Quasi-integrable mechanical systems*

Giovanni Gallavotti

sec.1 1. Basic definitions on integrability and canonical integrability. Examples.

A mechanical system (H, W) with Hamiltonian H on the phase space region W is called *integrable* if all motions in W are quasi-periodic in the following sense. There is a system of coordinates $(\mathbf{A}, \boldsymbol{\varphi})$ on W in which W appears as $\mathcal{V} \times \mathbb{T}^\ell$, where $\mathcal{V}$ is some open set in $\mathbb{R}^\ell$ and $\mathbb{T}^\ell$ is the ℓ–dimensional torus and, furthermore, the motion in these coordinates has the form $(\mathbf{A}, \boldsymbol{\varphi}) \to (\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{A})\, t)$, where $\boldsymbol{\omega}(\mathbf{A})$ is analytic in $\mathcal{V}$.

More formally (see Ref. [1], p.287): there is a map $\mathcal{I} :\ W \to \mathcal{V} \times \mathbb{T}^\ell$ analytic, invertible with non singular Jacobian and such that, if S_t denotes the solution map for the system,

1.1 $$S_t(\mathbf{p}, \mathbf{q}) \equiv\equiv S_t(\mathcal{I}(\mathbf{A}, \boldsymbol{\varphi})) = \mathcal{I}(\mathbf{A}, \boldsymbol{\varphi} + \boldsymbol{\omega}(\mathbf{a})\, t) \tag{1.1}$$

where $\boldsymbol{\omega}$ is an analytic function on $\mathcal{V}$.

Of course, if a system is integrable then its phase space is foliated into ℓ–dimensional invariant tori and the motions on them are quasi periodic.

Another notion of integrability is that of *canonical integrability*, arising when the map $\mathcal{I}$ integrating the system can be chosen to be a completely canonical map $\mathcal{C}$: since completely canonical maps transform solutions of the Hamilton equations into solutions of Hamilton equations with the same Hamiltonian (computed in the new coordinates) it follows that a system is canonically integrable of and only if there is a completely canonical map $\mathcal{C} :\ W \longleftrightarrow \mathcal{V} \times \mathbb{T}^\ell$ such that

1.2 $$H(\mathcal{C}^{-1}(\mathbf{A}, \boldsymbol{\varphi})) = h(\mathbf{A}), \qquad (\mathbf{A}, \boldsymbol{\varphi}) \in V \times \mathbb{T}^\ell, \tag{1.2}$$

for a suitable analytic function h.

In fact (1.2) implies that in the new coordinates the equation of motion become:

1.3 $$\dot{\mathbf{A}} = \mathbf{0}, \qquad \dot{\boldsymbol{\varphi}} = \partial_{\mathbf{A}} h(\mathbf{A}) = \boldsymbol{\omega}(\mathbf{A}) \tag{1.3}$$

It is convenient to think of $\mathbb{T}^\ell$ as a product of ℓ circles: the positions on them are described by ℓ angles $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_\ell)$. Since two Hamiltonian systems related by a completely canonical trans-

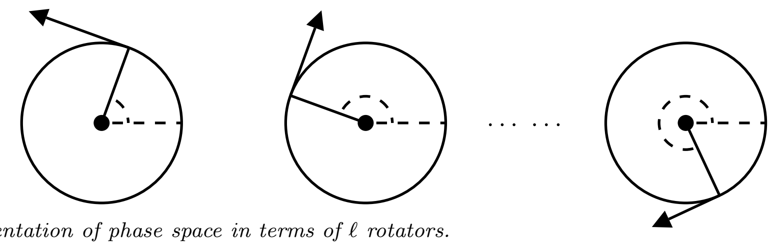


Fig.1 *A representation of phase space in terms of ℓ rotators.*

formation are called *conjugate*, one can interpret canonical integrability by saying that a system is conjugate to a system of *free rotators*; its phase space can then be represented as in Fig. 1.

---

* In *Phénomènes critiques, systèmes aléatoires, théories de jauge*, Lectures at the XLIII summer school in Les Houches, 1984, p. 541-624, Ed. K. Osterwalder, R. Stora, Elseviers.

