the Casimirs of B; this implies the statement, on the basis of a well known fact used in the proof of the Liouville–Arnol'd theorem. The proof of statements (ii) and (iii) of Theorem 4.6 will be given later, in Section 4.5. We will explain later, in subsection C below, what is the integer affine structure of the base A, which will be hereafter called the *action space* of the bifibration $A \overset{\pi_2}{\leftarrow} M \overset{\pi_1}{\rightarrow} B$.

As we now show, Theorem 4.1 is nothing else than a local version of Theorem 4.6. This is made precise in the following Proposition, whose proof is interesting because clarifies the geometric meaning of the hypotheses entering Theorem 4.1:

Proposition 4.7 *(i) The fibers of a submersion $F : M \to F(M)$ as in Theorem 4.1 are isotropic and have a polar foliation.*

(ii) Every fibration which satisfies the hypotheses of Theorem 4.6 can be described locally (i.e., in a neighbourhood of every fiber) by $2d - n$ functions as in Theorem 4.1.

Proof. On account of property B1, condition (4.1.1a) precisely means that the fibration by the level sets of F is a bifoliation. Consequently, by B4, $F(M)$ with the induced brackets is a Poisson manifold. Moreover, on account of property B5, the isotropy of the fibers of F amounts to the fact that the rank of the Poisson structure of $F(M)$ equals everywhere $2 dim\, B - dim\, M = 2d-2n$. This is in fact assured by the hypothesis on the rank of the matrix P entering (4.1.1a), as one verifies by observing that, in the coordinates y^i on $F(M)$ defined by $y_i(F(x)) = F_i(x)$ for every $x \in M$, one has $\{y^i, y^j\}_{f(M)}(f(x)) = \{F_{,}F_j\}_M(x) = P_{ij}(f(x))$.

We now prove statement (ii). To this end, we only need to prove that, if $\pi_1 : M \to B$ is a fibration which satisfies the hypotheses of Theorem 4.6, then every point $y \in B$ has a neighbourhood $U \subset B$ such that the restriction of π_1 to $\pi_1^{-1}(U)$ coincides (=has the same fibers) with a submersion $F : \pi_1^{-1}(U) \to \mathbb{R}^{2d-n}$ which satisfies the hypotheses of Theorem 4.1. To this purpose, just consider a local system of coordinates $\hat{f} : U \to \mathbb{R}^{2d-n}$. Then, the lift $F = \hat{f} \circ \pi_1$ does the job, since it is a submersion whose fibers coincide with the fibers of π_1 and, moreover, by properties B1 and B5, the Poisson brackets of its components satisfy the hypotheses of Theorem 4.1. ∎

Proposition 4.7 implies that Theorems 4.1 and 4.6 are equivalent. Indeed, since the statements of Theorem 4.6 are local on the base of the fibration, the local statements of Theorem 4.1 are sufficient to prove Theorem 4.6. However, the global hypotheses of Theorem 4.6 have the advantage that they allow a global characterization of noncommutative integrability.

Remarks: (i) This geometric point of view makes very easy to prove that, under the hypotheses of Section 4.1.B, the $2d - 1$ integrals of motion of a completely degenerate Hamiltonian system satisfy the rank condition (4.1.1b). Indeed, as observed in the proof of Proposition 4.7, such a condition is equivalent to the isotropy of the level sets of the map F, which is obviously verified since they are one–dimensional.

(ii) Condition (4.1.1b) entering Theorem 4.1 can be restated as

$$\dim B \;+\; \mathrm{corank} B \;=\; \dim M$$

(where corank$B = \dim B - \operatorname{rank} B$; compare with (4.1.3)).

B. Noncommutative integrability: geometric definition. Adopting a geometric point of view, we shall say that a Hamiltonian system on a manifold M_* is *integrable in the noncommutative sense in an open subset M of M_** if there is an isotropic and symplectically complete fibration of M whose fibers are compact, connected, and invariant under the flow of the given system.

As one sees by referring to local generalized action–angle coordinates, the flow of a system with these properties is linear on every torus, and all the tori based over the same symplectic leaf support motions with equal frequencies.† These tori are generated by the Hamiltonian vector fields of the (lifts of the) Casimirs of the base manifold.

Figure 4.1.b is an attempt to symbolically represent the bifoliation structure of the phase space of a noncommutatively integrable system. As a visual help for imagination, we suggest of thinking of each fiber $\pi_2^{-1}(a)$ as being a flower, whose center is the symplectic leaf $\pi_3^{-1}(a)$ and whose petals are the tori $\pi_1^{-1}(b)$. Correspondingly, the action space is the meadow on which they grow. The flow of an integrable system is quasi–periodic on the petals, and all petals in the same flower have equal frequencies. (This analogy is not perfect, since the sets $\pi_2^{-1}(a)$ need not have the simple product structure '*center* $\times$ *petal*' of true flowers: each fibration $\pi_1 : \pi_2^{-1}(a) \to \pi_3^{-1}(a)$ may instead be topologically non–trivial).

In any local system of generalized action–angle coordinates of a noncommutatively integrable system, the angles α are coordinates on the invariant tori, so that the local coordinates (a, p, q) can be regarded as local coordinates on the base manifold B.† More specifically, a are local Casimirs of B, so that (p, q) are local coordinates on the symplectic leaves of B (the centers of the flowers of Figure 4.1.b). In cases in which the foliation polar to the foliation by the invariant tori is a fibration $\pi_2 : M \to A$, then the actions a can be regarded as local coordinates on the action space A.

These considerations should clarify that what one does when regarding a noncommutatively integrable systems as completely integrable is to introduce pairs of action–angle coordinates on the symplectic leaves. Depending on the geometry of the symplectic leaves, this cannot be done globally, nor uniquely.

C. Non–uniqueness of the generalized action–angle coordinates. The group of allowed transformations of the generalized action–angle coordinates of a bifibration is larger than that of the action–angle coordinates of a Lagrangian fibration:

Proposition 4.8 ([37,14]) *Any two different sets (a, p, q, α) and (a', p', q', α') of generalized action–angle coordinates of $A \overset{\pi_2}{\leftarrow} M \overset{\pi_1}{\rightarrow} B$ are related, in each connected component of the intersection*

† This statement has an intrinsic character even if the fiber is not covered by a single system of generalized action–angle coordinates, as is seen by observing that the transition functions between different local systems of these coordinates have the form (4.3.1). The geometric reason behind this fact is that the fibration of $\pi_2^{-1}(a)$ by the tori $\pi_1^{-1}(b)$ is (locally) generated by an action of $\mathbb{T}^n$.

† Here there is a minor abuse, since (a, p, q) are defined on M, not B. The correct statement is that there exist functions $(\tilde{a}, \tilde{p}, \tilde{q})$ on B such that $(a, p, q) = (\tilde{a}, \tilde{p}, \tilde{q}) \circ \pi_1$.

of their domains, by equations of the form

$$\begin{aligned} a' &= Za + z \\ (p',q') &= \mathcal{D}(a,p,q) \\ \alpha' &= Z^{-T}\alpha + \mathcal{F}(a,p,q) \quad (mod 2\pi) \end{aligned} \tag{4.3.1}$$

with some matrix $Z \in SL_\pm(\mathbb{Z}, n)$*, some vector* $z \in \mathbb{R}^n$*, and some maps* $\mathcal{D}$ *and* $\mathcal{F}$*.*

Proof. Since the coordinates (a,p,q) and (a',p',q') are transversal to the fibers of π_1, in the intersection $U \cap U'$ of their domains, the transition functions between the two sets of coordinates are of the form

$$a' = a'(a,p,q), \qquad (p',q') = \mathcal{D}(a,p,q), \qquad \alpha' = \alpha'(a,p,q,\alpha). \tag{4.3.2}$$

On the other hand, both coordinate systems are symplectic, so that

$$\sum_i da'_i \wedge d\alpha'_i + \sum_s dp'_s \wedge dq'_s = \sum_i da_i \wedge d\alpha_i + \sum_s dp_s \wedge dq_s .$$

The only terms coming from the l.h.s. which contain the wedge product of some da_i's and of some $d\alpha_i$'s are

$$\sum_{ijh} \frac{\partial a'_i}{\partial a_j}(a) \, \frac{\partial \alpha'_i}{\partial \alpha_h}(a,p,q,\alpha) \, da_j \wedge d\alpha_h . \tag{4.3.3}$$

Equating this expression to $\sum_i da_i \wedge d\alpha_i$ one gets

$$\frac{\partial \alpha'}{\partial \alpha}(a,p,q,\alpha) = \left[\frac{\partial a'}{\partial a}(a) \right]^{-T}, \tag{4.3.4}$$

which implies that $\frac{\partial \alpha'}{\partial \alpha}$ is independent of α. Thus, there exist a matrix $C(a,p,q)$ and a vector $\mathcal{F}(p,q,a)$ which depend smoothly on (p,q,a) and are such that

$$\alpha' = C(a)\,\alpha + \mathcal{F}(a,p,q) \quad (mod 2\pi).$$

Since α and α' are both coordinates on the torus $(a,p,q) = const$, the matrix $C(a)$ belongs to $SL_\pm(n, \mathbb{Z})$ and so it is constant in connected sets; in (4.3.2) we have written $Z = C^{-T}$. Therefore, (4.3.4) implies $a' = Za + z(p,q)$. But here, the function z must be a constant. This can be seen in two ways. The first way refers to equation (4.3.3): equating to zero all terms at the l.h.s. which contain $dp'_i \wedge d\alpha'_j$ or $dq'_i \wedge d\alpha'_j$, and using the fact that the matrix C is nonsingular, one gets $\partial a'/\partial p = \partial a'/\partial q = 0$. The other way consists in observing that the polar foliation of π_1 is locally described by the equations $a_\lambda = const$, so that $a_\mu = a_\mu(a_\lambda)$. ∎

Remark: One can completely characterize the global structure of the bifoliation of M in terms of the transition functions (4.3.2) — an approach used by Nekhoroshev in [14].

