Appendix B

Background

In this appendix, we introduce the essential facts on Hamiltonian dynamical systems which will be necessary along the text. First, classical Hamiltonian systems are introduced, then extended through a geometric point of view, using symplectic geometry. Special emphasis is put on the transformation theory, and on the subsequent normal form reduction, because this will be the main tool we shall use along chapter 1.

Most of the theorems –in particular, those concerning with the geometric approach to the mechanics–, are stated without proof. The interested reader may access to a wide amount of literature related with this subject, specially significant for us are the book of Arnol'd (1974) and the one of Abraham and Marsden (1978).

B.1 Hamiltonian systems

Consider first the following system of ordinary first order differential equations on $\mathbb{R}^{2n}$,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{B.1.1}$$

with $i = 1, \dots, n$. The variables $\boldsymbol{q}^* = (q_1, \dots, q_n) \in \mathbb{R}^n$, are the *positions*, while $\boldsymbol{p}^* = (p_1, \dots, p_n) \in \mathbb{R}^n$ are referred as the *momenta*. Both of them, taken together, are often called *coordinates*. The system (B.1.1) is said to be a *Hamiltonian system* of differential equations, while the function H in (B.1.1) is the *Hamiltonian function* or many times, simply, the *Hamiltonian*. The number n (half times the dimension of the space) is the number of *degrees of freedom* of the system.

If now we introduce the notation $\boldsymbol{\zeta}^* = (\boldsymbol{q}^*, \boldsymbol{p}^*)$, identifying $\zeta_i = q_i$ and $\zeta_{i+n} = p_i$ for $i = 1, \dots, n$, the equations (B.1.1) can be written in vectorial form as,

$$\dot{\boldsymbol{\zeta}} = J_n \cdot \mathrm{grad} H(\boldsymbol{\zeta}), \tag{B.1.2}$$

where J_n is the matrix of the standard symplectic form (see definition B.2 and theorem B.12 below):

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \tag{B.1.3}$$

being I_n the $n \times n$ identity matrix. The vector field of the Hamiltonian equations (B.1.2), which we shall note by $X_H = J_n \cdot \mathrm{grad} H$, is the *Hamiltonian vector field* associated to the Hamiltonian H.

Now suppose that we make a change of coordinates given by $\boldsymbol{z} = \Phi(\boldsymbol{\zeta})$, with $\Phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ an smooth map. Then, if $\boldsymbol{\zeta}(t)$ is a solution (B.1.2), $\boldsymbol{z}(t) = \Phi(\boldsymbol{\zeta}(t))$ must satisfy, $\dot{\boldsymbol{z}} = S\dot{\boldsymbol{\zeta}} = SJ\mathrm{grad}_{\boldsymbol{\zeta}} H(\boldsymbol{\zeta}) = SJS^{*}\mathrm{grad}_{\boldsymbol{z}} H\left(\Phi^{-1}(\boldsymbol{z})\right)$, where $S^{i}_{j} = D_j\Phi_i(\boldsymbol{\zeta})$ is the Jacobian matrix of Φ, S^{*} the transpose matrix of $\mathbf{S}$ and Φ^{-1} the inverse of Φ. Therefore, the equations for $\boldsymbol{z}$ are of Hamiltonian type, with a new Hamiltonian function given by $K = H \circ \Phi$, if $SJS^{*} = J$[(1)]. A transformation satisfying this condition is called *canonical* or *symplectic*. If this is the case, the transformed equations can be written explicitly,

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial K}{\partial Q_i},$$

for $i = 1, \ldots, n$ and $\boldsymbol{z} = (\boldsymbol{Q}, \boldsymbol{P})$. It is said that "canonical transformations take Hamiltonian systems into Hamiltonian systems".

The above description corresponds to the classical definition of Hamiltonian systems and symplectic transformations. There, the space of the positions and momenta (the *phase space*) is $\mathbb{R}^n \times \mathbb{R}^n$. For many mechanical systems, though, their natural phase space is not an Euclidean space, but a manifold. For example, the phase space for the motion of a rigid body about a fixed point is $T^{*}SO(3)$, the cotangent bundle of the group of rotations $SO(3)$. See Marsden and Ratiu (1999).

To define a Hamiltonian system on a manifold we need to introduce first several concepts from the symplectic geometry, but before proceeding we fix some useful notation.

Let M be a manifold, then $\mathfrak{F}(M)$ denotes the set of smooth mappings from M into $\mathbb{R}$ and $\mathfrak{X}(M)$, $\Omega^k(M)$ the smooth vector fields and $k-$forms on M respectively. Also, if N is another manifold and $\varphi : M \to N$ is an smooth map, $d\varphi$ stands for the corresponding differential map between the tangent spaces i. e., $d\varphi_m : T_mM \to T_{\varphi(m)}N$, with $m \in M$.

When we have: maps, vector fields, forms, defined on a manifold N, and a regular map $\varphi : M \to N$, from another manifold M to N, there is a standard mechanism to extend these objects into M. This is known as the *pullback* by φ and is usually denoted by φ^{*}.

First, if $\alpha \in \Omega^k(N)$, the pullback of α by φ, $\varphi^{*}\alpha$, is a k-form on M (so $\varphi^{*}\alpha \in \Omega^k(M)$), given by

$$(\varphi^{*}\alpha)_m(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) = \alpha_{\varphi(m)}(d\varphi_m \cdot \boldsymbol{v}_1, \ldots, d\varphi_m \cdot \boldsymbol{v}_k),$$

with $m \in M$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in T_mM$. In the same way we define the pull back of a map $g \in \mathfrak{F}(N)$ by $\varphi^{*}g = g \circ \varphi$ and if φ is a diffeomorphism, the pullback of a vector field $Y \in \mathfrak{X}(N)$ as

$$(\varphi^{*}Y)(m) = (d\varphi_m)^{-1} \cdot Y(\varphi(m)),$$

for any $m \in M$. Note that $\varphi^{*}Y \in \mathfrak{X}(M)$.

If X is a continuous vector field on M, a finite dimensional differentiable manifold, by the theorem of Peano (see Sotomotor, 1979), there exists, for each $m \in M$, an $\varepsilon = \varepsilon(m) > 0$ and a C^1-map $c : (-\varepsilon, \varepsilon) \to M$ such that $c(0) = m$ and $\dot{c}(t) = \frac{d}{dt}c(t) = X(c(t))$, for all $t \in (-\varepsilon, \varepsilon)$. In addition, if the vector field is smooth, i. e., $X \in \mathfrak{X}^r(M)$ (the set of C^r-fields defined on M) with $r \geq 1$, given two C^1 curves on M, $\alpha_i : I_i \to$ such that $c_i(0) = m$, $\dot{c}_i(t) = X(c_i(t))$ for $i = 1, 2$, then $c_1 = c_2$ on $I_1 \cap I_2$.

The maps c with the properties described in this last paragraph, are called *integral curves* of the field X at the point m. The image of an integral curve at m is known as the *orbit* or *trajectory* of the vector field X through the point m.

---

(1) This is equivalent to the more usual condition $S^{*}JS = S$

Now let $\mathfrak{I}_m$ be the set

$$\mathfrak{I}_m = (\omega_-(m), \omega_+(m)) = \bigcup_{I \in \mathfrak{I}} I, \tag{B.1.4}$$

where $\mathfrak{I}$ is the class of open intervals, $I \subset \mathbb{R}$, such that $0 \in I$ and there exists an integral curve $c : I \longrightarrow M$ of the vector field X at m (so $c(0) = m$). $\mathfrak{I}_m$ is the *maximal interval* of definition of the integral curves through m at $t = 0$. Then we can define $c : \mathfrak{I}_m \to M$ of class C^1 and such that $c(0) = m$, $\dot{c}(t) = X\,(c(t))$ for all $t \in \mathfrak{I}_m$. This is the *maximal integral curve* through m at $t = 0$. It is denoted usually by $t \mapsto \phi(t; m) = \phi_t(m) \in M$.

Consider the set $\mathcal{D}_X = \{(t, m) \in \mathbb{R} \times M : t \in \mathfrak{I}_m\}$. The *flow* of X on the manifold M (we denote it again with ϕ), is the map

$$\begin{array}{rcl} \phi : \mathcal{D} \subset \mathbb{R} \times M & \to & M \\ (t, m) & \mapsto & \phi(t; m) = \phi_t(m), \end{array}$$

defined by the properties $\dot{\phi}(t; m) = X\,(\phi(t; m))$ and $\phi(0; m) = m$. Here, as it has been introduced before, the dot symbol denotes the derivative with respect to the parameter (the time) t.

It can be proved (see the book of Sotomayor referenced above), that $\mathcal{D}_X$ is an open set of $\mathbb{R} \times M$ holding $\{0\} \times M$, and also that the flow ϕ is a C^r map if X is a field of class C^r on M.

When $\mathfrak{I}_m = \mathbb{R}$ for all $m \in M$, the field X is said to be *complete*. This takes place, for example when M is a compact manifold (see the book of Palis and Melo, 1982, for a proof). Then, $\phi_t : M \to M$, $t \in \mathbb{R}$, defined by $m \in M \mapsto \phi_t(m) = \phi(t; m) \in M$ is a diffeomorphism and $\phi_t \circ \phi_s = \phi_{t+s}$ for all $t, s \in \mathbb{R}$. Moreover, the set $\{\phi_t\}_{t \in R}$ defines the *one parameter group of diffeomorphisms*.

If X is not complete, the relation $\phi_t\,(\phi_s(m)) = \phi_{t+s}(m)$ or equivalently, $\phi\,(t; \phi(s; m)) = \phi(t + s; m)$ with $m \in M$ holds only whenever both members are defined (that is only if $(s, m), (t, \phi(s, m)) \in \mathcal{D}$). In this case, we say that the flow is *local*.

Remark B.1. In many books, ϕ_t is used also to denote the flow. When t is fixed, and X is complete then it denotes a diffeomorphism $\phi_t : M \to M$. We shall use this convention because both meanings can be usually distinguished from the context. ♠

Let $\alpha \in \Omega^k(M)$, be a k-form on M, $X \in \mathfrak{X}(M)$ and ϕ_t the (local) flow of X. The dynamic definition of the *Lie derivative* of α along X is given by

$$\mathcal{L}_X \alpha = \left.\frac{d}{dt}\right|_{t=0} \phi_t^* \alpha$$

This definition, together with the properties of pullbacks, leads to the *Lie derivative theorem*,

$$\frac{d}{dt} \phi_t^* \alpha = \phi_t^* \mathcal{L}_X \alpha. \tag{B.1.5}$$

(see Abraham, Marsden and Ratiu, 1983, chapter 5).