The phase space foliation into invariant tori parameterized by $\mathbf{A} \in \mathcal{V}$ is described by the equations for the torus $\mathbb{T}_\mathbf{A}$:

1.4 $$(\mathbf{p}, \mathbf{q}) = \mathcal{C}^{-1}(\mathbf{A}, \varphi), \qquad \varphi \in \mathbb{T}^\ell \tag{1.4}$$

It is perhaps useful to recall the notion of canonical map. If W, W' are open subsets of two phase spaces, an analytic change of coordinates

1.5 $$\mathcal{C}(\mathbf{p}, \mathbf{q}) = (\mathbf{p}', \mathbf{q}') \tag{1.5}$$

mapping W onto W' is called *completely canonical and action preserving* (see, for instance. Ref. [1], p.289) if one can define an analytic function Φ on the graph $G(\mathcal{C}) = \big\{\mathbf{p}, \mathbf{q}, \mathbf{p}', \mathbf{q}' | (\mathbf{p}', \mathbf{q}') = \mathcal{C}(\mathbf{p}, \mathbf{q})\big\} \subset W \times W'$ such that

1.6 $$\mathbf{p} \cdot d\mathbf{q} = \mathbf{p}' \cdot d\mathbf{q}' + d\Phi \qquad \text{on} \quad G(\mathcal{C}) \tag{1.6}$$

Remark: If $\mathbf{p} \cdot d\mathbf{q} - \mathbf{p}' \cdot d\mathbf{q}'$ is only locally integrable on $G(\mathcal{C})$ one says that $\mathcal{C}$ is *completey canonical*, only. For an example see Appendix F in Ref. [2]. However here I shall deal only with completely canonical action preserving maps calling them, simply, completely canonical as no confusion can arise between the two notions.

To avoid misunderstandings let me say explicitly that a *subset of phase space* is a subset of the cotangent bundle $T^*\Sigma$ to some analytic boundaryless surface of dimension ℓ and (1.5) descriibes the map $\mathcal{C}$ in two given local systems of coordinates.

It is easy to see that if $\mathcal{C}$ satisfies (1.6) it is completely canonical in the sense that, given any Hamiltonian H on W abd defining $H'(\mathbf{p}', \mathbf{q}')$ the function $H'(\mathbf{p}', \mathbf{q}') = H(\mathcal{C}^{-1}(\mathbf{p}', \mathbf{q}')$ the Hamiltonian otions of H on W are mapped by $\mathcal{C}$ onto the Hamiltonian motions of H' on W'; this follows immediately from Hamilton's principle whereby the Hamiltonian motions of (H, W) are those which make stationary the functional

1.7 $$\mathcal{A} = \int_{t_1}^{t_2} \big(\mathbf{p} \cdot d\mathbf{q} - H(\mathbf{p}, \mathbf{q}) \tag{1.7}$$

among all motions $t \to (\mathbf{p}(t), \mathbf{q}(t)) \in W$ leading from $(\mathbf{p}_1, \mathbf{q}_1)$ to $(\mathbf{p}_2, \mathbf{q}_2)$ in the time interval $[t_1, t_2]$ with $(\mathbf{p}_1, \mathbf{q}_1), (\mathbf{p}_2, \mathbf{q}_2)$ fixed *a priori.*

If $\mathcal{C}$ satisfies (1.6) and H' is related to H as above, then any motion $t \to (\mathbf{p}(t), \mathbf{q}(t))$ developing in W from $(\mathbf{p}_1, \mathbf{q}_1)$ to $(\mathbf{p}_2, \mathbf{q}_2)$ in $[t_1, t_2]$ has a corresponding motion $t \to (\mathbf{p}'(t), \mathbf{q}'(t))$ developing in W from $\mathcal{C}(\mathbf{p}_1, \mathbf{q}_1)$ to $\mathcal{C}(\mathbf{p}_2, \mathbf{q}_2)$ in $[t_1, t_2]$ and (1.6) implies

1.8 $$\int_{t_1}^{t_2} \big(\mathbf{p} \cdot d\mathbf{q} - H(\mathbf{p}, \mathbf{q})\, dt = \int_{t_1}^{t_2} \big(\mathbf{p}' \cdot d\mathbf{q}' - H'(\mathbf{p}', \mathbf{q}')\, dt + \Phi_2 - \Phi_1 \tag{1.8}$$

with Φ_2, Φ_1 depending only on $(\mathbf{p}_1, \mathbf{q}_1), (\mathbf{p}_2, \mathbf{q}_2)$: hence the stationarity of $\mathcal{A}$ is equivalent to that of the corresponding $\mathcal{A}'$ and our assertion follows.

Whatever definition of integrability one chooses, it is clear that integrable systems have the following two properties:

(1) All motions stay bounded during their evolution, *i.e.* they remain at positive distance from ∂W, uniformly bounded from below.

(2) All motions are quasi periodic.