Exercise: Let (a,p,q,α) be generalized action–angle coordinates of a given bifibration. Establish under which conditions on the maps $\mathcal{D}$ and $\mathcal{F}$ the coordinates (a',p',q',α') as in (4.3.2) are generalized action–angle coordinates, too.

4.4 Examples

We illustrate now the bifoliation structure of the phase space of the noncommutatively integrable systems considered in Section 4.1.

A. Completely degenerate systems. Consider a system whose all orbits, in some subset M of the phase space, are periodic. Being one–dimensional, the periodic orbits define an isotropic foliation of M. We have already observed that this foliation does have a polar foliation, whose leaves are the connected components of the Hamiltonian H. Indeed, the vector field X_H is both tangent to the orbits of the system and symplectically orthogonal to the level sets of H.

If the periodic orbits are the fibers of a fibration, we have a noncommutatively integrable system in the strong (geometric) sense of Section 4.3. In particular, one can cover the whole phase space with an atlas with charts with generalized action–angle coordinates.

B. The isotropic oscillator. We study now in some details the two–dimensional isotropic oscillator of Sections 1.7 and 4.1. We restrict ourselves since the very beginning to the subset M of the phase space $\mathbb{R}^4$ fibered by the periodic orbits, that is, $\mathbb{R}^4 \setminus \{0\}$.

The bifibration. First of all, we show that the fibration by the periodic orbits is described by the three integrals of motion F_1, F_2 and F_3 as in (4.1.5).

Proposition 4.9 *The map*

$$F = (F_1, F_2, F_3) \; : \; \mathbb{R}^4 \setminus \{0\} \; \to \mathbb{R}^3 \setminus \{0\}$$

is a surjective submersion with level sets diffeomorphic to S^1.

Proof. It is immediate to verify that the Jacobian $\partial(F_1, F_2, F_3)/\partial(p_1, p_2, q_1, q_2)$ has four 3×3 minors equal to, respectively, $2q_1H$, $2p_1H$, $-2q_2H$, $-2p_2H$; hence, it has rank three at every point in $\mathbb{R}^4$ but the origin, and so F is a submersion.

In order to show that F is onto $\mathbb{R}^3 \setminus \{0\}$, we recall from (1.7.2) that one has

$$F_1^2 \; + \; F_2^2 \; + \; F_3^2 \; = \; H^2 \, . \tag{4.4.1}$$

This implies that F maps each level set of H, that is, the three–dimensional sphere $\{(p, q) \in \mathbb{R}^4 : \|p^2\| + \|q\|^2 = 2H\}$, into the two–dimensional sphere $\{F \in \mathbb{R}^3 : \|F\| = H^2\}$. Therefore, it suffices here to show that the restriction of F to the spheres $\{(p, q) \in \mathbb{R}^4 : \|p^2\| + \|q\|^2 = 2H\}$ is surjective. We show this by using an equivariance property of F.

Let us denote by $\mathcal{R}^{(2)}_\varphi$ and $\mathcal{R}^{(3)}_\varphi$ the matrices of a rotation of an angle φ around the axes F_2 and F_3, respectively, i.e.

$$\mathcal{R}^{(2)}_\varphi \; = \; \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} , \qquad \mathcal{R}^{(3)}_\varphi \; = \; \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

A direct computation shows that, for any φ, one has

$$\mathcal{R}^{(2)}_\varphi \circ F \; = \; F \circ \widehat{\mathcal{R}}^{(2)}_{\varphi/2} \, , \qquad \mathcal{R}^{(3)}_\varphi \circ F \; = \; F \circ \widehat{\mathcal{R}}^{(3)}_\varphi$$

where

$$\widehat{\mathcal{R}}^{(2)}_{\varphi} \;=\; \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & \cos\varphi & -\sin\varphi \\ 0 & 0 & \sin\varphi & \cos\varphi \end{pmatrix}, \qquad \widehat{\mathcal{R}}^{(3)}_{\varphi} \;=\; \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi & 0 \\ 0 & 1 & 0 & 0 \\ \cos\varphi & 0 & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

both belong to $SO(4)$. Hence, for any φ_2 and φ_3, one has

$$\mathcal{R}^{(3)}_{\varphi_3} \circ \mathcal{R}^{(2)}_{2\varphi_2} \circ F \;=\; F \circ \widehat{\mathcal{R}}^{(3)}_{\varphi_3} \circ \widehat{\mathcal{R}}^{(2)}_{\varphi_2}\,.$$

This proves the surjectivity of $F : S^3 \to S^2$, since any point $F \in S^2$ can be reached from any other point (for instance, from $(H,0,0) = F(\sqrt{H/2}, \sqrt{H/2}, \sqrt{H/2}, \sqrt{H/2}))$ with two rotations around the axes F_2 and F_3.

In order to show that the level sets of F are circles, we identify $\mathbb{R}^4$ with $\mathbb{C}^2$, letting

$$z_j \;=\; \frac{p_j + i q_j}{\sqrt{2}}\,, \qquad j = 1, 2.$$

Correspondingly

$$F_1(z) \;=\; \overline{z}_1 z_2 + \overline{z}_2 z_1\,, \qquad F_2(z) \;=\; i\,(\overline{z}_1 z_2 - \overline{z}_2 z_1)\,, \qquad F_3(z) \;=\; \overline{z}_1 z_1 - \overline{z}_2 z_2\,. \tag{4.4.2}$$

If z and z' belong to the same level set of F, one easily sees (using also $H = const$) that

$$|z_1'| \;=\; |z_1|\,, \qquad |z_2'| \;=\; |z_2|\,, \qquad \overline{z}_1' z_2 \;=\; \overline{z}_1' z_2'\,.$$

The first two equations imply $z_1' = e^{i\alpha_1} z_1$, $z_2' = e^{i\alpha_2} z_2$ for some $\alpha_1, \alpha_2 \in S^1$. The last equation then implies $\alpha_1 = \alpha_2$, showing that the level set of F is a circle. ∎

We can now describe the bifibration structure of the phase space of the two–dimensional isotropic oscillator. The level sets of the map F (being invariant and connected sets) obviously coincide with the orbits, so the isotropic fibration is given by F. The polar foliation is given by the energy, $H : \mathbb{R}^4 \setminus \{0\} \to \mathbb{R}_+$, and is a fibration with fiber S^3. So, we have the bifibration

$$A \overset{H}{\leftarrow} M \overset{F}{\rightarrow} B \tag{4.4.3}$$

with

$$A \;=\; \mathbb{R}_+\,, \qquad M \;=\; \mathbb{R}^4 \setminus \{0\} \approx S^3 \times \mathbb{R}_+\,, \qquad B \;=\; \mathbb{R}^3 \setminus \{0\} \approx S^2 \times \mathbb{R}_+ \tag{4.4.4}$$

The symplectic leaves of the base B are obtained by projecting with F the level sets of H, so they are the two–dimensional spheres $\{F \in \mathbb{R}^3 : \|F\| = H\}$; as seen in Section 1.7, the restriction of F to the level set of H is the Hopf fibration $S^3 \to S^2$.

The generalized action–angle coordinates. We construct now a set of (local) generalized action–angle coordinates of the bifibration (4.4.3), follwing [38]. We start from a system of action–angle coordinates of the fibration by the Lagrangian two–dimensional tori $a_1 = const$, $a_2 = const$, where

$$a_i \;=\; \frac{1}{2}\,(p_i^2 + q_i^2)$$

As seen in Section 1.4, as actions of this fibration one can take precisely a_1 and a_2, so we denote $(a_1, a_2, \alpha_1, \alpha_2)$ these coordinates. Obviously, these coordinates are defined in the subset of $\mathbb{R}^4 \setminus \{0\}$ where $a_1 \neq 0$ and $a_2 \neq 0$, that is, everywhere but on those special motions in which the oscillation takes place in the direction of one of the coordinate axes q_1 and q_2.

Since $H = a_1 + a_2$, $(a_1, a_2, \alpha_1, \alpha_2)$ are *not* generalized action–angle coordinates of the system. Partial action–angle coordinates (a, α, p, q) are however easily constructed out of them, by taking, for instance,

$$a = a_1 + a_2 , \quad \alpha = \alpha_1 , \quad p = -a_2 , \quad q = \alpha_1 - \alpha_2 ,$$

so that $H = a$. In the coordinates (a, α, p, q) the first integrals F_1, F_2, F_3 become

$$F_1 = 2\sqrt{-p(a+p)}\, cos\, q , \qquad F_2 = 2\sqrt{-p(a+p)}\, sin\, q , \qquad F_3 = a + 2p , \tag{4.4.5}$$

so that (F_3, q) are cylindrical coordinates for the vector $F \in \mathbb{R}^3 \setminus \{0\}$. The coordinates (p, q) on every symplectic leaf $\|F\| = a$ are obtained by replacing F_3 with $p = (F_3 - a)/2$, and so they are still cylindric–like coordinates, with singularities on the axis F_3 (see Figure 4.2.a).

The presence of some singularities of the generalized action–angle coordinates is unavoidable, since, as seen in Section 1.7, the fibration F is nontrivial (involving a Hopf fibration $S^3 \to S^2$); hence, there cannot be a global system of generalized action–angle coordinates adapted to it; and moreover, the symplectic leaves, being compact, cannot obviously be covered with a single system of coordinates. However, one can construct an atlas for the set M made of two charts with generalized action–angle coordinates.