If $f \in \mathfrak{F}(M)$, the *Lie derivative* of f along X is the *directional derivative*

$$\mathcal{L}_X f = df \cdot X,$$

and the formula (B.1.5) of the Lie derivative theorem has the same expression for functions, i. e.

$$\frac{d}{dt} \phi_t^* f = \phi_t^* \mathcal{L}_X f, \tag{B.1.6}$$

where, as above, ϕ_t is the flow of the vector field X and $f \in \mathfrak{F}(M)$.

For a k-form α on a manifold M and a vector field X, the *interior product* or the *contraction* of X and α, denoted $i_X\alpha$, is defined by

$$(i_X\alpha)_m = \alpha_m\left(X(m), \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}\right),$$

with $m \in M$, and $\boldsymbol{v}_1, \ldots, v_{k-1} \in T_mM$. From this last definition, it is possible to prove that "the pullback of a contraction, $\varphi^* i_X\alpha$, is equal to the contraction of the pullback", i. e.

$$\varphi^* i_X\alpha = i_{\varphi^* X}\varphi^*\alpha. \tag{B.1.7}$$

The Lie derivative and the interior product are related by the *Cartan's magic formula*

$$\mathcal{L}_X\alpha = di_X\alpha + i_Xd\alpha, \tag{B.1.8}$$

where d stands for the *exterior derivative* of the corresponding forms. The reader can find a proof of (B.1.8) in the chapter 6 of the same referred book of Abraham, Marsden and Ratiu (1983).

From the commutation of the pullback and the exterior derivative, i. e., $d\varphi^* = \varphi^* d$ and from the Cartan's magic formula (B.1.8), it is easy to obtain the following relation for the pullback of a Lie derivative

$$\varphi^*\mathcal{L}_X\alpha = \mathcal{L}_{\varphi^* X}\varphi^*\alpha. \tag{B.1.9}$$

($\alpha \in \Omega^k(M)$, $X \in \mathfrak{X}(M)$ and $\varphi : M \to N$, a diffeomorphism). So "the pullback of a Lie derivative is the Lie derivative of the pullback".

For a function $f \in \mathfrak{F}(M)$, and directly from the definitions of pullback and Lie derivative: $\left(\varphi^*\mathcal{L}_X f\right)(m) = \left(\mathcal{L}_X f\right)(\varphi(m)) = df_{\varphi(m)} \cdot X\left(\varphi(m)\right)$, with $m \in M$, but this can be expressed as

$$\begin{aligned} df_{\varphi(m)} \cdot d\varphi_m \cdot d\varphi^{-1}_{\varphi(m)} \cdot X\left(\varphi(m)\right) &= d\left(f \circ \varphi\right)_m \cdot \left(\varphi^* X\right)(m) \\ &= d\left(\varphi^* f\right)_m \cdot \left(\varphi^* X\right)(m) = \left(\mathcal{L}_{\varphi^* X}\varphi^* f\right)(m), \end{aligned}$$

for all $m \in M$. So the same formula (B.1.9) works also for the pullback of a Lie derivative of a function; i. e.

$$\varphi^*\mathcal{L}_X f = \mathcal{L}_{\varphi^* X}\varphi^* f. \tag{B.1.10}$$

A Hamiltonian system is born from a *symplectic structure* defined on a manifold. Next we introduce this and other related concepts.

Definition B.2. *Let M be a regular or smooth manifold. A symplectic form or a symplectic structure, is a two-form ω^2 on M, such that*

(i) *ω^2 is closed: $dw^2 = 0$, and*

(ii) *for each $m \in M$, $\omega^2_m : T_mM \times T_mM \to \mathbb{R}$ is nondegenerate, i. e.: if $\omega^2_m(\boldsymbol{u}, \boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in T_mM$ then $\boldsymbol{u} = 0$.*

The pair (M, ω^2), of a manifold M together with a symplectic form ω^2 on M is a *symplectic manifold*.

Definition B.3. *Let (M, ω^2) be a symplectic manifold, and $H \in \mathfrak{F}(M)$. The vector field X_H determined by the condition*

$$i_{X_H}\omega^2 = dH, \tag{B.1.11}$$

is called the Hamiltonian vector field of the Hamiltonian function H, and (M, ω^2, X_H) is a Hamiltonian system.

Remark B.4. The condition (B.1.11) in the definition above, is equivalent to

$$\omega_m^2\left(X_H(m), \boldsymbol{v}\right) = dH_m \cdot \boldsymbol{v}$$

for all $m \in M$ and $\boldsymbol{v} \in T_m M$. Thus, Nondegeneracy of ω^2 guarantees that X_H exists. ♠

From (B.1.11) it also follows that, on a connected symplectic manifold, any two Hamiltonians for the same X_H have the same differential, so they must differ by a constant.

To each Hamiltonian vector field X_H we associate its Hamilton's equations $\dot{x} = X_H$, whose solutions are integral curves of the field X_H. This is the "natural" extension of the Hamilton's equations (B.1.1) on manifolds.

Example B.5. For $M = \mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$) and if ω^2 is the *standard canonical two form* $\omega^2 = \sum_{i=1}^{n} dq_i \wedge dp_i$ (so $\omega^2(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^* J\boldsymbol{v}$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$)), it is readily checked that in this case $X_H = J_n \cdot \mathrm{grad}H$, and the Hamiltonian equations are the ones given by (B.1.2). ◇

Next we introduce the notion of symplectic map. It just generalizes our early definition of canonical transformation given on page 158 (see proposition B.9 below).

Definition B.6. *A map $\varphi : M \to N$ between symplectic manifolds (M, ω^2) and (N, α^2) is called symplectic if $\varphi^*\alpha^2 = \omega^2$. Note: when $\varphi : M \to M$, then φ is symplectic if $\varphi^*\omega^2 = \omega^2$.*

Example B.7. For $M = \mathbb{R}^{2n}$ and ω^2 the standard symplectic two form given in example B.5, and a diffeomorphism $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$; the just given definition of symplectic map reduces to the condition $S^*JS = J$, for in this case

$$\left(\varphi^*\omega^2\right)_x(\boldsymbol{u}, \boldsymbol{v}) = \omega^2\left(d\varphi_x \cdot u, d\varphi_x \cdot v\right) = \boldsymbol{u}^* S^* JS\boldsymbol{v},$$

where S is the Jacobian matrix of φ i. e., $S_j^i = D_j\varphi_i(x)$, as before. But the last term should be equal to $\omega^2(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^* J\boldsymbol{v}$, for all $x \in \mathbb{R}^{2n}$ and for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{2n}$, so it must be $S^*JS = J$, which is equivalent to our first definition of canonical transformation, $SJS^* = J$,

$$SJS^* = J \Leftrightarrow \left(JS^{-1}\right) SJS^*\left(JS\right) = \left(JS^{-1}\right) J\left(JS\right) \Leftrightarrow S^*JS = J,$$

and the property $J^2 = -I_{2n}$ is used. ◇

The following theorem provides an important result: the flow of a Hamiltonian vector field is, for each fixed time t, a symplectic map which leaves the Hamiltonian function invariant.

Theorem B.8. *Let X_H be a Hamiltonian vector field on the symplectic manifold (M, ω^2), and let ϕ_t be the flow of X_H. Then*

(i) ϕ_t is symplectic; i. e.: $\phi_t^*\omega^2 = \omega^2$,

(ii) H is constant along the the flow, i. e. $H \circ \phi_t = H$.

Proof. The proof of (i) follows immediately from the Lie derivative theorem (B.1.5), and the application of the Cartan's formula

$$\frac{d}{dt}\left(\phi_t^*\omega^2\right) = \phi_t^*\mathcal{L}_{X_H}\omega^2 = \phi_t^*\left(di_{X_H}\omega^2 + i_{X_H}d\omega^2\right),$$

but $d\omega^2 = 0$, because ω^2 is a closed form and $di_{X_H}\omega^2 = d(dH) = 0$. Thus, $\phi_t^*\omega^2 = \phi_{t=0}^*\omega^2 = \omega^2$.

To prove (ii), consider $c(t)$ to be a integral curve of X_H. Then applying the chain rule and the formula (B.1.11) of the definition B.3,

$$\begin{aligned}\frac{d}{dt}(H\circ c)(t) &= dH_{c(t)}\cdot X_H\left(c(t)\right)\\ &= \omega^2_{c(t)}\left(X_H\left(c(t)\right), X_H\left(c(t)\right)\right) = 0.\end{aligned}$$

by the skew-symmetry of ω^2. So H is constant along any integral curve of X_H. In particular along those ones given by the flow. □

The Hamiltonian H is often referred as the *energy* of the Hamiltonian system (M, ω^2, X_H). So (ii) states that the energy is conserved.

Now, we define for any two functions $f, g \in \mathfrak{F}(M)$, their *Poisson bracket* by

$$\{f, g\} = \omega^2(X_f, X_g), \tag{B.1.12}$$

where X_f and X_g are the Hamiltonian vector fields associated to the functions f and g as given by (B.1.11).

The Poisson brackets can also be expressed in terms of the Lie derivative, since

$$\mathcal{L}_{X_f}g = i_{x_f}dg = i_{x_f}i_{x_g}\omega^2 = \omega^2(X_f, X_g) = -\omega^2(X_g, X_f) = -\mathcal{L}_{X_g}f,$$

and therefore

$$\{f, g\} = -\mathcal{L}_{X_f}g = \mathcal{L}_{X_g}f.$$

Note that, for $M = \mathbb{R}^{2n}$, and for the standard symplectic 2-form, the Poisson bracket of two functions f and g can be expressed as

$$\{f, g\} = \sum_{i=1}^{n}\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}, \tag{B.1.13}$$

or, also

$$\{f, g\} = (\operatorname{grad} f)^*\, J\cdot \operatorname{grad} g. \tag{B.1.14}$$

in vectorial notation. Here $(\operatorname{grad} f)^*$ is the transpose of the gradient of f.

Now, let (M, ω^2) and (N, α^2) be symplectic manifolds. A diffeomorphism $\varphi : M \to N$ preserves the Poisson bracket if $\varphi^*\{f, g\} = \{\varphi^* f, \varphi^* g\}$, for all $f, g \in \mathfrak{F}(N)$.

The Poisson bracket gives an straight characterization of the symplectic maps as those which leave it invariant.

Proposition B.9. *With the notation introduced in the paragraph above, the following are equivalent:*

(i) φ *is symplectic.*

(ii) φ *preserves the Poisson bracket of any functions* $f, g \in \mathfrak{F}(N)$.

(iii) $\varphi^* X_f = X_{\varphi^* f}$*, for any* $f \in \mathfrak{F}(N)$.