The variables $(\mathcal{A}, \boldsymbol{\varphi})$ defined by a map $\mathcal{C}$ integrating canonically a given system are called *action–angle coordinates*.

In studying the integrability of a system one usually first proves that its motions are all quasi periodic and evolve on invariant tori: successively one tries to build an action preserving completely canonical integrating map. In a later section I shall illustrate a useful algorithm for this purpose, allowing one, in most cases, to deduce a pair $\mathcal{C}, h$ satisfying (1.2).

Equation (1.2) for the pair $\mathcal{C}$ (an action preserving completely canonical map), h (an analytic function) is the *Hamilton–Jacobi equation* and the well known algorithm just mentioned will be the *Arnold–Liouville theorem.*

In the remaining part of this lecture I shall review a few remarkable examples of integrable systems, leaving aside the question of canonical integrability.

The example of a system of free rotators is, tautologically, an example of an integrable system: therefore thie first example that we shall consider, slightly less trivial, is that of *one dimensional systems*:

1.9
$$H(p,q) = \frac{p^2}{2m} + V(q) \tag{1.9}$$

with V analytic, $V(0) = 0, V'(q) \neq 0$ for $q \neq 0$, $V(q) \xrightarrow[q\to\pm\infty]{} \infty$, $m > 0$.

In this case if $q_1(E) < q_2(E)$ are the two roots of $V(q) = E$ with $E > 0$, one easily finds that the system is integrable in the region $H(p,q) > 0$ and all its motions are periodic with period

1.10
$$T(E) = 2\int_{q_-(E)}^{q_+(E)} \frac{dq}{\left(\frac{2}{m}(E - V(q))\right)^{\frac{1}{2}}} \tag{1.10}$$

As integrating transformation one can take the (noncanonical) map $\mathcal{I}(p,q) = (E, \varphi)$:

1.11
$$E = H(p,q), \qquad \varphi = \frac{2\pi}{T(E)} t(p,q) \tag{1.11}$$

where $t(p,q)$ is the time needed for the initial datum $p = 0, q = q_-(E)$ to reach the datum (q,p). This time is obviously defined modulo $T(E)$ so that φ is naturallly an angle (*average anomaly*): see Fig.2 and the motion in the coordinates (1.11) is $(E, \varphi) \to (E, \varphi + \frac{2\pi}{T(E)} t$, so that $\omega(E) = \frac{2\pi}{T(E)}$.

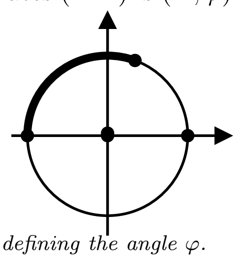


Fig.2 *The arc on the constant energy curve defining the angle* φ.

A second example is the two body problem:

1.12
$$H(p_\rho, p_\theta, \rho, \theta) = \frac{p_\rho^2}{2m} + \frac{p_\theta^2}{2m\,\rho^2} - \frac{K\,m}{\rho} \tag{1.12}$$

which is integrable in the region $H < 0, p_\theta > 0$ (and $H < 0, p_\theta < 0$).

In this case $p_\theta = G$ is a constant of motion (*angular momentum*) and the ρ coordinate evolves as described by the one dimensional Hamiltonian ($p \stackrel{def}{=} m\dot\rho$)

1.13 $$H = \frac{p^2}{2m} + \frac{G^2}{2m\rho^2} - \frac{K\,m}{\rho} \tag{1.13}$$

By the preceding example we can describe the motion of ρ in terms of two coordinates E, λ where E is the value of H and

1.14 $$\lambda = \frac{2\pi}{T(E,G)}\, t(\rho,\dot\rho) \tag{1.14}$$

where $t(\rho,\dot\rho)$ is the time it takes a motion of (1.13) to reach the datum $(\rho,\dot\rho)$, starting from $\rho(0) = \rho_-(E,G)$, $\dot\rho(0) = 0$, see Fig.3, where $\rho_\pm(E,G)$ are the roots of $\frac{2}{m}\big(E - \frac{G^2}{2m\rho^2} + \frac{K\,m}{\rho}\big) = 0$. The motion of the λ coordinate is simply described by $\lambda \to \lambda + \frac{2\pi}{T(E,G)}$, obviously by construction.