To this purpose, we begin by considering another set of action–angle coordinates $(a_1', a_2', \alpha_1', \alpha_2')$, adapted to a *different* Lagrangian fibration and with singularities in a *different* location. That this is possible should be quite clear: the underlying idea is that the location of the singularity of the coordinates (a, α, p, q) has nothing of intrinsic, since oscillations along the axes q_1 and q_2 do not have anything which distinguishes them from all other orbits—the oscillator is isotropic, so there are no preferred direction in which it moves. Thus, we make a rotation of an angle ψ in the plane (q_1, q_2), thus obtaining new coordinates

$$q_1' = q_1\, cos\, \psi - q_2\, sin\, \psi , \qquad q_2' = q_1\, sin\, \psi + q_2\, cos\, \psi , \tag{4.4.6}$$

and we then repeat the above construction of the generalized action–angle coordinates, the singularity should now appear on those motions in which the oscillator moves at an angle ψ with the axes q_1 and q_2.

Specifically, we extend the change of coordinates (4.4.6) to the momenta so as to obtain new canonical coordinates (p', q') in $\mathbb{R}^4$, that is,

$$p_1' = p_1\, cos\, \psi - p_2\, sin\, \psi , \qquad p_2' = p_1\, sin\, \psi + p_2\, cos\, \psi .$$

In the new coordinates, the Hamiltonian retains its form, i.e.

$$H = \frac{1}{2}\left({p_1'}^2 + {q_1'}^2 + {p_2'}^2 + {q_2'}^2\right)$$

and we introduce action–angle coordinates $(a_1', a_2', \alpha_1', \alpha_2')$ for the Lagrangian fibration

$$a_1' = \frac{1}{2}\left({p_1'}^2 + {q'^2}_1\right) = const , \qquad a_2' = \frac{1}{2}\left({p_2'}^2 + {q'^2}_2\right) = const ,$$

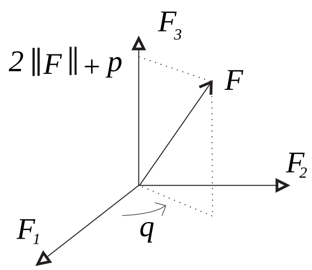


Figure 4.2.a

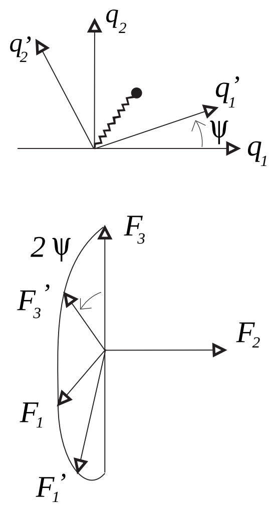


Figure 4.2.b

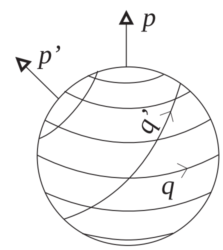


Figure 4.2.c

and out of them we construct the generalized action–angle coordinates

$$a' \;=\; a_1' + a_2' \,, \quad \alpha' \;=\; \alpha_1' \,, \quad p' \;=\; -a_2' \,, \quad q' \;=\; \alpha_1' - \alpha_2' \,,$$

Clearly $a' = a$, while (p', q') are coordinates on the symplectic leaves of the base manifold. Specifically, (p', q') are cylindric–like coordinates for the vector $F' = (F_1', F_2', F_3')$, where

$$F_1'(q', p') \;=\; q_1'q_2' + p_1'p_2' \,, \qquad F_2'(q', p') \;=\; p_1'q_2' - p_2'q_1' \,, \qquad F_3'(q', p') \;=\; \frac{1}{2}({p_1'}^2 + {q_1'}^2 - {p_2'}^2 - {q_2'}^2) \,.$$

As seen in the course of the proof of Proposition 4.9, F' is obtained by rotating F of an angle 2ψ around the axis F_2 (Figure 4.2.b). Hence, the coordinates (p', q') are cylindric–like coordinates relative to a rotated basis in the base manifold $B = \mathbb{R}^3 \setminus \{0\}$: on every symplectic leaf, the curves $p = const$ and $\varphi' = const$ are circles forming an angle 2ψ, as illustrated in Figure 4.2.c.

C. The rigid body. As a second example, we consider the symmetric Euler–Poinsot system. As discussed in Sections 1.8 and 3.5, the fibration by the invariant tori is given by the map

$$\pi_1 \;=\; L \times m^s \,:\, M_* \;\to\; B_* \,,$$

where L is the component of the angular momentum along the symmetry axis of inertia and

$$B \;=\; \big\{(L, m^s) \in \mathbb{R} \times \mathbb{R}^3 \,:\, |L| < \|m^s\|\big\} \,.$$

From the Poisson brackets of the functions L, m_x, m_y and m_z one sees that $G = \|m^s\|$ and L are two independent Casimirs of B. Hence, we can take

$$\pi_2 \;=\; L \times G \,:\, M \;\to\; A$$

with

$$A \;=\; \big\{(L,G)\in \mathbb{R}^2 :\, |L|<G\big\}\,.$$

The symplectic leaf over a point $(\overline{L},\overline{G})\in A$ is the two–dimensional sphere

$$\big\{(L,m^s)\in B:\, L=\overline{L},\ \|m^s\|=\overline{G}\big\}\,.$$

For fixed $\overline{L}$, this set can be identified with the two–dimensional sphere of constant modulus $\overline{G}$ of the angular momentum vector in space.

As local generalized action–angle coordinates, we can use a system of action–angle coordinates (G,L,J,g,l,j) as constructed in Section 3.5; G and L are the actions, g and l are the angles, and J and j are (local) coordinates on the symplectic leaves (indeed, as already observed, G, J and j are cylindric–like coordinates for the angular momentum vector in space). These coordinates are only locally defined in M but, as already observed in Section 3.5, one obtains a two–charts atlas for M just by considering the action–angle coordinates (G',L',J',g',l',j') relative to a different spatial frame $\{e_{x'},e_{y'},e_{z'}\}$.

D. Kepler system. Finally, we consider the Kepler system in the plane. Skipping all details (which are left as an exercise), we observe that the fibration by the invariant circles is given in this case by the map F already introduced in Section 4.1, that is, the map with components $F_1=\tilde{A}_1$, $F_2=\tilde{A}_2$, $F_3=G$. (Of course, one should restrict himself to the subset of the phase space where $H<0$ and remove the collisions).

We now regard the Delaunay coordinates (L,G,l,g) as generalized action–angle coordinates of this fibration. The mean anomaly l is the angle, and (L,G,g) are coordinates on the base manifold. Specifically, $(G,\sqrt{L^2-G^2},g)$ are cylindric coordinates for the vector F with respect to an orthonormal basis with the axes e_1,e_2 in the plane of the orbit and e_3 orthogonal to it. This is seen by observing that, since

$$\|F\| \;=\; L\,,$$

one has $F_1^2+F_2^2 \;=\; L^2-G^2$, and that, since the Laplace vector is directed towards the perihelion, the argument of the perihelion g gives precisely the azimuth of F. Note also that the symplectic leaves are the spheres $\|F\|=\mathit{const}$. The circular orbits sit on the north and south poles of these spheres, what explains the singularity of the Delaunay angles.

Concerning the polar foliation, we mention that it has been proven by Moser [39] that the sets of negative constant energy are diffeomorphic to $\mathbb{R}P^3$, that is, to the unit tangent bundle of S^2. (Moreover, if one removes the collisions, on every such level set the flow is conjugate to the geodesic flow on the (punctured) sphere S^2).

4.5 Proof of Theorem 4.6

A. Another formulation of the theorem on noncommutative integrability. We prove here Theorem 4.6. Specifically, we prove the following Theorem, which is equivalent Theorem 4.6:

Theorem 4.10 *Let M be a symplectic manifold of dimension $2d$. Assume that there exist $n \leq d$ functions $F_1, \ldots, F_n : M \to \mathbb{R}$ with the following properties:*

a) $F_1, \ldots, F_n$ are pairwise in involution and have differentials which are linearly independent at every point of M.

b) In M, the integral manifolds of the distribution of the Hamiltonian vector fields $X_{F_1}, \ldots, X_{F_n}$ are compact, and are the fibers of a fibration.†

Then, each connected component of these integral manifolds is diffeomorphic to $\mathbb{T}^n$ and has a tubular neighbourhood U endowed with a set of generalized action–angle coordinates (a, p, q, α) such that:

i) the integral manifolds of the distribution generated by $X_{F_1}, \ldots, X_{F_n}$ are the level sets of the map (a, p, q);

ii) the coordinates $a_1, \ldots, a_n$ are functions of $F_1, \ldots, F_n$: that is, there exists a diffeomorphism $\tilde{a} : F(U) \to a(U)$ such that $a = \tilde{a} \circ f$, where $F = F_1 \times \ldots \times F_n$.

Remarks: (i) The hypothesis that the leaves of the foliation generated by $X_{F_1}, \ldots, X_{F_n}$ form a fibration is independent of the other hypotheses, as is see on the example of the 2 : 1 resonance of Section 1.4.

(ii) A careful inspection of the proof will show that it is not sufficient to require that the vector fields $X_{F_1}, \ldots, X_{F_n}$ pairwise commute—one precisely needs that the functions $F_1, \ldots, F_n$ are in involution.