Proof. First we see that (ii) is equivalent to (iii), since

$$\varphi^* \{f, g\} = \varphi^* \mathcal{L}_{X_g} f = \mathcal{L}_{\varphi^* X_g} \varphi^* f = \{\varphi^* f, \varphi^* g\} .$$

Now consider

$$i_{X_{\varphi^* f}} \omega^2 = d(\varphi^* f) = \varphi^*(df) = \varphi^* i_{X_f} \alpha^2 = i_{\varphi^* X_f} \varphi^* \alpha^2$$

then, by the nondegeneracy of the 2-forms ω^2 and α^2, and by the fact that any $\boldsymbol{v} \in T_m M$ equals to some $X_h(z)$ for a function $h \in \mathfrak{F}(N)$, it follows that φ is symplectic if and only if $\varphi^* X_f = X_{\varphi^* f}$, for all $f \in \mathfrak{F}(N)$. So, it is proved that (i) and (iii) are equivalent. This ends the proof of the proposition. □

Remark B.10. (iii) implies that Hamilton's equations are preserved under canonical transformations. Recall that from our initial definition of symplectic transformation (page 158 and see also example B.7), we saw directly the conservation of the Hamiltonian equations. (iii) adds that the converse is also true. ♠

Proposition B.9 allows to generalize the conservation of the energy established in theorem B.8. We do this through the following corollary.

Corollary B.11 (of proposition B.9). *Consider a Hamiltonian vector field* X_H *on the symplectic manifold* (M, ω^2)*, and let* ϕ_t *be its corresponding flow; then, for any* $f \in \mathfrak{F}(M)$*, we have*

$$\frac{d}{dt}(f \circ \phi_t) = \{f, H\} \circ \phi_t = \{f \circ \phi_t, H\} \tag{B.1.15}$$

Proof. Using formula (B.1.6) of the Lie derivative theorem and the definition of the Poisson bracket in terms of the Lie derivative, $\{g, f\} = \mathcal{L}_{X_f} g$, we obtain

$$\begin{aligned} \frac{d}{dt}(f \circ \phi_t) &= \frac{d}{dt}\phi_t^* f = \phi_t^* \mathcal{L}_{X_H} f \\ &= \phi_t^* \{f, H\} = \{\phi_t^* f, \phi_t^* H\} = \{f \circ \phi_t, H\}, \end{aligned}$$

where the preservation of the Poisson bracket and the Hamiltonian under the flow ϕ_t has been applied. □

A function $g \in \mathfrak{F}(M)$ is in *involution* or *Poisson commute* if $\{g, H\} = 0$. Then, by (B.1.15), g is constant along the flow of the Hamiltonian vector field X_H. This functions are called *integrals* or *constants of the motion*. A classical theorem of Liouville states that, if a Hamiltonian n-degrees of freedom system has k functionally independent integrals in involution, then the number of degrees of freedom can be reduced to $n - k$. When $k = n$ it is said that the system is *integrable*. Then, and under certain additional hypotheses, the trajectories of the system in the phase space are straight lines on high-dimensional cylinders or tori and the Hamiltonian equations can be integrated by quadratures. When the motion takes place on tori, it is possible (though not always trivial) to introduce the so called *action-angle variables*.

For a more precise formulation and for the proof of the Liouville theorem, beyond the outline given here, the reader is aimed to consult the book of Abraham and Marsden (1978) and also the one of Arnol'd (1974). In Goldstein (1980), chapter 10, there are several examples and exercises on integration of mechanical systems by changing to action-angle variables.

The next theorem states that, locally, it is possible to write the Hamilton's equations in the form (B.1.1) of page 157.

Theorem B.12 (Darboux). *Consider the symplectic manifold (M, ω^2) and $m \in M$. Then there exists a chart (U, ψ), with $\psi(m) = 0$; such that*

$$\omega^2\big|_U = \sum_{i=1}^{n} dq_i \wedge dp_i,$$

being $(q_1, \dots, q_n; p_1, \dots, p_n)$ the coordinate functions of the chart map ψ. This coordinates are often referred as symplectic *or* canonical coordinates.

In particular, it follows that the manifold M is even dimensional. The Darboux's theorem allows us to extend, over the whole manifold M, any local result proved for the standard symplectic manifold $(\mathbb{R}^{2n}, \omega^2 = \sum_{i=1}^{n} dq_i \wedge dp_i)$, whenever it is invariant under canonical transformations. For a constructive proof see Arnol'd (1974).

Henceforth we shall consider that the Darboux's theorem has been applied so in the rest of the present monograph we shall work by default with the standard symplectic manifold; so if nothing is stated in the contrary sense, $M = \mathbb{R}^{2n}$ and $\omega^2 = \sum_{i=1}^{n} dq_i \wedge dp_i$.

We also want to stress that all the definitions and theorems presented on these section are valid on *finite dimensional* manifolds. In the different books quoted along the text, it is possible to find generalizations of the corresponding notions and results on infinite dimensional manifolds.

B.2 Poincaré maps

Poincaré maps are a useful trick in the study of dynamical systems. They transform continuous dynamical systems (flows) into discrete (mappings) at the same time that reduce the dimension. The material we include here is taken from Delshams (1994). Also good references are the books of Sotomotor (1979) and Palis and Melo (1982).

Consider $p_0 \in \mathbb{R}^m$ a nonsingular point of a smooth vector field $X \in \mathfrak{X}(R^m)$ (not necessarily Hamiltonian); this means that $X(p_0) \neq 0$. Let $p_1 = \phi(T; p_0)$ with $T \neq 0$ be another point on the integral curve of X at p_0 and Σ_1 a *transversal section* of the field at p_1, i. e., Σ_1 is a smooth $m-1$ hypersurface and $X(p_1) \notin T_{p_1}\Sigma_1$.

We shall suppose that, there exists a regular function $h : U_1 \to \mathbb{R}$, with U_1 a neighborhood of p_1, such that $\Sigma_1 \cap U_1 = \{x \in \mathbb{R}^m : h(x) = 0\}$ and with $Dh(x) \neq 0$ for all $x \in U_1$. Now consider the map $(x, t) \mapsto h(\phi(t; x))$, defined in a neighborhood of p_0. By the implicit function theorem, we have neighborhoods U_0 of p_0, I of T and an unique *time map* $\tau : U_0 \to I$ such that

$$\phi(t; x) \in \Sigma_1 \text{ with } (t, x) \in I \times U_0 \Leftrightarrow t = \tau(x).$$

(in particular, $\tau(p_0) = T$).

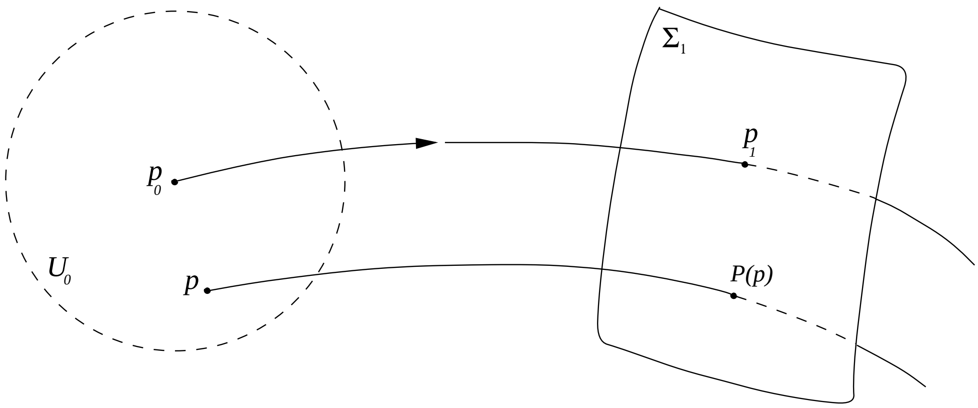


Figure B.1: Poincaré map

The map $x \in U_0 \mapsto P(x) = \phi(\tau(x); x) \in \Sigma_1$ is called the *Poincaré map*. It is a smooth map, though *degenerate* in the following sense: given $p \in U_0$, on the piece of the orbit at p

$$\mathcal{O}_{U_0}(p) = \{x = \phi(t; p) : \phi(s; p) \in U_0 \text{ for all } s \in [0, t_p]\}$$

contained in U_0, the map P is constant, i. e., $P(p) = P(x)$, for all $x \in \mathcal{O}_{U_0}(p)$ (see figure B.2). To avoid this degeneration, we select another transversal section, Σ_0, at p_0 and restrict $P|_{\Sigma_0} : \Sigma_0 \cap U_0 \to P(\Sigma_0 \cap U_0)$ (see figure B.2). In fact, the *restricted* map is the so called Poincaré map. Furthermore, it is not difficult to prove that P is a diffeomorphism.

An interesting case is for periodic orbits, $p_1 = \phi(T; p_0) = p_0$. Then we can take $\Sigma_0 = \Sigma_1$, because $P(\Sigma_0 \cap U_0) \cap (\Sigma_0 \cap U_0) \neq \emptyset$. For this cases we could iterate P and consider P^n, when possible. Note that we have lower in one unit the dimension of the dynamical system, which now is discrete for it is described by a diffeomorphism $P : \Sigma' = \Sigma_0 \cap U_0 \to P(\Sigma') \subset \Sigma_0$.

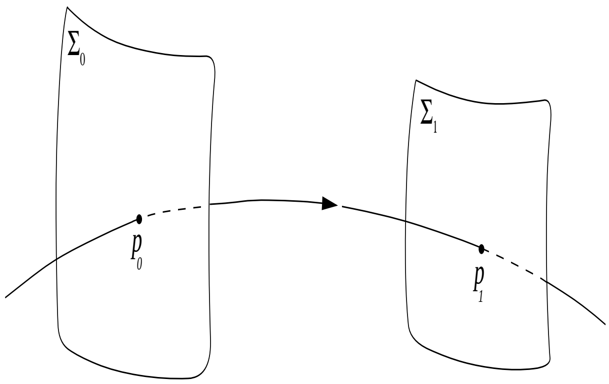


Figure B.2: Transversal section Σ_0 at p_0.

Moreover, the dynamics properties of the flow X are translated to the map P. Thus, periodic points on the map correspond to periodic orbits of X of the same hyperbolic type (see below) and if $\mathcal{A}$ is an invariant set under the flow of X, then the intersection $\Sigma' \cap \mathcal{A}$ is invariant under P, and so on.

For example, it can be seen the relation between the differential matrices DP and $D\phi = \frac{\partial \phi}{\partial x}$ (for the notation, see remark B.14 at the next section).

For a given orbit (not necessarily periodic), consider the Poincaré map $P : \Sigma_0 \to \Sigma_1$ of the field X and the transversal sections Σ_0, Σ_1 defined above. We have $p_1 = \phi(\tau(p_0); p_0)$ and $X(p_0) \neq 0$. Σ_i are transversal sections at p_i so $X(p_i) \notin T_{p_i}\Sigma_i$ with $i = 0, 1$.