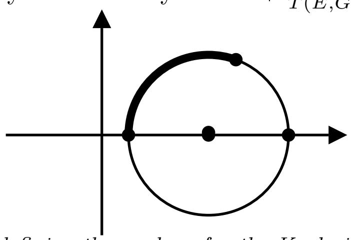


Fig.3 *The arc on the constant energy curve defining the angle φ for the Keplerian motion.*

The complete motion is found from the equation $m\rho^2\dot\theta = G$: since $\rho(t) = R(\lambda + \frac{2\pi}{T(E,G)})$, where R is the motion of (1.13) with initial datum $\dot\rho(0) = 0, \rho(0) = \rho_-(E,G)$ one finds

1.15 $$\theta(t) = \theta(0) + \int_0^t \frac{G}{m\,R\big(\lambda + \frac{2\pi}{T(E,G)}\big)^2}\, dt \tag{1.15}$$

which implies

$$\theta(t) = \theta(0) + \chi_0(E,G)t + \sum_{k\ne0} \chi_k(E,G)\frac{\exp(\frac{2\pi\,i\,k}{T(E,G)}t) - 1}{\frac{2\pi\,i\,k}{T(E,G)}} e^{ikt}$$

where $\chi_k(E,G)$ are the Fourier coefficients of the function $G/m\,R(\lambda)^2$.
Defining $\mu = \theta - \sum_{k\ne0}\chi_k(E,G)\frac{2\pi ik}{T(E,G)}e^{ik\lambda}$ (which is real in spite of being define in terms of complex numbers) one sees that the coordinates $(E,G,]l,\mu)$ evolve as

1.16 $$(E,G,\lambda,\mu) \to (E,G,\lambda + \frac{2\pi}{T(E,G)}t, \mu + \chi_0(E,G)t) \tag{1.16}$$

showing the integrability.

In fact it turns out that $\frac{2\pi}{T(E,G)} = \chi_0(E,G)$ and all motions are periodic. They evolve on ellipses with polar equation

1.17 $$\frac{1}{\rho} = (\frac{1}{\rho_-} + \frac{1}{\rho_+}) + (\frac{1}{\rho_-} - \frac{1}{\rho_+})\cos\theta \tag{1.17}$$

I shall not go through the well known calculations to obtain (1.17).

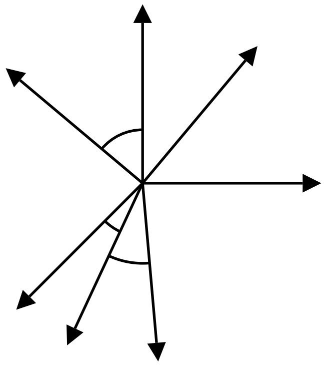


Fig.4 *The Euler angles.*

A third example is a solid with a fixed point and inertia moments $0 < I_1 < I_2 < I_3$ on the principal inertia axes $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$. The Euler equations are

1.18 $$I_1\omega_1 = (I_2 - I_3)\omega_2\omega_3, \qquad I_2\omega_1 = (I_3 - I_1)\omega_2\omega_3, \qquad I_3\omega_1 = (I_1 - I_2)\omega_2\omega_3 \tag{1.18}$$

where $\boldsymbol{\omega}$ is the angular velocity vector, which in terms of the Euler angles is

1.19 $$\boldsymbol{\omega} = \dot\theta\,\mathbf{n} + \dot\varphi\,\mathbf{k} + \dot\psi\,\mathbf{i}_3 \tag{1.19}$$

whose components on the comoving frame are denoted $\omega_1, \omega_2, \omega_3$.

We take the z-axis to be paprallel to the (constant) angula momentum $\mathbf{K} = A\,\mathbf{k} = \mathbf{I}\boldsymbol{\omega}$, where $\mathbf{I}$ is the (diagonal) inertia matrix.

The three comoving components of $\mathbf{k}$ are

1.20 $$\begin{aligned} K_3 =& I_3\omega_3 = A\cos\theta\\ K_2 =& I_2\omega_2 = A\sin\theta\cos\psi\\ K_1 =& I_3\omega_3 = A\sin\theta\sin\psi\\ &\Rightarrow\ \cos\theta = \frac{I_3}{A}\omega_3,\ \tan\psi = \frac{I_1\omega_1}{I_2\omega_2} \end{aligned} \tag{1.20}$$

The energy and angular momentum are E and A, respectively

1.21 $$2E = I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2, \qquad A = I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2 \tag{1.21}$$

which allows us to express ω_1, ω_3 in terms of ω_2; after substitution into (1.18) one finds

1.22 $$\dot\omega_2^2 + V_{E,A}(\omega_2) = 0 \tag{1.22}$$

telling us that ω_2 varies as a one–dimensional motion of a mass 1 particle with zero energy and potential energy $\frac12 V_{E,A}$ whic, as is easy to compute, is given by

1.23 $$V_{E,A}(\omega) = -\frac{\big[(2EI_3 - A^2) - (I_3 - I_2)I_2\omega_2^2\big]\big[(A^2 - 2EI_1) - (I_2 - I_1)I_2\omega_2^2\big]}{I_1 I_2^2 I_3} \tag{1.23}$$

This implies that ω_2 varies periodically with period $T(E,A)$ and so do ω_1, ω_3 and, by (1.20), θ, ψ.