The equivalence with Theorem 4.6 is seen as follows. On the one hand, under the hypotheses of Theorem 4.10, the fibration by the integral manifolds of the distribution spanned by $X_{F_1}, \ldots, X_{F_n}$ possesses a polar foliation, which is given by the level sets of the map F. On the other hand, given a fibration π_1 as in Theorem 4.6, we locally construct functions $F_1, \ldots, F_n$ whose Hamiltonian vector fields generate the fibers by lifting to the manifold any system of local coordinates on the action space.

Theorem 4.10 is of interest in itself, as a theorem on noncommutative integrability. One may then wonder how to verify, in practical cases, the hypotheses entering its statement. The following Theorem provides an useful criterion:

Proposition 4.11 *Let M be a symplectic manifold of dimension $2d$. Assume that there exists $2d - n$ functions $F_1, \ldots, F_{2d-n}$ which have differentials everywhere linearly independent in M, and with the properties that each of the first n of them is in involution with all the others:*

$$\{F_i, F_j\} \;=\; 0 \;, \qquad i = 1, \ldots, n \,, \quad j = 1, \ldots, 2d-n \,. \tag{4.5.1}$$

Then:

i) The common level sets of $F_1, \ldots, F_{2d-n}$ coincide with the integral manifolds of the distribution of $X_{F_1}, \ldots, X_{F_n}$. If they are compact, they are the fibers of a fibration.

ii) Let $H : U \to \mathbb{R}$ be a function in involution with $F_1, \ldots, F_{2d-n}$. Then, H is functionally dependent on $F_1, \ldots, F_n$.

Proof of Proposition 4.11. It follows from the hypotheses that the level sets of the map

† The integrability of the distribution spanned by $X_{F_1}, \ldots, X_{F_n}$ follows by Frobenius theorem, since these vector fields commute.

$(F_1, \ldots, F_n)$ are n–dimensional submanifolds of M, and that the vector fields $X_{F_1}, \ldots, X_{F_n}$ are tangent to them (equations (4.5.1) can indeed be written $L_{X_{F_i}} F_j = 0$). This proves the first statement in i), because of the uniqueness of the integral manifolds of a distribution. The second statement in ii) follows from the fact that any submersion with compact level sets is a fibration (Ehresmann's theorem).

We now prove statement ii). Let $\hat{H}(a, p, q, \alpha)$ be the local representative of H in a local system of generalized action–angle coordinates. Since $\{H, F_j\} = 0$ for all $j = 1, \ldots, 2d - n$, and since, by i), $F_1, \ldots, F_{2d-n}$ are functions of a, p, q alone, one gets $\{H, a_i\} = \{H, p_s\} = \{H, q_s\} = 0$ for all $i = 1, \ldots, n$ and all $s = 1, \ldots, d - n$. This implies that H is a function of the a_i's alone, which in turn are functions of $F_1, \ldots, F_n$. ■

Remark: Under the hypotheses of Theorem 4.10, there always exist *locally* functions $F_1, \ldots, F_{2d-n}$ as in Proposition 4.11: just take $F_1, \ldots, F_n$ and the coordinates (p, q). However, it may happen that no set of such functions exists *globally* – that is, in the whole subset of the phase space fibered by the invariant tori of dimension n. This is the reason why one needs to resort to the characterization given in the Proposition. The characterization through $2d - n$ functions in involution, which appeared in [14], is unable to account for the global structure of degenerate systems.

B. Proof of Theorem 4.10. Every connected component N of one of the integral manifolds of the distribution spanned by $X_{F_1}, \ldots, X_{F_n}$ is an n–dimensional compact connected manifold which has n linearly independent and pairwise commuting tangent vector fields. Hence, according to Proposition 2.4, it is diffeomorphic to the torus $\mathbb{T}^n$. In the sequel, we denote by $\mathcal{L}(N)$ the period lattice of N, defined as in the course of the proof of the Liouville–Arnol'd theorem (with reference to the action generated by the Hamiltonian vector fields of $F_1, \ldots, F_n$.

The proof of Theorem 4.10 follows the same lines of that of the Liouville–Arnol'd theorem, but there a few relevant differences. We pick a point $x_* \in M$, and denote by N_{x_*} the connected component of the integral manifold of $X_{F_1}, \ldots, X_{F_n}$ through it. First of all, we use a theorem by Caratheodory [7,14] to complete the functions $F_1, \ldots, F_n$ to a local system of symplectic coordinates: since these functions are in involution and their differentials are linearly independent at x_*, there exist a neighbourhood V of x_* and functions $T_1, \ldots, T_n, P_1, \ldots, P_{d-n}, Q_1, \ldots, Q_{d-n}$ such that:

- the symplectic 2–form of M, in V, is given by

$$\sum_{i=1}^{n} dF_i \wedge dT_i \;+\; \sum_{s=1}^{d-n} dP_s \wedge dQ_s\,; \tag{4.5.2}$$

- the map $F \times P \times Q \times T$ is a diffeomorphism of V onto $\mathcal{W} \times \mathcal{T}$, where $\mathcal{W} = (F \times P \times Q)(V) \subset \mathbb{R}^{2d-n}$ and $\mathcal{T} = T(V) \subset \mathbb{R}^n$ are open sets.†

We shall denote by $\mathcal{C} : \mathcal{W} \times \mathcal{T} \to V$ the inverse of the map $F \times P \times Q \times T$ (as well as its restrictions and extensions introduced in the sequel).

It follows from (4.5.2) that, in the local coordinates (F, P, Q, T), the flow of X_{F_i} is a shift of the coordinate T_i. Thus, in V, the integral manifolds of $X_{F_1}, \ldots, X_{F_n}$ are the sets

† The construction of the 'energy–time' coordinates in Section 2.2.B provides a proof of this theorem for the case $n = d$. Here, for shortness, we give it for granted.

$(F,P,Q) = const$ and carry coordinates $(T_1,\dots,T_n)$. As a consequence, the submanifold of V described by $T=0$ intersects transversally the leaves of this foliation. The important point is that, since this foliation is by hypothesis a fibration, if $\mathcal{W}$ is sufficiently small, then the submanifold $T=0$ intersects each leaf in exactly one point. Thus, the submanifold $T=0$ is the image of a (local) section

$$\sigma \;:\; \mathcal{W} \;\to\; V\,, \tag{4.5.3}$$

of this fibration, defined by $\sigma(b) = \mathcal{C}(b,0)$. Using this fact, we now extend the coordinates (F,P,Q,T) from V to a neighbourhood $N_{\mathcal{W}}$ of the torus N_{x_*}. To this purpose, we extend the map $\mathcal{C}$ to a map $\mathcal{W}\times\mathbb{T}^n \to M$, still denoted by $\mathcal{C}$, defined by

$$\mathcal{C}(b,\tau) \;=\; \Phi_\tau(\sigma(b))\,.$$

Let $N_{\mathcal{W}} := \mathcal{C}(\mathcal{W}\times\mathbb{R}^n)$. Note that $\mathcal{C}$ maps surjectively the subspaces $b = const$ onto the connected components N_x of the integral manifolds of $X_{F_1},\dots,X_{F_n}$ which intersect the section σ. Thus, the set $N_{\mathcal{W}}$ is the union of these leaves (which are diffeomorphic to tori). Proceeding exactly as in the proof of the Liouville–Arnol'd theorem, one sees that $\mathcal{C}$ is symplectic and is a local diffeomorphism, even though it is of course not injective: all points (b,τ), (b,τ') with $\tau'-\tau \in \mathcal{L}(N_{\sigma(b)})$ are mapped into the same point of $N_{\mathcal{W}}$. In the sequel, we denote by (F,P,Q,T) the local inverses of $\mathcal{C}$, and by the small letters (f,p,q,τ) the points and the coordinates in $\mathcal{W}\times\mathbb{R}^n$.

Our purpose now is to perform a symplectic change of coordinates $(f,p,q,\tau)\mapsto(a,p,q,\hat\alpha)$ on $\mathcal{W}\times\mathbb{R}^n$,[‡] with

$$\hat\alpha \;=\; S^{-1}\tau\,, \qquad S \;=\; 2\pi\, L(N)^{-1}\,, \tag{4.5.4}$$

in such a way to normalize the periods. Of course, the period lattice and the matrix S vary from leaf to leaf but, as we now show, they depend only on F; this is important, because this is precisely what will allow us to keep the p's and q's unchanged and to take the a's function of the f's alone.