On U_0 we can define $P(x) = \phi(\tau(x); x)$ with $\tau(p_0) = p_1$. Consider now the direct sum $\mathbb{R}^m = \operatorname{Span}\{X(p_i)\} \oplus T_{p_i}\Sigma_i$; $i = 0, 1$. To fix ideas, suppose we choose two bases for $\mathbb{R}^m$, say $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$; given by $\boldsymbol{u}_1 = X(p_0)$, $\operatorname{Span}\{\boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\} = T_{p_0}\Sigma_0$ and in the same way, $\boldsymbol{v}_1 = X(p_1)$, $\operatorname{Span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\} = T_{p_1}\Sigma_1$.

Next, we apply the relations (see remark B.14 below):

$$D\phi\left(\tau(p_0);p_0\right)\cdot X(p_0)=X(p_1), \tag{B.2.1}$$

and

$$DP(p_0)=X(p_1)D\tau(p_0)+D\phi\left(\tau(p_0);p_0\right), \tag{B.2.2}$$

to the vectors of the basis $\{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m\}$, to obtain

$$\begin{aligned}D\phi\left(\tau(p_0);p_0\right)\cdot\boldsymbol{u}_1 &= \boldsymbol{v}_1,\\ D\phi\left(\tau(p_0);p_0\right)\cdot\boldsymbol{u}_i &= DP(p_0)\cdot\boldsymbol{u}_i-\left(D\tau(p_0)\cdot\boldsymbol{u}_i\right)X(p_1),\end{aligned}$$

for $i=2,\ldots,m$ in the second equation. But note that $\left(D\tau(p_0)\cdot\boldsymbol{u}_i\right)X(p_1)\in\operatorname{Span}\{X(p_1)\}$ and $DP(p_0)\cdot\boldsymbol{u}_i\in T_{p_1}\Sigma_1$. So the following proposition is proved.

Proposition B.13. *Let $P:\Sigma_0\cap U_0\to\Sigma_1$ be the Poincaré map of the flow of $X\in\mathfrak{X}(\mathbb{R}^m)$, with the transversal sections Σ_0 and Σ_1. Consider any point $p_0\in\Sigma_0\cap U_0$, $p_1=P(p_0)=\phi(T;p_0)$. Then in the bases $\{\boldsymbol{u}_i\}_{i=1,\ldots,m}$, $\{\boldsymbol{v}_i\}_{i=1,\ldots,m}$ described above, the matrices $D\phi(T;p_0)$ and $DP(p_0)$ are related by*

$$D\phi(T;p_0)=\left(\begin{array}{c|c}1 & \alpha\\ \hline 0 & DP(p_0)\end{array}\right),$$

with $\alpha=-D\tau(p_0):T_{p_0}\Sigma\to\mathbb{R}$.

Remark B.14. $D\phi(t;p_0)=\frac{\partial\phi}{\partial x}(t;p_0)$, so it is the derivative with respect to the initial conditions of the flow, and satisfies the following initial value problem:

$$\frac{d}{dt}D\phi(t;p_0)=DX\left(\phi(t;p_0)\right)\,D\phi(t;p_0),\qquad D\phi(0;p_0)=I_m$$

(and I_m is the $m\times m$ identity matrix). Thus, $D\phi(t;p_0)$ is a *fundamental* (in fact the *principal* at $t=0$) matrix of the linear system

$$\dot{x}=A(t)\,x, \tag{B.2.3}$$

with $A(t)=X\left(\phi(t;p_0)\right)$. We say that the equations (B.2.3) are –for the vector field X–, the *first variational equations* of the orbit at p_0. ♠

B.2.1 Poincaré maps of periodic orbits

For a Poincaré map associated to a periodic orbit i. e., when $p_1=\phi(T;p_0)=p_0$, with $X(p_0)\neq 0$ and taking $\Sigma_0=\Sigma_1$, the matrix $A(t)$ in (B.2.3) is T-periodic: $A(t)=A(t+T)$. Then $D\phi(T;p_0)$, is a *monodromy matrix* (since $D\phi(t+T;p_0)=D\phi(t;p_0)\,D\phi(T;p_0)$) and its eigenvalues are called the *characteristic multipliers* of the orbit. Strictly speaking, the characteristic multipliers of a system (B.2.3) with $t\mapsto A(t)$ continuous and T-periodic are defined to be the eigenvalues of *any* monodromy matrix.

In fact, the characteristic multipliers do not depend on the particular monodromy matrix chosen; that is, the particular fundamental solution used to define the monodromy matrix: if $\Psi(t)$ is a fundamental matrix solution with monodromy matrix C (so $\Psi(t+T)=\Psi(t)\,C$)

and $\widetilde{\Psi}(t)$ is another fundamental matrix; then there exists a nonsingular matrix, D, such that $\widetilde{\Psi}(t) = \Psi(t)D$. Hence $\widetilde{\Psi}(t+T) = \widetilde{\Psi}(t)D^{-1}CD$ so the monodromy matrix for $\widetilde{\Psi}(t)$ is $D^{-1}CD$; and similar matrices have the same eigenvalues.

Applying the Floquet theorem, any fundamental matrix $\Psi(t)$ of (B.2.3) can be written as the product of two $m \times m$ matrices

$$\Psi(t) = P(t)\exp(tB), \tag{B.2.4}$$

with $P(t)$ T-periodic and B a constant matrix given by $M = \exp(TB)$, where M is the monodromy matrix of ϕ (so $\Psi(t+T) = \Psi(t)\,M$). As M is not singular, the matrix B exists though it will be complex in general. For a more complete account on Floquet's theorem see Smale (1974), or practically any text book on differential equations.

If λ is a characteristic multiplier (i. e. an eigenvalue of the monodromy matrix M), each complex number μ such that

$$\lambda = e^{T\mu},$$

is called *characteristic* or *Floquet* exponents. Note that the imaginary part of μ is not determined uniquely, since $\mu + i2k\pi/T$ also verifies the above condition for $k \in \mathbb{Z}$. As the characteristic multipliers are determined uniquely, it is usual to choose the characteristic exponents such that they coincide with the eigenvalues of B in (B.2.4).

With this choice, μ_i is a characteristic exponent if and only if $\lambda_i = e^{T\mu_i}$ is a characteristic multiplier, so the solution is asymptotically stable if and only if:

$$\text{Re}\,\mu_i < 0.\ \text{for all } i = 1,\dots,m\ (\Leftrightarrow |\lambda_i| < 1,\ \text{for all } i = 1,\dots,m)$$

It turns out that, if γ is a periodic orbit, its associated Poincaré map $P : \Sigma_0 \cap U_0 \to \Sigma_0$ does not depend upon the selected point p_0, $p_0 \in \gamma$ "up to C^r-conjugations"; more precisely, in the sense of the following.

Proposition B.15. *Let γ be a T-periodic orbit of $X \in \mathfrak{X}^r(M)$ and Σ_0 a transversal section of X at a point $p_0 \in \gamma$. We define the Poincaré map $P_0 : \Sigma_0 \cap U_0 \to \Sigma_0 \cap U_0'$ by $x \mapsto \phi\left(\tau(x); x\right)$, $\tau(p_0) = T$. Then, if P_1 is another Poincaré map associated to γ at another point $p_1 \in \gamma$, there exists a C^r-map $h : W_0 \to W_1$, with W_i a neighborhood of p_i such that $W_i \subset U_i \cap U_i'$, $i = 0, 1$; satisfying $P_1 \circ h = h \circ P_0$ on W_0. Furthermore h is a C^r-diffeomorphism; i. e., there exists a C^r-map $g : W_1 \to W_0$ such that $g = h^{-1}$ on $W_1 \cap \Sigma_1$.*

So, if $A(t) = DX\left(\phi(t; p_0)\right)$ with $\phi(t+T; p_0) = \phi(t; p_0)$, the set

$$\{\lambda \in \mathbb{C},\ \text{characteristic multiplier of } \dot{x} = A(t)\,x\} = \{1\} \cup \text{Spec}\left(DP(p_0)\right),$$

because 1 is always an eigenvalue of $D\phi(T; p_0)$ with eigenvector $X(p_0)$, since

$$D\phi(T; p_0)X(p_0) = X(p_0).$$

Therefore the characteristic multipliers which determine the behavior of the periodic orbit γ are those ones of $DP(p_0)$; and the study of the linear stability of the periodic orbit γ is reduced to the stability of the fixed point p_0 of the Poincaré map P. For example, p_0 (and thus γ) is hyperbolic when $\text{Spec}\left(DP(p_0)\right) \cap S^1 = \emptyset$, i. e., if all the characteristic multipliers, but one lie outside the unit circle in the complex plane.

As through the present work we shall deal with Hamiltonian systems, and they always have at least the integral corresponding to the energy (see theorem B.8), the following proposition becomes of interest.

Proposition B.16. *Let X be a smooth vector field on $\mathbb{R}^m$, $X \in \mathfrak{X}^r(\mathbb{R}^m)$ with $r \geq 1$ and $F : M \to \mathbb{R}$ of class C^1, a first integral of X. Consider the flow $\phi : \mathcal{D} \subset \mathbb{R} \times M \to M$ of the vector field X, a T-periodic orbit $\gamma = \{\phi(t, p_0), t \in [0, T]\}$ and the Poincaré map associated to $p_0 \in \gamma$. Then 1 is an eigenvalue of $DP(p_0)$, and an eigenvalue of $D\phi(T; p_0)$ with multiplicity at least two.*

To prove this proposition we need a previous result

Lemma B.17. *The map*

$$\begin{array}{rcl} W : \mathfrak{J}_{p_0} \times \mathbb{R}^m & \to & \mathbb{R} \\ (t, \boldsymbol{v}) & \mapsto & DF\left(\phi(t; p_0)\right) \boldsymbol{v}, \end{array}$$

is a time-dependent first integral of the variational equation $\dot{x} = A(t)\, x$ associated to the periodic orbit of the vector field X at p_0.