It remains to determine φ. First one remarks that the simple relation

1.24 $$\omega_3 = \dot\psi + \dot\varphi \cos\theta \tag{1.24}$$

that $\dot\varphi$ is periodic in time.

Hence, as in the two body problem, it follows that φ is quasi–periodic with two periods $T(E, A)$ and $2\pi/\overline{\dot\varphi}$ where $\overline{\dot\varphi}$ is the average of $\dot\varphi$ (which is periodic). More precisely let $\Omega_\pm(E, A)$ be the two roots of $V_{E,A} = 0$. If $\Omega(t, E, A)$ is the solution to the equation (see (1.24))

1.25 $$\ddot\omega + \frac{1}{2}\partial_\omega V_{E,A}(\omega) = 0 \tag{1.25}$$

with initial datum $\omega(0) = \Omega_-(E, A), \dot\omega(0)$, we have

1.26 $$\omega_2(t) = \Omega(t(\omega_2, \dot\omega_2) + t, E, A) \tag{1.26}$$

where $t(\omega_2, \dot\omega_2)$ is the time needed by the solution Ω of (1.25) to reach the datum $\omega_2, \dot\omega_2$.

By (1.20),(1.24) and (1.26) it is clear that

1.27 $$\dot\varphi = \Phi(E, A, t(\omega_2, \dot\omega_2) + t), \tag{1.27}$$

where Φ is a suitable function with period $T(E, A)$ in the third variable. Denoting by $\Phi_k(E, A)$ the Fourier transform of Φ, we find:

1.28 $$\varphi(t) = \varphi(0) + t\,\Phi_0(E, A) + \sum_{k\neq 0} \Phi_k(E, A)\frac{e^{2\,\pi\,i\,k\,t/T(E,A)} - 1}{2\,\pi\,i\,k\,t/T(E, A)} \tag{1.28}$$

showing that the angle

1.29 $$\eta = \varphi - \sum_{k\neq 0} \Phi_k(E, A)\frac{e^{2\,\pi\,i\,k\,t/T(E,A)}}{2\,\pi\,i\,k\,t/T(E, A)} \tag{1.29}$$

varies uniformly in time.

At this point the system is integrated and one can use the following coordinates E = energy, A = angular momentum, δ = angle between the z–axis of the fixed reference frame and $\mathbf{K}$, $\varepsilon = \frac{2\,\pi t(\omega_2,\dot\omega_2)}{T(E,A)}$, η = (see]equ(1.29), γ = angle between the node line of the plane orthogonal to $\mathbf{K}$ and the xy–plane of a fixed reference frame and the x–axis.

It is easy to find explicit expressions for $T(E, A)$ and $\Phi_0(E, a)$ in terms of elliptic integrals.

As a fourth example consider the geodesic motion on an ellipsoid for the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. The natural (Jacobi) coordinates are defined by

1.30 $$\begin{aligned} x =& \sqrt{a}\,\cos\theta\,\sqrt{\varepsilon + (1-\varepsilon)\cos^2\varphi}, \\ y =& \sqrt{b}\,\sin\theta\,\cos\varphi, \qquad \varepsilon = \frac{b-a}{c-a} \\ z =& \sqrt{c}\,\sin\varphi\,\sqrt{1-\varepsilon\cos^2\theta}. \end{aligned} \tag{1.30}$$

and the ellipsoid is covered twice as (θ, φ) vary in $\mathbb{T}^2$.