Lemma 4.12 *$\mathcal{L}(N_x) = \mathcal{L}(N_{x'})$ whenever $x,x'\in N_{\mathcal{W}}$ and $F(x) = F(x')$.*

Proof. Let us consider the map $\Psi_t : N_{\mathcal{W}} \to M$ defined by

$$\Psi_t \;=\; \Phi^{X_{F_1}}_{t_1}\circ\ldots\circ\Phi^{X_{F_n}}_{t_n}\circ\Phi^{X_{P_1}}_{t_{n+1}}\circ\ldots\circ\Phi^{X_{P_{d-n}}}_{t_d}\circ\Phi^{X_{Q_1}}_{t_{d+1}}\circ\ldots\circ\Phi^{X_{Q_{d-n}}}_{t_{2d-n}}\,,$$

where $t\in\mathbb{R}^{2d-n}$. Since the vector fields X_{F_i}, X_{P_s}, X_{Q_s} commute with X_{F_j} $(i,j=1,\dots,n,$ $s=1,\dots,d-n)$, one has

$$\Psi_t\circ\Phi_\tau \;=\; \Phi_\tau\circ\Psi_t \qquad \textit{for all } t\in\mathbb{R}^{2d-n} \textit{ and all } \tau\in\mathbb{R}^n\,,$$

where $\Phi_\tau = \Phi^{X_{F_1}}_{\tau_1}\circ\cdots\circ\Phi^{X_{F_n}}_{\tau_n}$. This implies that any two points x, x' of $N_{\mathcal{W}}$ which are such that $x' = \Psi_t(x)$ for some $t\in\mathbb{R}^{2d-n}$ have the same period lattice. It is easy to verify that this is the case whenever $F(x') = F(x)$: indeed, if (f,p',q',τ') and (f,p,q,τ) are the

[‡] More precisely, in some subset of it of the form $\mathcal{W}'\times\mathbb{R}^n$, $\mathcal{W}'\subset\mathcal{W}$, but for the sake of notation we shall keep writing $\mathcal{W}$ instead than $\mathcal{W}'$.

coordinates of x and x', respectively, then, since $X_{F_i} = -\frac{\partial}{\partial T_i}$, $X_{P_s} = -\frac{\partial}{\partial Q_s}$, and $X_{Q_s} = \frac{\partial}{\partial P_s}$, one has $x' = \Psi_t(x)$ with $t = (-\tau_1' + \tau_1, \dots, -q_1' + q_1, \dots, p_1' - p_1, \dots)$. ∎

From now on, we shall denote the period lattice $\mathcal{L}(N_x)$ by $\mathcal{L}(f)$, where $f = F(x)$, and the corresponding period matrix by $L(f)$. We establish now under which conditions there exists a symplectic diffeomorphism

$$\mathcal{S} \,:\, \mathcal{W} \times \mathbb{R}^n \to \hat{B} \times \mathbb{R}^n \,, \qquad (f,p,q,\tau) \mapsto (a(f),p,q,S^{-1}\tau) \tag{4.5.5}$$

with some map a. To this purpose observe that the condition for $\mathcal{S}$ to be symplectic is $\sum_i da_i \wedge d\hat{\alpha}_i = \sum_i df_i \wedge d\tau_i$ and that one has

$$da_i \wedge d\hat{\alpha}_i \;=\; \sum_j \frac{\partial a_i}{\partial f_j} df_j \wedge d\hat{\alpha}_i \,,$$

$$df_i \wedge d\tau_i \;=\; \frac{1}{2\pi}\, df_i \wedge \left(\sum_{jl} \frac{\partial L_{ij}}{\partial f_l} \hat{\alpha}_j df_l \;+\; \sum_j L_{ij} d\hat{\alpha}_j \right) .$$

Since the period matrix depends only on f, the symplecticity condition reduces to

$$\frac{\partial a_i}{\partial f_j} \;=\; \frac{1}{2\pi} L_{ji}(f) \,, \qquad i,j \;=\; 1,\dots,n \tag{4.5.6a}$$

$$\frac{\partial L_{ij}}{\partial f_l}(f) \;=\; \frac{\partial L_{lj}}{\partial f_i}(f) \,, \qquad i,j,l \;=\; 1,\dots,n \,. \tag{4.5.6b}$$

Equations (4.5.6b) are the closedness conditions which assure that (4.5.6a) can be integrated to get a map $a = a(f)$ defined in some subset of $F(N_{\mathcal{W}})$. Since, as we show below, the period matrix satisfies (4.5.6b), we conclude that, restricting $\mathcal{W}$ if needed, there exists a symplectic diffeomorphism $\widehat{\mathcal{S}}$ as in (4.5.5). The proof of Theorem 4.10 is completed by considering the quotient $(a,p,q,\hat{\alpha}) \mapsto (a,p,q,\hat{\alpha}(mod 2\pi))$, which provides the required diffeomorphism (a,p,q,α) between $N_{\mathcal{W}}$ and $\hat{B} \times \mathbb{T}^n$.

Equations (4.5.6b) are a consequence of the fact that the section σ defined in (4.5.3) is coisotropic (since $T = 0$ is obviously coisotropic, and $\mathcal{C}$ is symplectic) and from the following

Lemma 4.13 *The period matrix satisfies (4.5.6b) if and only if the section $\sigma : \mathcal{W} \to N_{\mathcal{W}}$ is coisotropic.*

Proof. The pre–images under $\mathcal{C}$ of the submanifold $\sigma(\mathcal{W})$ are the submanifolds of $\mathcal{W} \times \mathbb{R}^n$ given by

$$\Sigma_\nu \;:\; (f,p,q) \mapsto (f,p,q,L(f)\nu) \qquad \nu \in \mathbb{Z}^n \,.$$

Since $\mathcal{C}$ is symplectic, σ is coisotropic if and only if Σ_ν is coisotropic, for some $\nu \in \mathbb{Z}^n$. By definition, this means that the symplectic orthogonal $(T_{(b,\tau)}\Sigma_\nu)^\perp$ to every tangent space $T_{(b,\tau)}\Sigma_\nu$, $(b,\tau) \in \Sigma_\nu$, is isotropic. Now, since a basis for the tangent spaces $T_{(b,\tau)}\Sigma_\nu$ is provided by the vector fields

$$W_i \;=\; \frac{\partial}{\partial f_i} \;+\; \sum_{jl+1}^{n} \frac{\partial L_{jl}}{\partial f_i}\, \nu_l \,\frac{\partial}{\partial \tau_i} \,, \qquad Y_s \;=\; \frac{\partial}{\partial p_s} \,, \qquad Z_s \;=\; \frac{\partial}{\partial q_s} \,, \tag{4.5.7}$$

with $i = 1, \ldots, n$ and $s = 1, \ldots, d-n$, a vector V belongs to $(T_{(b,\tau)}\Sigma_\nu)^\perp$ if and only if $\Omega(V, W_i) = \Omega(V, Y_s) = \Omega(V, Z_s) = 0$ for all i and s (here, it is understood that the symplectic 2–form Ω and all the vector fields are evaluated at the point (b, τ)). Writing

$$V \;=\; \sum_i \left(V_{f_i} \frac{\partial}{\partial f_i} + V_{\tau_i} \frac{\partial}{\partial \tau_i} \right) \;+\; \sum_s \left(V_{p_s} \frac{\partial}{\partial p_s} + V_{q_s} \frac{\partial}{\partial q_s} \right) ,$$

and using the expression (4.5.2) for Ω, one computes

$$\Omega(V, W_i) \;=\; \sum_{jl} V_{f_j} \frac{\partial P_{jl}}{\partial f_i} \nu_l \;-\; V_{\tau_i} \,, \qquad \Omega(V, Y_s) \;=\; -V_{q_s} \,, \qquad \Omega(V, Z_s) \;=\; V_{p_s} \,.$$

Thus, V belongs to $(T_{(b,\tau)}\Sigma_\nu)^\perp$ if and only if $V = \sum_{i=1}^n V_{f_i} W_i$. In this way, we have shown that the vector fields $W_1, \ldots, W_n$ provide a base for the spaces $(T_{(b,\tau)}\Sigma_\nu)^\perp$. It follows that $(T_{(b,\tau)}\Sigma_\nu)^\perp$ is isotropic if and only if $\Omega(X_i, X_j) = 0$ for all $i, j = 1, \ldots, n$. Since

$$\Omega(W_i, W_j) \;=\; \sum_l (W_i)_{f_l} (W_j)_{\tau_l} \;-\; (W_i)_{\tau_l} (W_j)_{f_l} \;=\; \sum_k \nu_k \left(\frac{\partial L_{ik}}{\partial f_j} \;-\; \frac{\partial L_{jk}}{\partial f_i} \right)$$

this condition coincides with (4.5.6b). ∎

Chapter 5

Symmetry and Integrability

This short final chapter is devoted to a few considerations about a topic which has a foundational interest, namely, under which conditions the existence of symmetries implies the integrability of a Hamiltonian system. Specifically, we ask under which conditions the level sets of the momentum map (or else, of the energy–momentum map) are invariant tori. We have seen in section 1.8, for instance, that this happens for the Euler–Poinsot system. A more delicate question is the converse question, namely, whether integrability can always be explained as due to the existence of symmetries, in the precise sense that the fibration by the invariant tori is the momentum (or the energy–momentum) map of a suitable group action. The notion of noncommutative integrability, and its geometric picture, play an essential role in this analysis.

5.1 Hamiltonian actions

In the course of this analysis, we will use a number of properties of Lie groups, group actions, and reduction. We assume that the reader is familiar with these topics, and we limit ourselves to recall the facts that we need. For an exhaustive treatment, we refer to [7,6,9].

We consider a connected Lie group G, with Lie algebra $\mathcal{G}$. The dual $\mathcal{G}^*$ of $\mathcal{G}$ carries a natural Poisson structure, called the (+) Lie–Poisson brackets, which we denote $\{\ ,\ \}_{\mathcal{G}^*}$; the symplectic leaves of this Poisson structure are the orbits of the coadjoint representation, which we denote $\mathcal{O}_\mu$, $\mu \in \mathcal{G}^*$. Note that

$$dim\, \mathcal{O}_\mu \ = \ dim\, G - dim\, G_\mu \tag{5.1.1}$$

where G_μ is the isotropy subgroup of the point $\mu \in \mathcal{G}^*$ in the coadjoint representation.

Consider now an action $\Phi : G \times M \to M$ of G on a symplectic manifold (M, Ω). As is customary, we write $\Phi_g(m)$, or else $g.x$, for $\Phi(g, m)$, and we denote by $G.x$ the orbit through the point $x \in M$ and by G_x the isotropy subgroup of x.