Proof Of the lemma. $\boldsymbol{v}(t)$ is a solution of the variational equation if and only if $\boldsymbol{v}(t) = D\phi(t; p_0)\boldsymbol{v}(0)$. Therefore,

$$W\left(t, \boldsymbol{v}(t)\right) = DF\left(\phi(t; p_0)\right) D\phi(t; p_0)\boldsymbol{v}(0) = DF(p_0)\boldsymbol{v}(0) = W\left(0, \boldsymbol{v}(0)\right),$$

as follows by deriving both sides of $F\left(\phi(t; x)\right) = F(x)$ with respect to x and then substitute $x = p_0$. □

Proof of proposition B.16. By the lemma we have that

$$DF\left(\phi(t; p_0)\right) D\phi(t; p_0) = DF\left(\phi(0; p_0)\right) D\phi(0; p_0) = DF(p_0).$$

Taking $t = T$, (and therefore $\varphi(T; p_0) = p_0$), in the expression above, it reduces to $DF(p_0)D\phi(T; p_0) = DF(p_0)$, or equivalently,

$$\left(D\phi(T; p_0)\right)^* \operatorname{grad} F(p_0) = \operatorname{grad} F(p_0).$$

So $\operatorname{grad} F(p_0)$ is an eigenvector of $\left(D\phi(T; p_0)\right)^*$ associated to an eigenvalue equal to 1. By the other hand, we already know that $D\phi(T; p_0)X(p_0) = X(p_0)$. Also, the relation: $\langle \operatorname{grad} F(p_0), X(p_0)\rangle = DF(p_0)X(p_0) = 0$ must be satisfied since F is a first integral. From here, it follows that the generalized eigenspace $E_1 = \{\boldsymbol{v} \in \mathbb{R}^m : \left(D\phi(T; p_0) - I_m\right)^m = 0\}$, has, at least, dimension 2, i. e., 1 is an eigenvalue of the monodromy matrix with multiplicity at least two. If this is not so, and $\dim E_1 = 1$, then the rest of the eigenvalues should be different from one and therefore, in a suitable basis the monodromy could be expressed as

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A_{m-1} \end{array}\right);$$

thus, in this basis $DF(p_0) = e_1^*$ and $X(p_0) = e_1$ so $DF(p_0) \cdot X(p_0) \neq 0$; which cannot be possible because F is a first integral of the vector field X. □

Without loss of generality, we can suppose $\operatorname{grad} F(p_0) \in T_{p_0}\Sigma_0$; e. g., we can take a surface of section Σ_0 such that, $T_{p_0}\Sigma_0 = X(p_0)^{\perp}$ (see proposition B.15). In this case, taking $\boldsymbol{v}_1 =$

$X(p_0)$, $\boldsymbol{v}_2 = \mathrm{grad}F(p_0)$ and $\boldsymbol{v}_3, \dots, \boldsymbol{v}_m$ such that $\mathrm{Span}\,\{\boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_m\} = T_{p_0}\Sigma_0$. Then in the basis $\{\boldsymbol{v}_i\}_{1\le i\le m}$, the matrix of monodromy reads

$$D\phi(T;p_0) = \left(\begin{array}{c|cccc} 1 & * & * & \cdots & * \\ \hline 0 & 1 & 0 & \cdots & 0 \\ \cline{3-5} \vdots & * & \multicolumn{3}{|c}{} \\ \vdots & \vdots & \multicolumn{3}{|c}{DP_{F_0}(p_0)} \\ 0 & * & \multicolumn{3}{|c}{} \end{array}\right).$$

The box $(m-2)\times(m-2)$ noted as $DP_{F_0}(p_0)$ in the matrix above, corresponds to the Jacobian matrix at the point p_0 of the *reduced Poincaré map* $P_{F_0} : \Sigma_{F_0} \cap U_0 \to \Sigma_{F_0}$, $x \mapsto \phi\left(\tau(x); x\right)$, where

$$\Sigma_{F_0} = \{x \in \Sigma : F(x) = F(p_0) = F_0\}\,, \tag{B.2.5}$$

and it is a $(m-2)$-dimensional manifold of $\mathbb{R}^m$.

If we have $r \ge 1$ first integrals $F_1, \dots, F_r$ with $\mathrm{grad}F_1(p_0), \dots \mathrm{grad}F_r(p_0)$ linear independent, it is often possible to reduce the dynamical system in $r+1$ dimensions and work on

$$\Sigma_{F_1^0,\dots,F_r^0} = \left\{x \in \Sigma : F_j(x) = F_j(p_0) = F_j^0, j = 1, \dots, r\right\}.$$

When this is possible, 1 is an eigenvalue of multiplicity at least $r+1$ of $D\phi\left(T;p_0\right)$ (multiplicity r of $DP(p_0)$) and in such cases, periodic orbits are not isolate, but usually we have *families* of periodic orbits.

B.2.2 Stability for three degrees of freedom Hamiltonian systems

In this section we introduce some notations and conventions for the study of the (linear) stability of Hamiltonian systems with three degrees of freedom. The approach we present here is classical and it can be found mainly in Broucke (1969).

In particular we are interested in the study of (linear) stability of periodic orbits. Since autonomous (i. e., not time dependent) Hamiltonian systems have always a first integral corresponding to the energy, we can fix the energy level $H = h_0$, and work with the reduced Poincaré map $P_{h_0} : \Sigma_{h_0} \to \Sigma_{h_0}$, at some convenient point p_0 on the periodic orbit. Here, the *isoenergetic* surface of section Σ_{h_0} is given by (B.2.5) with $F_0 = h_0 = H(p_0)$. By definition $P_{h_0}(p_0) = p_0$, and the study of the stability of the periodic orbit reduces to the study of the stability of the fixed point p_0 of the map P_{h_0}. Thus, the linear normal behavior of the periodic orbit is determined by the eigenvalues of the differential of the Poincaré map at p_0, $DP_{h_0}(p_0)$.

But P_{h_0} is a symplectic map, so $DP_{h_0}(p_0)$ is a linear symplectic map. The following is a well known result (see Arnol'd, 1974).

Proposition B.18. *The characteristic polynomial of a real symplectic transformation $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$,*

$$p(\lambda) = \det(A - \lambda I_{2n})$$

is cyclic, i. e., $p(\lambda) = \lambda^{2n} p\left(\frac{1}{\lambda}\right)$.

Then, if λ is a eigenvalue of A, $1/\lambda$ must also be an eigenvalue of A. On the other hand, the characteristic polynomial is real, so if λ is a complex eigenvalue, its complex conjugate $\overline{\lambda}$ must also be an eigenvalue. Thus, in a real symplectic map, the eigenvalues appear;

(i) in 4-tuples: λ, $\overline{\lambda}$, $1/\lambda$, $1/\overline{\lambda}$, ($|\lambda| \neq 1$, Im $\lambda \neq 0$),

(ii) pairs on the real axis: $\lambda = \overline{\lambda}$, $1/\lambda = 1/\overline{\lambda}$,

(iii) pairs on the unit circle: $\lambda = 1/\overline{\lambda}$, $\overline{\lambda} = 1/\lambda$.

A linear map A is *stable* if, for any given $\varepsilon > 0$, there exists $\delta > 0$ such that $|x| < \delta \Rightarrow |A^q x| < \varepsilon$, for all $q \in \mathbb{N}$. With this definition, it is straightforward to deduce that if all the $2n$ eigenvalues of a linear symplectic map A, are distinct and lie on the unit circle of the complex plane, then the map is stable. Even more, then, A is proved to be *strongly stable*, which means that any other symplectic map, A_1, "close enough" to A is also stable.

In our particular case, $A = DP_{h_0}(p_0)$ and $n = 2$. Proposition B.18 implies thus that all the possible distributions of the four eigenvalues λ_1, $1/\lambda_1$, λ_2, $1/\lambda_2$ on the complex plane are the ones plotted in figure B.4. Another consequence of the proposition is that the characteristic polynomial may be written in the following form

$$p(\lambda) = \lambda^4 + \alpha\lambda^3 + \beta\lambda^2 + \alpha\lambda + 1 \tag{B.2.6}$$

An important fact is that the stability only depends on the coefficients α and β of this polynomial.

In astronomy and in celestial mechanics, it is usual to define the *stability indices* of the map (and so, of the corresponding orbit) by

$$b_i = \lambda_i + \frac{1}{\lambda_i}, \tag{B.2.7}$$

($i = 1, 2$), and it can be immediately seen that these indices are related with the coefficients of the characteristic polynomial (B.2.6), α, β through

$$\alpha = -(b_1 + b_2), \quad \beta = 2 + b_1 b_2. \tag{B.2.8}$$

From these two relations it follows that the stability indices are given by the solutions of

$$x^2 + \alpha x + (\beta - 2) = 0, \tag{B.2.9}$$

which are,

$$b_{1,2} = \frac{-\alpha \pm \sqrt{\Delta}}{2}$$

where we have introduced the quantity

$$\Delta = \alpha^2 - 4\beta + 8. \tag{B.2.10}$$

By the definition of the stability indices (B.2.7), it is seen immediately that the eigenvalues λ_1, $1/\lambda_1$ are the solutions of the two degree equation

$$x^2 - b_1 x + 1 = 0, \tag{B.2.11}$$

given by

$$\lambda_1, \lambda_1^{-1} = \frac{b_1 \pm \sqrt{b_1^2 - 4}}{2},$$

and in the same manner, the other reciprocal pair λ_2, $1/\lambda_2$ are the solutions of

$$x^2 - b_2 x + 1 = 0, \tag{B.2.12}$$

and given by

$$\lambda_2, \lambda_2^{-1} = \frac{b_2 \pm \sqrt{b_1^2 - 4}}{2}.$$

Then, the problem of finding the roots of the characteristic polynomial (and hence, the eigenvalues) is reduced to solving the three two order equations (B.2.9), (B.2.11) and (B.2.12). The discriminant Δ of (B.2.9) is zero for $\beta = \alpha^2/4 + 2$. In the plane (α, β), this is the equation of a parabola with its apex at $(\alpha, \beta) = (0, 2)$, and the β axis the symmetry axis.

The equations for b_1 and b_2 have zero discriminant along the straight lines $\beta = 2\alpha - 2$ and $\beta = -2\alpha - 2$ respectively. These are tangents to the parabola $\Delta = 0$ at the points $(4, 6)$ and $(-4, 6)$. The parabola, together with these two straight lines bound seven regions in the plane (α, β). They correspond to the seven possible distributions of the eigenvalues with respect the unit circle in the complex plane plotted in the figures B.4(a) to B.4(g).

Thus, periodic orbits (or equivalently, the fixed points of their associated Poincaré maps) can be classified by the position in the *Broucke diagram* (figure B.3 above) of their characteristic polynomial coefficients (α, β) –see (B.2.6)–. Depending upon the region in the diagram the point (α, β) belongs to, the periodic orbit may be stable or present six different types of instability.

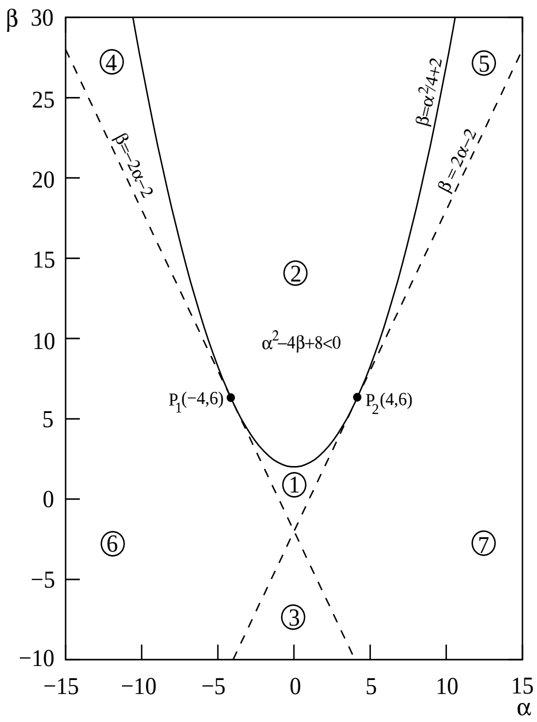


Figure B.3: Broucke Diagram.