Setting $u = a + (b-a)\cos^2\theta$, $v = b + (c-b)\cos^2\varphi$, then (1.30) takes the well known form:

1.31 $$\pm x=\sqrt{a}\Big(\frac{(u-a)(v-a)}{(b-a)(c-a)}\Big)^{\frac{1}{2}},\ \pm y=\sqrt{a}\Big(\frac{(b-u)(v-b)}{(b-a)(c-b)}\Big)^{\frac{1}{2}},\ \pm z=\sqrt{a}\Big(\frac{(c-u)(c-v)}{(c-a)(c-b)}\Big)^{\frac{1}{2}}, \tag{1.31}$$

After a few calculations one finds that th eLagrangian (*i.e.* the kinetic energy) is:

1.32 $$\mathcal{L}=\frac{1}{2}\big[\dot\theta^2\,\alpha(\theta)+\dot\varphi^2\,\beta(\varphi)\big]\big[g(\theta)+h(\varphi)\big] \tag{1.32}$$

where

1.33 $$\begin{aligned}\alpha(\theta)&=\frac{(c-a)+(b-a)\,\cos^2\theta}{a+(b-a)\cos^2\theta}, \qquad g(\theta)=(b-a)\sin^2\varphi\\ \beta(\varphi)&=\frac{(b-a)+(c-b)\,\cos^2\varphi}{b\sin^2\varphi+c\cos^2\varphi}, \qquad h(\varphi)=(c-b)\cos^2\varphi\end{aligned} \tag{1.33}$$

Hence the Hamiltonian is

1.34 $$H=\frac{1}{2}\Big(\frac{p_\theta^2}{\alpha(\theta)}+\frac{p_\varphi^2}{\beta(\varphi)}\Big)\frac{1}{g(\theta)+h(\varphi)} \tag{1.34}$$

Without using the special form of α,β,h,g (1.34) implies

1.35 $$\frac{p_\theta^2}{2\alpha(\theta)}-Eg(\theta)=\eta=-\frac{p_\varphi^2}{2\beta(\varphi)}-Eh(\varphi) \tag{1.35}$$

(with $H=E$) is a constant of motion.

One easily expresses $M=\eta E/b$ in terms of the original angles:

1.36 $$\begin{aligned}M=&\frac{1}{2b}\Big[a+(b-a)\cos^2\theta\Big]\Big[b+(c-b)\cos^2\varphi\Big]\cdot\Big[(b-a)\sin^2\theta+(c-b)\cos^2\varphi\Big]\cdot\\ &\cdot\Big[\frac{\dot\theta^2}{(c-a)+(b-a)\cos^2\theta}+\frac{\dot\varphi^2}{(b-a)+(c-b)\cos^2\varphi}\Big].\end{aligned} \tag{1.36}$$

and one also finds that $M\ge E$ and if $M>E$ the motion cannot visit the singularities of the coordinate system (1.30) which, when $a<b<c$ are found only for $\sin\theta=\cos\varphi=0$, *i.e.* for $x=\pm\sqrt{a\varepsilon}, y=0, z=\pm\sqrt{(1-\varepsilon)c}$.

If one considers the angle variables ψ_1,ψ_2 defined as the average anomalies for the one dimensional systems

1.37 $$H_1=\frac{1}{2}p_\theta^2-\big(E\,\alpha(\theta)\,g(\theta)-\eta\,\alpha(\theta)\big), \qquad \frac{1}{2}p_\beta^2-\big(E\,\beta(\varphi)\,h(\varphi)-\eta\,\beta(\varphi)\big), \tag{1.37}$$

corresponding to their zero energy motions (*i.e.* $H_1=0$, $H_2=0$) one sees that the system is integrated by the coordinates (E,M,ψ_1,ψ_2) and that ψ_1 rotates with period

1.38 $$T_1(E,M)=2\int_{\theta_-(E,M)}\theta_+(E,M)\frac{d\theta}{\sqrt{2(E\,\alpha(\theta)\,g(\theta)-\eta\,\alpha(\theta))}} \tag{1.38}$$

with the usual notations, and ψ_2 rotates with a period given by a similar expression.

Finally we consider the one dimensional Schrödinger equation in a periodic potential

1.39 $$-\ddot q + \varepsilon\, q\, V(\omega t) = \, E\, q \tag{1.39}$$

where V is *periodic and analytic on* $\mathbb{T}^1$. Eq. (1.39) corresponds to the Hamiltonian system

1.40 $$H(p,q,T,\psi) = \omega T + \frac{p^2}{2} + (E - \varepsilon V(\psi))\,\frac{q^2}{2} \tag{1.40}$$

Using the completely canonical map $\mathcal{C}(p,q) = (A,\varphi)$ with

1.41 $$\begin{aligned} A =& \sqrt{E}\,\big(\frac{p^2}{2} + \frac{E\,q^2|}{2}\big), \qquad \varphi = \arctan\frac{p}{E\,q},\\ p =& (2A\sqrt{E})^{\frac12}\,\cos\varphi, \qquad q = (2A\sqrt{E})^{\frac12}\,\sin\varphi, \end{aligned} \tag{1.41}$$

Eq. (1.40) becomes, if (A,φ) and (T,ψ) are conjugate pairs,

1.42 $$\sqrt{E}A - \frac{\varepsilon A}{\sqrt{E}} V(\psi)\sin^2\varphi + \omega T. \tag{1.42}$$

Integration of (1.42) is equivalent to the construction of the Bloch waves for (1.39) and works for E inside the *bands* of (1.39).