The action Φ is called *Hamiltonian* (or Poisson [4] or strongly Hamiltonian [7]) if it has the following properties:

i. It is symplectic, i.e., $\Phi_g^*\Omega = \Omega$ for all $g \in G$.
ii. For every $\xi \in \mathcal{G}$, the infinitesimal generator ξ_M has a global Hamiltonian J_ξ.
iii. For every $\eta, \xi \in \mathcal{G}$, the Hamiltonian of the Lie derivative $[\eta_M, \xi_M]$ is the Poisson bracket $\{J_\eta, J_\xi\}_M$.†

A Hamiltonian action possesses a momentum map $\mathbf{J} : M \to \mathcal{G}^*$ which is equivariant with respect to the coadjoint action of G on $\mathcal{G}^*$, i.e., $\mathbf{J} \circ \Phi_g = ad_g^* \circ \mathbf{J}$.

We now review a few basic properties of momentum maps:

(G1) ker $(T_x\mathbf{J}) = \big(T_x(G.x)\big)^\perp$ *at every point* $x \in M$.

In the sequel, we shall assume that, for every $\mu \in \mathbf{J}(M) \subset \mathcal{G}^*$, the level set $\mathbf{J}^{-1}(\mu)$ is a submanifold of M. In such a case, property G1 becomes $(T_x\mathbf{J}^{-1}(\mu)) = \big(T_x(G.x)\big)^\perp$ for all $x \in \mathbf{J}^{-1}(\mu)$, namely, the fibers of $\mathbf{J}$ and the orbits of G are symplectically orthogonal. This implies

$$\begin{aligned} dim\,\mathbf{J}^{-1}(\mu) &= dim\,M - dim\,G.\,x && (5.1.2a) \\ &= dim\,M - dim\,G + dim\,G_x && (5.1.2b) \end{aligned}$$

since $dim\,G.\,x = dim\,G - dim\,G_x$.‡

The orbits of a (connected) group are the leaves of a foliation whenever they have all the same dimension. In the present case, by equation (5.1.2), this happens if and only if the level sets of $\mathbf{J}$ have all the same dimension, in which case property G1 states that the level sets of the momentum map and the orbits of the group are the leaves of two foliations which are polar to each other.

Note that this situation is met in the special case in which $\mathbf{J} : M \to \mathcal{G}^*$ is a submersion (which happens if and only if the isotropy subgroup G_x is discrete for any $x \in M$) and also in the more general situation in which $\mathbf{J}(M)$ is a submanifold of $\mathcal{G}^*$ and $\mathbf{J} : M \to \mathbf{J}(M)$ is a submersion.

(G2) If $\mathbf{J}^{-1}(\mu)$ *is a submanifold of* M*, then*

$$T_x\mathbf{J}^{-1}(\mu) \,\cap\, T_x(G.\,x) \;=\; T_x(G_\mu.\,x)\,, \qquad \forall x \in \mathbf{J}^{-1}(\mu)\,.$$

Together with property G1, this implies

$$T_x\mathbf{J}^{-1}(\mu) \,\cap\, \big[T_x\mathbf{J}^{-1}(\mu)\big]^\perp \;=\; T_x(G_\mu.\,x)\,, \qquad \forall x \in \mathbf{J}^{-1}(\mu)\,,$$

so that $\mathbf{J}^{-1}(\mu)$ is isotropic if and only if

$$\begin{aligned} dim\,\mathbf{J}^{-1}(\mu) &= dim\,(G_\mu.\,x) \\ &= dim\,G_\mu - dim\,G_x \end{aligned} \qquad (5.1.3)$$

for any $x \in \mathbf{J}^{-1}(\mu)$. The latter equality in (5.1.3) follows form the fact that, as one immediately verifies, the equivariance of the momentum map implies $G_x \subset G_{\mathbf{J}(x)}$.

† We denote by $\{\ ,\ \}_M$ the Poisson brackets on M induced by its symplectic structure.

‡ Note that $\dim G_x$ is constant on every level set of the momentum map.

(G3) $\mathbf{J} : (M, \{\ ,\ \}_M) \to (\mathcal{G}^*, \{\ ,\ \}_{\mathcal{G}^*})$ *is a Poisson morphism.*

(G4) For any $x \in M$*, the restriction of* $\mathbf{J}$ *to the orbit* $G.x$ *is a surjective submersion onto the coadjoint orbit* $\mathcal{O}_{\mathbf{J}(x)}$*.*

Property G4 implies that $\mathbf{J}(M)$ is a union of coadjoint orbits. Hence, if $\mathbf{J}(M)$ is a submanifold of $\mathcal{G}^*$, then it is a Poisson submanifold of $\mathcal{G}^*$ with the Lie–Poisson structure. (Indeed, a submanifold of a Poisson manifold is a Poisson submanifold if and only if it entirely contains any symplectic leaf that it intersects).

Finally, we recall that the *reduced phase space* corresponding to the value $\mu \in \mathbf{J}(M)$ is the quotient space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$. If, for a given μ, $\mathbf{J}^{-1}(\mu)$ and P_μ are both manifolds, then one has

$$\begin{aligned} \dim P_\mu &= \dim \mathbf{J}^{-1}(\mu) - \dim (G_\mu . x) \\ &= \dim \mathbf{J}^{-1}(\mu) - \dim G_\mu - \dim G_x \,. \end{aligned} \tag{5.1.4}$$

5.2. Symmetry and noncommutative integrability

A. Momentum maps and integrability. We consider now the question of under which conditions the existence of a symmetry group implies the noncommutative integrability of the system. Specifically, the question that we consider is whether the momentum map provides the fibration by invariant tori—namely, if its level sets are isotropic and have a polar foliation. The analysis on the previous Section immediately leads to the following:

Proposition 5.1 *Consider a Hamiltonian action of a connected Lie group* G *on a symplectic manifold* M_* *and a* G*–invariant Hamiltonian* H *on* M_**. Let* M *be a* G*–invariant, open, connected subset of* M_* *and assume that*

- $\mathbf{J}(M)$ *is a submanifold of* $\mathcal{G}^*$ *of dimension* $2d - n$*, with some* $1 \le n \le d$*.*†
- *In* M*, the level sets of* $\mathbf{J}$ *are connected.*

Then:

(i) The level sets of $\mathbf{J}$ *in* M *are the leaves of a foliation which possesses a polar foliation (the leaves of which are the connected components of the group orbits).*

(ii) The level sets of $\mathbf{J}$ *in* M *are isotropic if and only if*

$$\begin{aligned} \dim \mathcal{O}_\mu &= 2 \dim \mathbf{J}(M) - \dim M \\ &= 2d - 2n \end{aligned} \tag{5.2.1}$$

for all $\mu \in \mathbf{J}(M)$*. Condition (5.2.1) is equivalent to the fact that* $\mathbf{J}(M)$ *has everywhere* n *independent Casimirs, or else to the fact that* $\mathcal{G}^*$ *has everywhere* $\dim G - 2d + 2n$ *independent Casimirs.*

---

† Hence, $\mathbf{J}|_M$ has constant rank equal to $2d - n$, and the orbits of G in M have all dimension $2d - n$.

If the reduced phase space P_μ is a manifold, then equation (5.2.1) is equivalent to $dim\, P_\mu = 0$ for all $\mu \in \mathbf{J}(M)$.

Proof. We have already proven that, under the stated hypotheses, the level sets of $\mathbf{J}$ are the leaves of a bifoliation. By property B4.ii of Proposition 4.4, the level sets of $\mathbf{J}$ are isotropic if and only if the rank of the induced Poisson structure of $\mathbf{J}(M)$ equals $2dim\, \mathbf{J}(M) - dim\, M$. On account of property G3, the rank of the Poisson structure of $\mathbf{J}(M)$ at a point μ equals $dim\, \mathcal{O}_\mu$. This proves condition (5.2.1). The equivalent statements follow by recalling that the rank of a Poisson manifold is the codimension of the symplectic leaf. The statement about the reduced phase space follows from (5.1.4). ∎

Under the hypotheses of the Proposition, if the level sets of $\mathbf{J}$ are compact, then we are exactly in the situation described by Theorem 4.6, so the system is integrable in the noncommutative sense and its invariant tori are the level sets of the momentum map. The leaves of the polar foliation (the flowers of Figure 4.1.b) are the level sets of the (lifts to M) of the Casimirs of $\mathbf{J}(M)$, so that the (local) actions a are functions of the Casimirs of $\mathbf{J}(M)$. Furthermore, the symplectic leaves of $\mathbf{J}(M)$ are the connected components of the level sets of the Casimirs, that is, the coadjoint orbits. Proposition 5.1 shows that the existence of a symmetry group implies or not the integrability of a system depending on general properties of the group, of the action, and of the manifold on which it acts. Instead, as is obvious, the only relevant property of the Hamiltonian is its invariance under the group action.

Remarks: (i) In the statement of the Proposition we have not assumed that the momentum map $\mathbf{J} : M \to \mathcal{G}^*$ is a submersion (i.e., that the action is locally free). This condition is in fact not verified in cases in which the group is “too large” (i.e., has dimension $> 2n - d$), as it happens for instance for the d–dimensional isotropic oscillator if $d \geq 3$ (see below).

(ii) If the action is only locally Hamiltonian (Hamiltonian, in the language of [7]), then Proposition 5.1 still holds true, with the coadjoint action replaced by the affine action on $\mathcal{G}^*$ which makes $\mathbf{J}$ equivariant.

Exercise: Show that (under suitable hypotheses) a system which has a symmetry group whose all orbits have codimension one is integrable in the noncommutative sense.