Next, we shall briefly describe the relation between the distribution of the (nontrivial) eigenvalues of the periodic orbit with respect the unit circle, and their representation in the Broucke diagram. For a more complete account, see Broucke (1969).

B.2.3 Regions in the Broucke diagram

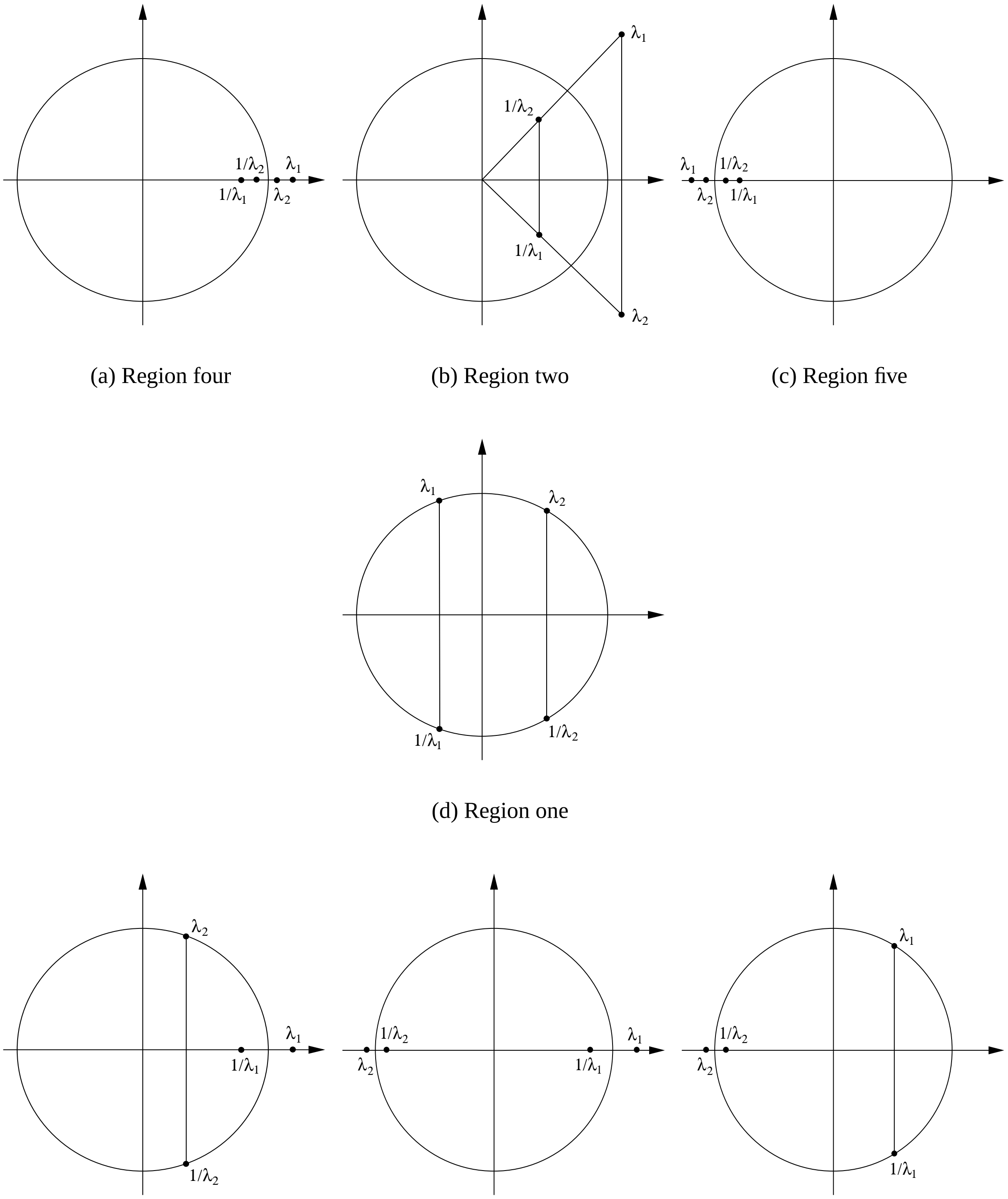
The connection between the seven regions on the diagram of figure B.3 and the seven types of stability depicted in the figure B.4, can be described briefly as follows:

Region 1. $\Delta > 0$ and $b_1, b_2 \in \mathbb{R}$ with $b_1^2 < 4$, $b_2^2 < 4$. Then $\lambda_1, 1/\lambda_1$ are complex conjugates, and so are $\lambda_2, 1/\lambda_2$. The four eigenvalues lie on the unit circle. We have thus *stability*.

Region 2. $\Delta < 0$ and b_1, b_2 are complex conjugates. The four multipliers are complex and lie outside the unit circle. Moreover $\overline{\lambda_1} = \lambda_2$. This type of instability is called *complex instability*.

Region 3. $\Delta > 0$ and $b_1^2 > 4$, $b_2^2 > 4$. All the eigenvalues are real, but $\lambda_1, 1/\lambda_1$ have signs which are opposite to the signs of $\lambda_2, 1/\lambda_2$. This is known as *even-odd instability*.

Region 4. $\Delta > 0$ b_1, b_2 are reals, positive and $b_1^2 > 4$, $b_2^2 > 4$. There are four positive eigenvalues. This is the *even-even instability*.



(a) Region four (b) Region two (c) Region five

(d) Region one

(e) Region six (f) Region three (g) Region seven

Figure B.4: Configuration of the roots with respect to the unit circle on the complex plane for each stability region in the Broucke Diagram.

Region 5. $\Delta > 0$. b_1, b_2 reals and negative with $b_1^2 > 4$ and $b_2^2 > 4$. There are four negative eigenvalues. This is the *odd-odd instability*.

Region 6. $\Delta > 0$. b_1 and b_2 are real with $b_1^2 > 4$ and $b_2^2 < 4$. b_1 is real and positive and $b_1 > 2$ and $\lambda_1, 1/\lambda_1$ are real and positive. The other pair $\lambda_2, 1/\lambda_2$ are complex conjugates and lie on the unit circle. This kind of instability is called *even semi-instability*.

Region 7. $\Delta > 0$. b_1, b_2 are real with $b_1^2 < 4$ and $b_2^2 > 4$. b_2 is real negative and < -2. The eigenvalues $\lambda_2, 1/\lambda_2$ are real and negative. $\lambda_1, 1/\lambda_1$ are complex conjugates on the unit circle. This type of instability is called the *odd semi-instability*.

Remark B.19. For complex instability, $\Delta < 0$, and the stability indices b_1 and b_2 are complex. Therefore it is advisable (see Pfenniger, 1985a) to use instead

$$c_1 = \frac{1}{2}(b_1 + b_2) = \mathrm{Re}\left(\lambda_1 + \frac{1}{\lambda_1}\right) = \frac{\alpha}{2}, \tag{B.2.13}$$

which measures the even or odd character of the complex instability, and

$$c_2 = \frac{1}{2}|b_1 - b_2| = \left|\mathrm{Im}\left(\lambda_1 + \frac{1}{\lambda_1}\right)\right| = \frac{|\Delta|^{\frac{1}{2}}}{2}, \tag{B.2.14}$$

which measures the degree of complex instability. ♠

B.3 The transformation algorithm

In chapter 1, we shall simplify the Hamiltonian function using changes of coordinates. The key point is that such transformations must preserve the structure of the Hamiltonian equations, so they must be canonical or symplectic transformations as defined in section B.1.

A practical way –from a computational point of view–, for generating canonical transformations is based on the fact that the flow of a Hamiltonian system at a fixed time is a symplectic map (see theorem B.8).

Consider a real (or complex) analytic function G, defined on a domain Ω of $\mathbb{R}^{2n}$ (or $\mathbb{C}^{2n}$) and the one-parameter family of transformations $\phi_t^G : \Omega \to \Omega$, $t \in \mathbb{R}$, verifying:

(i) $\phi_0^G = Id$ (the identity map),

(ii) $\frac{d}{dt}\phi_t^G = J_n \mathrm{grad}\, G \circ \phi_t^G$, for all $t \in \mathbb{R}$.

Then, $\{\phi_t^G\}_{t\in\mathbb{R}}$, is the *one-parameter group of symplectic transformations* generated by the function G. The element corresponding to $t = 1$, ϕ_1^G , is the *symplectic transformation generated time-one flow* (of the Hamiltonian G). Therefore, for any analytic $f : \Omega \to \mathbb{R}$ (or $\mathbb{C}$), one may Taylor-expand at $t = 0$ to obtain a *formal* development of the transformed function $f \circ \phi_1^G$, i. e.,

$$f \circ \phi_1^G = f + \frac{d}{dt}(f \circ \phi_t^G)\Big|_{t=0} + \frac{1}{2!}\frac{d^2}{dt^2}(f \circ \phi_t^G)\Big|_{t=0} + \frac{1}{3!}\frac{d^3}{dt^3}(f \circ \phi_t^G)\Big|_{t=0} + \cdots, \tag{B.3.1}$$

Now, on the space of all real (or complex) functions defined on Ω, we define the linear operator: $L_G = \{\cdot, G\}$ and, recursively,

$$\begin{aligned} L_G^0 f &= f, \\ L_G^k f &= L_G\left(L_G^{k-1} f\right), \qquad \text{for } k = 1, 2, \cdots, \end{aligned}$$

then, induction shows that,

$$\frac{d^k}{dt^k}\left(f\circ\phi_t^G\right) = \left(L_G^k f\right)\circ\phi_t^G. \tag{B.3.2}$$

Thus, the above expansion of $f\circ\phi_1^G$ can be expressed as next lemma shows.

Lemma B.20. *Under the conditions specified above on the functions f and G, the transformed of the function f through the time-one flow ϕ_1^G $f\circ\phi_1^G$, can be cast into*

$$f\circ\phi_1^G = \sum_{k=0}^{r-1}\frac{1}{k!}L_G^k f + \int_0^1 \frac{(1-t)^{r-1}}{(r-1)!}\left(L_G^r f\right)\circ\phi_t^G\,dt,$$

for any $r = 1, 2, \ldots$

Proof. If follows from the Taylor expansion (B.3.1), up to order $r-1$, adding the remainder in its integral form and with the time derivatives substituted by (B.3.2). □

Remark B.21. Note that, in particular applying the above lemma to the coordinate functions (i. e., taking $f = x_i$, for $i = 1, \ldots, 2n$) one determines the components of ϕ_1^G itself. ♠

The method just described is the most elementary version of the so called *Lie series methods* to generate canonical transformations. These are particularly well-suited for mechanized treatment, since they only require computation of Poisson brackets. The interested reader can find, in Jorba (1999), a practical implementation of a specialized software package able to deal with these questions.