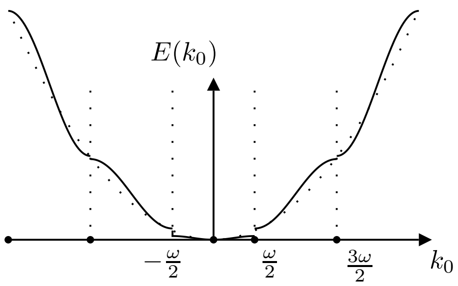
More generally a quasi periodic Schrödinger equation

1.43 $$-\ddot q + \varepsilon\, V(\omega_1 t, \omega_2 t, \ldots, \omega_\ell t) = E\, q \tag{1.43}$$

is equivalent to the $(1+\ell)$–dimensional system with Hamiltonian

1.44 $$\sqrt{E}A + \omega_1 T_1 + \ldots + \omega_\ell T_\ell - \frac{\varepsilon A}{\sqrt{E}} \sin^2\varphi\, V(\psi_1,\ldots,\psi_\ell). \tag{1.44}$$

Note that, in general, every value of E for which (1.43) admits a quasi periodic solution will be a point in the continuum spectrum of (1.43) and, in fact, this is a method that can be used to study the continuum spectum of (1.43).

1.45 

(1.45)

For completeness we give here the construction of the second integral of motion for (1.42), for E in the bands. So we consider the Hamiltonian system (1.42) and fix $k_0 \in (-\frac{\omega}{2}, \frac{\omega}{2})$. We define a *Bloch wave with wave number* k_0 a solution of the Hamiltonian equations for (1.42) of the form

1.46 $$q(t) = e^{ik_0 t} q_0(\psi_0 + \omega t) \tag{1.46}$$

with $q_0 \in C^\infty(\mathbb{T}^1)$. The equation for q_0 is

1.47 $$(k_0^2 - E)\, q_0 - 2ik_0 \dot q_0 + \varepsilon\, V(\omega t + \psi_0)\, q_0 - \omega^2 \ddot q_0 = 0 \tag{1.47}$$

Eq. (1.47) is a self adjoint equation on $L_2(\mathbb{T}^1)$ with discrete spectrum. For fixed k_0 we can find a sequence $E_n(k_0)$, $n = 0, 1, \ldots$ of eigenvalues for (1.47). The parametrization can be made so that $E_n(k_0)$ are, for each n, continuous functions of k_0. One finds the picture of Fig.(1.45). This picture is obtained by first drawing $E_0(k_0)$ over the interval $(-\frac{\omega}{2}, \frac{\omega}{2})$, (the *first Brillouin zone*); then we draw $E_1(k_0)$ over the interval $(\frac{\omega}{2}, \frac{3}{2}\omega)$ and $E_2(k_0)$ over the interval $-\frac{3}{2}\omega, -\frac{\omega}{2})$ *instead* of drawing them again on $(-\frac{\omega}{2}, \frac{\omega}{2})$. If $E_n(k_0)$ are appropriately labeled (continuing the above rule) one defines a function $E(k_0)$ with a graph as in Fig.(1.45) and one can prove that $E(k_0)/k_0^2$ approaches a constant limit as $k_0 \to \infty$.

Let $E = E(k_0)$, $k_0 \neq \pm\frac{1}{2}(2k+1)\omega$, $k = 0, 1, 2, \ldots$; the claim is that for such E the system (1.42) is an integrable system. Let $\overline{k}_0 = k_0 - \frac{1}{2}(2k+1)\omega$ with k so chosen that $|\overline{k}_0| < \frac{\omega}{2}$.