B. Energy–momentum maps and integrability. In certain cases, the energy is independent of the momentum map, and one should of course consider as primary object the energy–momentum map, rather than the momentum map, since it defines a finer invariant foliation. For this reason, we give here the analogue of Proposition 5.1 for this case. However, as we discuss in the proof, it is always possible to enlarge the symmetry group so as to get the energy–momentum map as momentum map of a new action, so this case does not add anything new.

Proposition 5.2 *Under the same hypotheses of Proposition 5.1, assume that the restriction $(H \times \mathbf{J})|_M$ of $H \times \mathbf{J}$ to M has constant rank, equal to $rank\,(\mathbf{J}|_M) + 1$, that it has connected level sets, and that $(H \times \mathbf{J})(M)$ is a submanifold of $\mathbb{R} \times \mathcal{G}^*$. Then, the level sets of $(H \times \mathbf{J})|_M$ are the leaves of a bifoliation. They are isotropic if and only if*

$$dim\, \mathcal{O}_\mu \;=\; 2 \;+\; 2\, dim\, \mathbf{J}(M) \;-\; dim\, M\,. \tag{5.2.2}$$

If the reduced phase space P_μ exist, then equation (5.2.2) is equivalent to $dim\, P_\mu \;=\; 2$.

Proof. Let us consider the product group $\tilde{G} = \mathbb{R} \times G$, with the composition $(t, g)(t', g') = (t + t', gg')$, and its action on M defined by

$$(t, g).\, x \;=\; \Phi_t^H(g.\, x)\,,$$

where Φ_t^H is the map at time t of the flow of the Hamiltonian vector field of H. This action is well defined on account of the commutation relation

$$\Phi_t^H(g.\, x) \;=\; g.\, \Phi_t^H(x)\,, \qquad \forall t \in \mathbb{R}\,,\; x \in M\,,$$

which follows from the fact that, since H is invariant under the action of G, $[X_H, \xi_M] = -L_{\xi_M} H = 0$ for every infinitesimal generator ξ_M of the G–action. The Lie algebra $\tilde{\mathcal{G}}$ of $\tilde{G}$ is $\mathbb{R} \times \mathcal{G}$ with the brackets $[(v, \xi), (v', \xi')]_{\tilde{\mathcal{G}}} = [\xi, \xi']_{\mathcal{G}}$. The momentum map of the action of $\tilde{G}$ is

$$\tilde{\mathbf{J}} \,:\, M \to \tilde{\mathcal{G}}^*\,, \qquad \tilde{\mathbf{J}}(x) \;=\; (H(x), \mathbf{J}(x))\,.$$

So, under the stated hypotheses, the action of $\tilde{G}$ satisfies all the hypotheses of Proposition 5.1, and equation (5.2.1) takes the form (5.2.2) since $dim\, \tilde{\mathbf{J}}(M) = dim\, \mathbf{J}(M) + 1$. In order to prove the statement relative to the reduced phase space, we compute, from (5.1.4), (5.1.1) and (5.1.2b),

$$\begin{aligned} dim\, P_\mu \;&=\; dim\, \mathbf{J}^{-1}(\mu) - dim\, G_\mu + dim\, G_x \\ &=\; dim\, \mathbf{J}^{-1}(\mu) + dim\, \mathcal{O}_\mu - dim\, G + dim\, G_x \\ &=\; 2\, dim\, \mathbf{J}^{-1}(\mu) + dim\, \mathcal{O}_\mu - dim\, M \\ &=\; dim\, M - 2\, dim\, \mathbf{J}(M) + dim\, \mathcal{O}_\mu \end{aligned}$$

which proves the claim. ∎

C. Is integrability always due to symmetry? The above results provide some sufficient criteria for symmetry to imply integrability. A few examples will be given in the next Section. Here, we want to spend at least a few warning words on the converse to this problem, namely, whether symmetry can be regarded as the reason of integrability of every integrable system.

Specifically, we ask whether the fibration by the invariant tori of a (noncommutatively) integrable system can always be identified with the momentum map, or with the energy–momentum map, of some Hamiltonian group action.

This is trivially true, at least *locally*, for completely integrable systems. Indeed, the flows of the Hamiltonian vector fields of the local actions define a Hamiltonian action of $\mathbb{T}^n$ whose momentum map is provided by the actions themselves. However, such an action can be extended to the whole subset fibered by the Lagrangian tori only under certain conditions, namely, if and only if the fibration has "trivial monodromy" [1,33].

In degenerate cases, however, the situation is not so obvious. The fact is that the invariant tori, being isotropic but not Lagrangian, do not coincide with the leaves of the polar foliation. The orbits of the group should be not the invariant tori, but the coisotropic leaves of the polar foliation. Therefore, it is not even clear whether such a group always exists.

Consider, for instance, the case of a 'completely degenerate' system, whose all orbits are periodic. In this case, as we have already observed, the group orbits should coincide with the connected components of the level sets of the Hamiltonian. So, the problem is whether there exists a group action whose orbits are the level sets of the Hamiltonian.

5.3 An example

As an example, we consider now the two–dimensional isotropic oscillator. We use the complex coordinates

$$z_j \;=\; p_j + i q_j\,, \qquad j = 1,2\,,$$

and regard the system as defined on $\mathbb{C}^2$ with the symplectic 2–form $\Omega = d\Theta$, with

$$\Theta \;=\; \frac{1}{2i}\sum_{j=1}^{2} \overline{z}_j dz_j\,.$$

The Hamiltonian is

$$H(z) \;=\; \frac{1}{2}\sum_{j=1}^{2} z_j \overline{z}_j\,.$$

The system is invariant under the linear action $z \mapsto gz$ of the group $SU(2) = \{g \in U(2) : det\, g = 1\}$. In fact, the system is also invariant under $U(2) = \{g \in GL(2,\mathbb{C}) : g^\dagger g = \mathbb{1}\}$ but, as we shall see, $U(2)$ is unnecessarily large.

The Lie algebra $su(2)$ is the set of all two–by–two complex matrices $\hat\xi$ which are anti–hermitian ($\hat\xi^\dagger = -\hat\xi$) and traceless, the Lie bracket being the commutator. As is well known, $su(2)$ is isomorphic to $\mathbb{R}^3$. Specifically, the isomorphism is realized by associating to any vector $\xi = (\xi_1,\xi_2,\xi_3) \in \mathbb{R}^3$ the matrix $\hat\xi \in su(2)$ given by

$$\hat\xi \;=\; \begin{pmatrix} \xi_3 & \xi_1 - i\xi_2 \\ \xi_1 + i\xi_2 & \xi_3 \end{pmatrix} \;=\; \sum_{j=1}^{3} \xi_j\,\sigma_j \;\equiv\; \xi\cdot\sigma\,, \tag{5.3.1}$$

where

$$\sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 \;=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

are the Pauli matrices.†

We compute now the momentum map of this action. Since the action is linear, the infinitesimal generator of the subgroup $exp\, t\hat\xi$ is the vector field $\hat\xi_{\mathbb{C}^2}(z) = \hat\xi z$. Since the symplectic 2–form is exact, $\hat\xi_{\mathbb{C}^2}(z)$ has the Hamiltonian $J_{\hat\xi} = \Theta(\hat\xi_{\mathbb{C}^2})$, that is,

$$\begin{aligned} J_{\hat\xi}(z) \;&=\; \frac{1}{2i}\sum_{j=1}^{2} \overline{z}_j dz_j(\hat\xi z) \\ &=\; \frac{1}{2i}\,(z,\hat\xi z) \\ &=\; \frac{1}{2i}\left[\xi_1(\overline{z}_1 z_2 + \overline{z}_2 z_1) - i\xi_2(\overline{z}_1 z_2 - \overline{z}_2 z_1) + \xi_3(\overline{z}_2 z_2 - \overline{z}_1 z_1)\right], \end{aligned}$$

where (,) denotes the Hermitian scalar product on $\mathbb{C}^2$ and we have used (5.3.1). Since $J_{\hat\xi}(z) = \mathbf{J}(z)(\hat\xi z)$, we conclude

$$\mathbf{J}(z) \;=\; \frac{1}{2i}\sum_{k=1}^{3} J_k(z)\sigma_k \;\equiv\; J(z)\cdot\sigma\,,$$

† In order to have a Lie algebras isomorphism between $su(2)$ and $\mathbb{R}^3$ with the vector product as bracket, one should identify $\xi \in \mathbb{R}^3$ with $\frac{1}{2i}\xi\cdot\sigma$.

where

$$\mathbf{J}(z) = \begin{pmatrix} -i(z_1\overline{z}_2 + \overline{z}_1 z_2)/2 \\ i(z_1\overline{z}_2 - \overline{z}_1 z_2)/2 \\ -i(z_1\overline{z}_1 - z_2\overline{z}_2)/2 \end{pmatrix} = \begin{pmatrix} -iF_1/2 \\ iF_2/2 \\ -iF_3/2 \end{pmatrix}$$

(see (4.4.2)).

So, we see that the momentum map of the $SU(2)$ action gives exactly the fibration by the periodic orbits of the system. In particular, the symplectic leaves of the base manifold, that is $\mathbf{J}(\mathbb{C}^2 \setminus \{0\}) \approx \mathbb{R}^3 \setminus \{0\}$ are the coadjoint orbits of $su(2)$, namely, the spheres $(2iJ_1)^2 + (2iJ_2)^2 + (2iJ^2)^2 = H^2$.

Exercise: Show that the momentum map $\mathbf{J} : \mathbb{C}^2 \to u(2)^* \approx \mathbb{R}^4$ of the $U(2)$ action is not a submersion but that $\mathbf{J}(\mathbb{C}^2 \setminus \{0\})$ is a submanifold of $u(2)^*$. (Hints: a basis for $u(2)$ is provided by the three Pauli matrices and $\sigma_4 = \mathbb{1}$. The momentum map is $(-iF_1/2, iF_2/2, -iF_3/2, iH)$ and its image is a three–dimensional cone).