Corollary B.22. *With the same assumptions of lemma B.20, but accept $f = f_0 + f_1$, then*

$$f\circ\phi_1^G = f + \{f - f_1, G\} + \int_0^1 \left(\{f_1, G\} + (1-t)\left\{\{f - f_1, G\}, G\right\}\right)\circ\phi_t^G\,dt \tag{B.3.3}$$

Proof. Directly from the lemma for $r = 2$, setting $\{f, G\} = \{f - f_1, G\} + \{f_1, G\}$ and taking into account that,

$$\{f_1, G\} = \int_0^1 \frac{d}{dt}\left((t-1)\left(L_G f_1\right)\circ\phi_t^G\right)\,dt = \int_0^1 \left(L_G f_1 + (t-1)L_G^2 f_1\right)\circ\phi_t^G\,dt,$$

one arrives to the result of the lemma. □

Formula (B.3.3) will be used in chapter 3 to derive bounds for norm of the the “bad terms” –i. e., those avoiding certain kind of solutions–, in the Lie-transformed Hamiltonians appearing along the iterative steps of the KAM method. The function G is the *generating function* or the *generator* of the transformation. In our context, the function to transform will be an initially given Hamiltonian, $H^{(0)}$. Therefore, one may ask for G such that the transformed new Hamiltonian $H^{(1)} = H^{(0)}\circ\phi_1^G$, will be –tied to some predefined criteria–, simpler than the initial one. This is, essentially, the idea on which the *normal form* and *normalizing transformation* computations are based upon. The aim of the example below is just to illustrate these concepts.

Example B.23. Let $H(\boldsymbol{\xi},\boldsymbol{\eta})$ be an n degree of freedom real analytic Hamiltonian, so $\boldsymbol{\xi}^* = (\xi_1,\dots,\xi_n)$ and $\boldsymbol{\eta}^* = (\eta_1,\dots,\eta_n)$. Furthermore, we suppose that H can be expanded as $H = H_2 + H_3 + \dots + H_k + \dots$, where H_k, is an homogeneous polynomial of degree $k > 2$ in the variables $(\boldsymbol{\xi},\boldsymbol{\eta}) \in \mathbb{C}^{2n}$. So

$$H_k(\boldsymbol{\xi},\boldsymbol{\eta}) = \sum_{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=k} h_{\boldsymbol{l},\boldsymbol{m}}\,\boldsymbol{\xi}^{\boldsymbol{l}}\boldsymbol{\eta}^{\boldsymbol{m}}, \tag{B.3.4}$$

and the following standard notation is used: $\boldsymbol{u}^{\boldsymbol{l}} = \prod_{j=1}^n u_j^{l_j}$, $\boldsymbol{u} \in \mathbb{C}^n$, while $|\cdot|_1$ denotes the norm $|\boldsymbol{r}|_1 = |r_1| + \cdots + |r_n|$.

From the development (B.3.4), it follows that $(\boldsymbol{\xi},\boldsymbol{\eta}) = (\mathbf{0},\mathbf{0})$ is an equilibrium point of the Hamiltonian system (since $\operatorname{grad} H(\mathbf{0},\mathbf{0}) = \mathbf{0}$). In order to simplify, we shall suppose that this equilibrium point is non-degenerate *elliptic*; i. e., that $\operatorname{Spec}(J\,D^2H(\mathbf{0},\mathbf{0})) = \{\mathrm{i}\omega_1,\dots,\mathrm{i}\omega_n\}$, with $w_j \in \mathbb{R}$, $\omega_j \neq \omega_k$ for $j \neq k$ and $\mathrm{i} = \sqrt{-1}$. In such cases (see Arnol'd, 1974), by means of a real linear canonical transformation, $\binom{\boldsymbol{x}}{\boldsymbol{y}} = S\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$ with $S^*JS = J$, the initial Hamiltonian may be transformed to its real linear normal form,

$$H'(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\sum_{j=1}^n \omega_j\left(x_j^2 + y_j^2\right) + H_3'(\boldsymbol{x},\boldsymbol{y}) + H_4'(\boldsymbol{x},\boldsymbol{y}) + \dots \tag{B.3.5}$$

Moreover, to get simpler *homological equations* (see below), it is useful to introduce the following (complex) linear symplectic change,

$$x_j = \frac{q_j + ip_j}{\sqrt{2}}, \qquad y_j = \frac{iq_j + p_j}{\sqrt{2}}, \tag{B.3.6}$$

$j = 1,\dots,n$. With this complex change, the Hamiltonian (B.3.5) transforms to,

$$H^{(0)}(\boldsymbol{q},\boldsymbol{p}) = \sum_{j=1}^n i\omega_j q_j p_j + \sum_{k>2} H_k^{(0)}(\boldsymbol{q},\boldsymbol{p}), \tag{B.3.7}$$

where $H_k^{(0)}(\boldsymbol{q},\boldsymbol{p})$, as in (B.3.4) are homogeneous polynomials of degree k in $\boldsymbol{q},\boldsymbol{p}$; so, as before, we write:

$$H_k^{(0)}(\boldsymbol{q},\boldsymbol{p}) = \sum_{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=k} h_{\boldsymbol{l},\boldsymbol{m}}^{(0)}\,\boldsymbol{q}^{\boldsymbol{l}}\boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.8}$$

and $H_2^{(0)}$ will stand for the quadratic part of $H^{(0)}$, i. e.,

$$H_2^{(0)}(\boldsymbol{q},\boldsymbol{p}) = \sum_{j=1}^n i\omega_j\, q_j p_j.$$

The Hamiltonian (B.3.7) is known as the *complexified Hamiltonian*. An explicit description of the linear normalization process which leads to $H^{(0)}$ can be found in Siegel and Moser (1971), chap. 2, § 15.

The inverse of the change (B.3.6) is

$$q_j = \frac{x_j - iy_j}{\sqrt{2}}, \quad p_j = \frac{y_j - ix_j}{\sqrt{2}} \tag{B.3.9}$$

so, when x_j, y_j are real, the complex conjugates of the complex positions $\boldsymbol{q}$ and momenta $\boldsymbol{p}$ satisfy, $\overline{q}_j = -ip_j$, $\overline{p}_j = -iq_j$ for $j = 1, \ldots, n$. This induces the following symmetries on the coefficients in (B.3.8),

$$\overline{h}^{(0)}_{\boldsymbol{l},\boldsymbol{m}} = i^{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1} h^{(0)}_{\boldsymbol{m},\boldsymbol{l}}. \tag{B.3.10}$$

Remark B.24. Before continuing with the nonlinear normalization of the Hamiltonian, it is worth mentioning two essential properties of the Poisson brackets.

P1. If f and g are homogeneous polynomials of degrees r and s respectively, the degree of their Poisson bracket is $\deg\{f, g\} = r + s - 2$, and

P2. if the expansions of f and g satisfy the symmetry (B.3.10), so does $\{f, g\}$.

By this last property, it is assured that, after the nonlinear reduction process, the change in (B.3.9) will transform the final Hamiltonian into a real analytic one. ♠

In principle, it is possible to "remove" (in the sense we specify below) the monomials on $H^{(0)}_3$ by taking a generating function $G_3(\boldsymbol{q}, \boldsymbol{p}) = \sum g_{\boldsymbol{l},\boldsymbol{m}} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}$, with the coefficients $g_{\boldsymbol{l},\boldsymbol{m}}$ ($|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3$), suitably chosen. First, by (B.20) the transformed Hamiltonian $H^{(1)} = H^{(0)} \circ \phi_1^{G_3}$ will be given by

$$H^{(1)} = H^{(0)}_2 + \{H^{(0)}_2, G_3\} + H^{(0)}_3 + \ldots \tag{B.3.11}$$

Note that, by the first of the properties on the remark B.24, the dots hold terms of degree greater than 3, i. e., there are no more terms of degree 3 than the sum: $\{H^{(0)}_2, G_3\} + H^{(0)}_3$. Ideally, we ask $H^{(1)}$ not to contain terms of degree 3 (in this sense we want to *remove* these terms). Hence, G_3 should satisfy the following *homological equations*:

$$\{H^{(0)}_2, G_3\} + H^{(0)}_3 = 0.$$

Let $\mathfrak{A}_k$ be the space of (complex) homogeneous polynomials of degree k in the variables $(\boldsymbol{q}, \boldsymbol{p}) \in \mathbb{C}^{2n}$ and define the operator $L_{H^{(0)}_2} = \{\cdot, H^{(0)}_2\}$. More precisely,

$$\begin{aligned} L_{H^{(0)}_2} : \mathfrak{A}_k &\longrightarrow \mathfrak{A}_k \\ f &\longmapsto L_{H^{(0)}_2} f = \{f, H^{(0)}_2\}, \end{aligned}$$

(we abbreviate $L_{H^{(0)}_2} = L$ in the text).

Remark B.25. Note that, by (31) $L_{H^{(0)}} f = \mathcal{L}_{X_{H^{(0)}}} f$, where $X_{H^{(0)}}$ is the Hamiltonian vector field associated to the function $H^{(0)}$. ♠

With this operator, the homological equations above, may be written as,

$$L_{H^{(0)}_2} G_3 = H^{(0)}_3$$

but for this equations to be compatible, it is necessary that $H^{(0)}_3 \in \text{Range}\,(L)$. This does not happen, in general, due to the presence of *resonant monomials* in $H^{(0)}_3$. A monomial $f = a_{\boldsymbol{l},\boldsymbol{m}} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}$, $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = k$, is said to be *resonant*, if $Lf = 0$ (equivalently, if $f \in \text{Ker}\,(L)$). It is thus necessary to add a compatibility term, $Z_3 \in \text{Ker}(L)$, satisfying $H^{(0)}_3 - Z_3 \in \text{Range}\,(L)$.

Therefore, the homological equations (B.23) for the degree $k=3$; must be completed in the form,

$$L_{H_2^{(0)}} G_3 + Z_3 = H_3^{(0)}. \tag{B.3.12}$$

We define the resonance modulus associated to $\boldsymbol{\omega}^* = (\omega_1, \ldots, \omega_n)$, the vector of the frequencies, as $\mathfrak{R} = \{\boldsymbol{r} \in \mathbb{Z}^n : \langle \boldsymbol{r}, \boldsymbol{\omega} \rangle = 0\}$ (the angular brackets $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{j=1}^n u_j v_j$, will be used for the ordinary scalar product). Thus, it becomes clear how to express the compatibility term Z_3,

$$Z_3(\boldsymbol{q}, \boldsymbol{p}) = \sum_{\substack{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3 \\ \boldsymbol{m} - \boldsymbol{l} \in \mathfrak{R}}} h_{\boldsymbol{l},\boldsymbol{m}}^{(0)} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.13}$$

When the homological equations (B.3.12) are written down explicitly, using the developments of G_3, $H_3^{(0)}$ and Z_3, and computing the Poisson bracket, on can realize that

$$\sum_{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3} \mathrm{i} \langle \boldsymbol{\omega}, \boldsymbol{m} - \boldsymbol{l} \rangle g_{\boldsymbol{l},\boldsymbol{m}} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}} = \sum_{\substack{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3 \\ \boldsymbol{m} - \boldsymbol{l} \notin \mathfrak{R}}} -h_{\boldsymbol{l},\boldsymbol{m}}^{(0)} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}.$$

This gives, for the unknowns $g_{\boldsymbol{l},\boldsymbol{m}}$, an algebraic linear diagonal system in the space $\mathbb{C}^d \simeq \mathfrak{A}_3$, $d = \binom{2n+2}{3}$; so, easily, one obtains

$$g_{\boldsymbol{l},\boldsymbol{m}} = \frac{-h_{\boldsymbol{l},\boldsymbol{m}}^{(0)}}{\mathrm{i} \langle \boldsymbol{\omega}, \boldsymbol{m} - \boldsymbol{l} \rangle}; \tag{B.3.14}$$

with $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3$ and indeed, $\boldsymbol{l} - \boldsymbol{m} \notin \mathfrak{R}$, so the divisors do not vanish. With this choice for the coefficients in G_3, the transformed Hamiltonian, does not hold three degree terms, but the resonant ones, contained in Z_3. Therefore,

$$H^{(1)} = H^{(0)} \circ \phi_1^{G_3} = H_2 + Z_3 + H_4^{(1)} + \ldots, \tag{B.3.15}$$

where $H_4^{(1)} + \ldots$ denotes the transformed terms of degree greater than three.

The general step

In the previous paragraphs, we have given an account of the linear reduction and some details of the first step in the nonlinear –or in the “normal form”–, reduction process of the initial Hamiltonian. Suppose now that the same has been repeated up to degree $k > 3$. So at the k-th step we have a Hamiltonian $H^{(k)} = H^{(0)} \circ \phi_1^{G_3} \circ \cdots \circ \phi_1^{G_{k+2}}$,

$$H^{(k)} = H_2^{(0)} + Z_3 + \ldots + Z_{k+2} + \sum_{s > k+2} H_s^{(k)}, \tag{B.3.16}$$

with $H_s^{(k)}$ homogeneous polynomials of degree $s > k$, in $(\boldsymbol{q}, \boldsymbol{p})$,

$$H_s^{(k)}(\boldsymbol{q}, \boldsymbol{p}) = \sum_{|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = s} h_{\boldsymbol{l},\boldsymbol{m}}^{(k)} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}.$$

In the $(k+1)$-th step we want to remove the terms of degree $k+3$ of $H^{(k)}$. We apply the change $\phi_1^{G_{k+3}}$ to obtain a new Hamiltonian, $H^{(k+1)} = H^{(k)} \circ \phi_1^{G_{k+3}}$. By the Lie transform formula (B.20), and writing only up to degree $k+3$,

$$H^{(k+1)} = H_2^{(0)} + Z_3 + \cdots + Z_{k+3} + \{H_2^{(0)}, G_{k+3}\} + H_{k+3}^{(k)} + \ldots,$$

for, again –with (B.20) in mind–, it is just a check on the degrees, to realize that the only terms of degree $k+3$ present in this new Hamiltonian are those in the sum $\{H_2^{(0)}, G_{k+3}\} + H_{k+3}^{(k)}$. In consequence, the homological equations will be,

$$L_{H_2^{(0)}} G_{k+3} + Z_{k+3} = H_{k+3}^{(k)}, \tag{B.3.17}$$

with, in the same way as for $k=3$, taking Z_k an homogeneous polynomial holding the resonant terms of $H_{k+3}^{(k)}$,

$$Z_{k+3}(\boldsymbol{q},\boldsymbol{p}) = \sum_{\substack{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1=k+3 \\ \boldsymbol{m}-\boldsymbol{l}\in\mathfrak{R}}} h_{\boldsymbol{l},\boldsymbol{m}}^{(k)} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{m}}, \tag{B.3.18}$$

The solutions of (B.3.17) will have, of course, the same form of (B.3.14),

$$g_{l,m} = \frac{-h_{\boldsymbol{l},\boldsymbol{m}}^{(k)}}{\mathrm{i}\langle \boldsymbol{\omega}, \boldsymbol{m}-\boldsymbol{l}\rangle}; \tag{B.3.19}$$

now, with $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = k+3$, but identically: $\boldsymbol{l}-\boldsymbol{m} \notin \mathfrak{R}$. Therefore, if we make r steps of the nonlinear reduction process, the resulting Hamiltonian splits into

$$H^{(r)} = Z^{(r)} + R^{(r)}, \tag{B.3.20}$$

where the *remainder* $R^{(r)}$ contains terms of degree $> r+2$, while

$$Z^{(r)} = Z_2 + Z_3 + \ldots + Z_{r+2},$$

(with $Z_2 = H_2^{(0)}$), is the (complex) *normal form* up to order $r+2$ of the Hamiltonian. It holds the quadratic part plus the resonant terms $Z_3, \ldots, Z_{r+2}$ from degree 3 up to $r+2$. Nevertheless, sometimes, the term *normal form* applies, by extension, to the whole transformed Hamiltonian (B.3.20).

Remark B.26. From (B.3.14) it follows that the generating functions coefficients $g_{\boldsymbol{l},\boldsymbol{m}}$, $|\boldsymbol{l}|_1 + |\boldsymbol{m}|_1 = 3$, satisfy the symmetries (B.3.10). By Induction and taking into account the second property of remark (B.24), it is possible to see that, at any degree $3 \le k \le r+2$, $g_{\boldsymbol{l},\boldsymbol{m}} = i^{|\boldsymbol{l}|_1+|\boldsymbol{m}|_1} g_{\boldsymbol{m},\boldsymbol{l}}$. Hence, the same is true for the coefficients of $H^{(r)}$. This implies that we can obtain a *real* normalized Hamiltonian, applying to (B.3.20) the change (B.3.9). ♠

Even in the simplest case, when $\mathfrak{R} = \emptyset$, there appear *inevitable* resonances when $\boldsymbol{l} = \boldsymbol{m}$. Then, the terms in $Z^{(r)}$ take the form,

$$\left.\begin{array}{rcl} Z_{2s}(\boldsymbol{q},\boldsymbol{p}) & = & \sum_{|\boldsymbol{l}|_1=s} h_{\boldsymbol{l},\boldsymbol{l}}^{(2s-3)} \boldsymbol{q}^{\boldsymbol{l}} \boldsymbol{p}^{\boldsymbol{l}}, \\ Z_{2s-1} & \equiv & 0, \end{array}\right. \qquad 2 \le s \le \lfloor r/2 \rfloor + 1,$$

where $\lfloor x \rfloor$ denotes the greatest integer function of $x \in \mathbb{R}$ (i. e, $\lfloor x \rfloor := \max\{z \in \mathbb{Z} : z \le x\}$). If –as pointed in last remark–, we apply first the linear symplectic change (B.3.9), and then introduce polar canonical coordinates in the form,

$$x_j = \sqrt{2I_j}\cos\theta_j, \qquad y_j = -\sqrt{2I_j}\sin\theta_j \tag{B.3.21}$$

($j = 1, \dots, n$), we obtain a real normalized Hamiltonian,

$$\widetilde{H}^{(r)}(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{j=1}^{n} \omega_j I_j + \sum_{j=1}^{\lfloor r/2 \rfloor + 1} \widetilde{Z}_{2j}(\boldsymbol{I}) + \widetilde{R}^{(r)}(\boldsymbol{I}, \boldsymbol{\theta}),$$

with $\boldsymbol{I} = (I_1, \dots, I_n)$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$, and $\widetilde{Z}_{2j}(\boldsymbol{I})$ are real homogeneous polynomials of degree j in $I_1, \dots, I_n$.

An interesting point is that the normal form, i. e., the first two sums in the expression above,

$$\widetilde{Z}^{(r)} = \sum_{j=1}^{n} \omega_j I_j + \sum_{j=1}^{[r/2]+1} \widetilde{Z}_{2j}, \tag{B.3.22}$$

does not depend on the angular variables θ_j, so if we skip the remainder off and consider the Hamiltonian system given by (B.3.22),

$$\dot{\theta}_j = \frac{\partial \widetilde{Z}^{(r)}}{\partial I_j}, \qquad \dot{I}_j = 0, \tag{B.3.23}$$

for $j = 1. \dots, n$, this system is immediately integrable, with solutions:

$$I_j = I_j^0 = \text{contant}, \quad \theta_j = \Omega_j(\boldsymbol{I}^0)\, t + \theta_j^0, \tag{B.3.24}$$

with the frequencies, $\Omega_j = \frac{\partial \widetilde{Z}^{(r)}}{\partial I_j} = \omega_j + \dots ; j = 1, \dots, n$. Therefore, the trajectory of the phase point $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ winds an n dimensional invariant torus defined by the first integrals $I_j = I_j^0 = \frac{1}{2}(x_j^2 + y_j^2)$. ◇

Normal forms around equilibrium points were studied by Birkhoff (1927). In fact, in the texts, the normal form (B.3.20) is known as the Birkhoff's normal form. In the description left here, it has been obtained applying successive canonical changes, each of them constructed to remove the terms one degree higher than the preceding one. The final transformation is a product of the r transformations: $\Psi^{(r)} = \phi_1^{G_3} \circ \phi_1^{G_4} \circ \cdots \circ \phi_1^{G_{r+2}}$.

There is, however, a vast literature for *Lie transformation algorithms*. See Deprit (1969) and the extensions of Kamel (1970); Henrard (1970a,b,c) for non Hamiltonian systems; or also the books of Chow and Hale (1982), chapter 12 and Meyer and Hall (1992) chapter 7.

Without digging deeper into the details: given a generating function G, which can be expanded as a sum $G = \sum_{k \geq 3} G_k$, there are algorithms which allows us to construct a canonical transformation T_G, such that if f is a function of $(\boldsymbol{q}, \boldsymbol{p})$, the transformed function $T_G f$ will be defined by

$$T_G f = \sum_{k \geq 1} F_k,$$

where the terms f_k are obtained recursively, that is, beginning with $F_1 = f_1$, for $k > 1$ is $F_k = F_k(G_1, \dots, G_{k+1}; f_0, \dots, f_k)$, so each term in the sum can be obtained from the generating function and the previous computed terms.

In particular in chapters 1 and 2 we shall use an algorithm of this type: the Giorgilli–Galgani algorithm, (see references there). With this short outline of the transformation theory, we close this appendix.