Finally, we make a remark about the d–dimensional isotropic oscillator, $d \geq 3$. If we use complex coordinates $z_j = \frac{p_j + iq_j}{\sqrt{2}}$, the system is defined on $(\mathbb{C}^d, d\Theta)$, $\Theta = i\sum_j \overline{z}_j dz_j$ by the Hamiltonian

$$H(z) = \sum_{j=1}^{d} z_j \overline{z}_j \,,$$

and is clearly invariant under the linear action $(g, z) \mapsto gz$ of $SU(d)$, as well as of $U(d)$. Now, $SU(d)$ has dimension $d^2 - 1$, which for $d \geq 3$ is greater than the dimension of the base of the fibration by the periodic orbits, which is $2d - 1$. Therefore, the momentum map $\mathbf{J} : \mathbb{C}^d \to so(d)^*$ of the action of $SU(d)$ cannot be a submersion and, as in Proposition 5.1, we should regard it as a map onto $\mathbf{J}(\mathbb{C}^d \setminus \{0\})$.

Exercise: Compute the momentum map of the action of $U(d)$ on $\mathbb{C}^d$. (Hints: $u(d)$ has a basis constituted by the d^2 matrices

$$I^-_{jk} = E_{jk} - E_{kj} \,, \qquad I^+_{jk} = i(E_{jk} + E_{kj}) \,, \qquad I^0_k = iE_{kk} \,, \qquad j < k = 1, \ldots, d \,,$$

where E_{jk} is the matrix with entries $(E_{jk})_{rs} = \delta_{jr}\delta_{ks}$. Writing $\hat{\xi} = \sum_{j<k} \left(\xi^-_{jk} I^-_{jk} + \xi^-_{jk} I^-_{jk}\right) + \sum_k \xi^0_k I^0_{kk}$ and using the fact that the infinitesimal generator $\hat{\xi}_{\mathbb{C}^d}(z) = \hat{\xi}z$ has the Hamiltonian $J_\xi(z) == i(z, \xi z)$, one obtains

$$\begin{aligned} J_{\hat{\xi}}(z) &= i \sum_{j<k} [\xi^-_{jk}(z, I^-_{jk} z) + \xi^-_{jk}(z, I^-_{jk} z)] - \sum_k \xi^0_k (z, I^0_k z) \\ &= \sum_{j<k} [i\xi^-_{jk}(\overline{z}_j z_k - \overline{z}_k z_j) - \xi^+_{jk}(\overline{z}_j z_k + \overline{z}_k z_j)] - \sum_k \xi^0_k \overline{z}_k z_k \,. \end{aligned}$$

Hence, using the basis $\{I^-_{jk}, I^+_{jk}, I^0_k\}$ in $u(d)^*$, the momentum map $\mathbf{J}(z)$ is given by

$$\mathbf{J}(z) = \sum_{j<k} [\, i(\overline{z}_j z_k - \overline{z}_k z_j) I^-_{jk} - (\overline{z}_j z_k + \overline{z}_k z_j) I^+_{jk} \,] - \sum_k \overline{z}_k z_k \, I^0_k \,.$$

Note: In the case $d = 3$, only five of the nine components of $\mathbf{J}$ are independent, see [35], Chapter 1, Section 2.4).

References

[1] J. J. Duistermaat, *On global action–angle coordinates.* Commun. Pure Appl. Math. **33**, 687-706 (1980).

[2] G. Gallavotti, *Quasi–Integrable Mechanical Systems,* in *Critical phenomena, Random Systems, Gauge Theories,* K. Osterwalder and R. Stora editors, *Les Houches,* Session XLIII, 1984 (North–Holland, Amsterdam 1986).

[3] R. Cushman and L. Bates *Global aspects of classical integrable systems.* (Birkhäuser,Basel, 1997)

[4] V.I. Arnold, *Méthodes Mathématiques de la Mécanique Classique* (MIR, Moscow 1976).

[5] V.I. Arnold, *Chapitres supplémentaires de la théorie des équations différentielles ordinaires* (MIR, Moscow 1980).

[6] R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, Reading 1978).

[7] P. Libermann and C.-M. Marle, *Symplectic geometry and analytical mechanics.* (D. Reidel, Dordrecht, 1987)

[8] V. Guillemin and S. Sternberg, *Symplectic techniques in physics* (Cambridge University Press, Cambridge, 1984)

[9] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry.* Texts in Applied Mathematics **17** (Springer, New York, 1994)

[10] C. Camacho and A.L. Neto, *Geometric theory of foliations* (Birkhäuser, Boston, 1985).

[11] V.I. Arnold, *Small denominators and problems of stability of motion in classical and celestial mechanics.* Usp. Mat. Nauk **18**:6, 91-192 (1963) [Russ. Math. Surv. **18**:6, 85-191 (1963)].

[12] V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York 1968).

[13] R. Jost, *Winkel- und Wirkungsvariable für allgemeine mechanische Systeme.* Helvetica Physica Acta **41**, 965-968 (1968).

[14] N.N. Nekhoroshev, *Action–angle variables and their generalizations.* Trudy Moskov. Mat. Obsc. **26**, 181-198 (1972) [Trans. Moskow Math. Soc. **26**, 180-198 (1972)].

[15] L. Markus and K.R. Meyer, *Generic Hamiltonian dynamical systems are neither integrable nor ergodic.* Memoirs of the Am. Math. Soc. **144**, 1-52 (1974)

[16] E. T. Whittaker, *A treatise on the analytical dynamics of particles & rigid bodies* (Cambridge Mathematical Library, Cambridge, 1937)

[17] L.M. Bates and J. Sniatycki, *On action–angle variables.* Arch. Rational Mech. Anal. **120**, 337-343 (1992).

[18] A. Weinstein, *Symplectic manifolds and their lagrangian submanifolds.* Advances in Math. **6**, 329-346 (1971).

[19] P.M. Gruber and C.G. Lekkerkerker, *Geomtery of numbers* (North Holland, Amsterdam, 1987)

[20] J. Vey, *Sur certain systemes dynamiques separables.* Amer. J. Math. **100**, 591-614 (1978).

[21] L.H. Eliasson, *Normal forms for Hamiltonian systems with Poisson commuting integrals – elliptic case.* Comment. Math. Helvetici **65**, 4-35 (1990).

[22] J.P. Dufour and P.Molino, *Compactification d'actions de* $\mathbb{R}^n$ *et variables action–angle avec singularitès,* in *Symplectic Geometry, Groupoids, and Integrable Systems. Séminaire Sud Rhodanien de Géométrie*, P. Dazord and A. Weinstein editors, 151-168 (Springer, New York, 1991).

[23] N. Desoulnex–Moulis, *Singular lagrangian foliation associated to an integrable hamiltonian vector field,* in *Symplectic Geometry, Groupoids, and Integrable Systems. Séminaire Sud Rhodanien de Géométrie*, P. Dazord and A. Weinstein editors, 129-136 (Springer, New York, 1991).

[24] H. Ito, *Action–angle coordinates at singularities for analytic integrable systems*, Math. Z. **206**, 363-407 (1991).

[25] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers* (Oxford University Press, Oxford, 1938, 1979)

[26] H. Poincaré, *Les méthodes nouvelles de la méchanique céleste,* Vols. 1-3 (Gauthier–Villars, Paris, 1892/1893/1899).

[27] H. Goldstein, *Classical Mechanics*, 2^{nd} edition (Addison–Wesley, Reading, 1986)

[28] F. Boigey, *Une théorie des perturbations en variable angles–actions. Application au mouvement d'un solide autour d'un point fixe. Précession–nutation.* J. de Méc. **11**, 521-543 (1972).

[29] A. Deprit, *Free rotation of a rigid body studied in phase plane.* Am. J. Phys. **55**, 424-428 (1967).

[30] H. Andoyer, *Cours de Mécanique Céleste* (Gauthier–Villars, Paris 1923).

[31] J.A. Serret, *Mémoire sur l'emploi de la méthode de la variation des arbitraires dans la théorie des mouvements des rotation.* Mémoires de l'Académie des Sciences de Paris, **35**, 585-616 (1866)

[32] A.S. Mischenko and A.T. Fomenko, *Generalized Liouville method of integration of Hamiltonian systems.* Funct. Anal. Appl. **12**, 113-121 (1978).

[33] P. Dazord and T. Delzant, *Le probleme general des variables actions-angles.* J. Diff. Geom. **26**, 223-251 (1987).

[34] F. Fassò, *The Euler–Poinsot top: a non–commutatively integrable system without global action–angle coordinates.* J. Appl. Math. Phys. (ZAMP) **47**, 953-976 (1996).

[35] M.V. Karasev and V.P. Maslov, *Nonlinear Poisson Brackets. Geometry and Quantization.* Translations of the AMS, Volume 119 (AMS, Providence, Rhode Island, 1993).

[36] A. Weinstein, *The local structure of Poisson manifolds.* J. Diff. Geom. **18**, 523-557 (1983).

[37] M. Born, *The Mechanics of the Atom* (Frederick Ungar Publishing, New York, 1960).

[38] F. Fassò e T. Ratiu, *Compatibility of symplectic structures adapted to noncommutatively integrable systems.* J. Geom. and Phys. **27**, 199-220 (1998).

[39] J. Moser, *Regularization of Kepler's problem and the averaging method on a manifold.* Commun. Pure Appl. Math. , **23**, 609-636 (1970).