The change of variables (1.41) and the correspondence between the solutions of (1.43) and (1.42) ($\ell = 1$) suggests introducing the following functions on $\mathbb{T}^2$:

1.48
$$\begin{aligned} Q(\alpha_1,\alpha_2) =& \mathrm{Re}\,(e^{i\alpha_1} q_{k_0}(\alpha_2)),\\ P(\alpha_1,\alpha_2) =& \mathrm{Re}\,(i\overline{k}_0 e^{i\alpha_1} q_{k_0}(\alpha_2) + e^{i\alpha_1}\omega q'_{k_0}(\alpha_2))\\ a(\alpha_1,\alpha_2) =& \frac{P^2 + E\,Q^2}{2\sqrt{E}}, \qquad \Gamma(\alpha_1,\alpha_2) = \frac{\varepsilon}{2} V'(\alpha_2) Q^2(\alpha_1,\alpha_2), \end{aligned} \tag{1.48}$$

where V' is the derivative of V and q'_{k_0} that of q_{k_0}. Then given $\rho \in \mathbb{R}$ the function

1.49
$$q(t) = \rho Q(\alpha_0 + \overline{k}_0 t, \psi_0 + \omega t), \quad \psi(t) = \psi_0 + \omega t) \tag{1.49}$$

gives a solution of the Hamiltonian equations with initial data

1.50
$$p = \rho\, P(\alpha_0, \psi_0), \qquad q = \rho\, Q(\alpha_0, \psi_0), \qquad \psi = \psi_0, \qquad B = B_0 \tag{1.50}$$

provided $B(t)$ is defined so that

1.51
$$\dot{B} = \frac{\rho}{2}\varepsilon\, V'(\psi_0 + \omega t) Q^2(\alpha_0 + \overline{k}_0 t, \psi_0 + \omega t) = \rho^2 \Gamma(\alpha_0 + \overline{k}_0 t, \psi_0 + \omega t) \tag{1.51}$$

Using that (1.48) implies that the Fourier transform $\Gamma_{r_1\,r_2}$ of Γ vanishes unless $r_1 = 0, \pm 2$ we see that $\overline{k}_0 r_1 + \omega r_2$ can vanish only if $r_1 = r_2 = 0$ or if $\pm 2\overline{k}_0 + \omega r_2 = 0$ and this can happen only if $\overline{k}_0$ is an extreme point of the Brillouin zone, against our assumption that E is inside a band. This means that, if $\mathbf{r} = (r_1, r_2)$,

1.52
$$B(t) = B_0 + \rho^2 \Gamma_{00} t + \rho^2 \sum_{\mathbf{r}\neq\mathbf{0}} \Gamma_{\mathbf{r}} \frac{e^{i(r_1\overline{k}_0 + r_2\omega)t} - 1}{i(r_1\overline{k}_0 + r_2\omega)}, \tag{1.52}$$

which implies that if $\Gamma_{\mathbf{0}} = 0$

1.53
$$\mathcal{B} = B - \rho^2 \sum_{\mathbf{r}\neq\mathbf{0}} \Gamma_{\mathbf{r}} \frac{e^{i(r_1\overline{k}_0 + r_2\omega)t}}{i(r_1\overline{k}_0 + r_2\omega)} \tag{1.53}$$

is a constant of motion. SO if $\Gamma_{\mathbf{0}} = 0$ we have shown that the system admits a third constant of motion (besides H, ρ) and *also* that the motions are quasi periodic; hence the system is integrable. Writing

1.54 $$Q(\alpha_1,\alpha_2)=\frac{1}{2}(e^{i\alpha_1}\,q_{k_0}(\alpha_2)+e^{-i\alpha_1}\,\overline{q}_{k_0}(\alpha_2)) \tag{1.54}$$

one realizes that

1.55 $$\Gamma_{\mathbf{0}}\propto\int_0^{2\pi}V'(\alpha)|q'_{k_0}(\alpha)|^2\,d\alpha=0 \tag{1.55}$$

which vanishes as it follows by considering (1.47), multiplying it by $\overline{q}'_{k_0}$ (where the $'$ denotes the derivative ∂_α) and taking the real part of the result and integrating over a period. Note that the three constants of motion are not independent as one can check that $\mathcal{B}=H/\omega$.

sec.2 2. Canonical integrability and the Arnold–Liouville theorem.

In the preceding lecture the rather typicalsituation has arisen in which integrability is checked by explicitly exhibiting foliatons into invariant tori of the relevant regions of phase space and the relative quasi p;eriodic motions.

Doubt may arise that this does not necessarily mean that the system is canonically integrable or that, even if the system is canonically integrable, the construction of the action angle variables may be a very hard task.

It is therefore interesting to present a few general considerations which eventually also lead to a simple constructive algorithm to define action–angle variables.

First we remark that if a system (H,W) is canoniclly integrable the completely canonical map $\mathcal{C}:\,W\longleftrightarrow V\times\mathbb{T}^\ell$

2.1 $$(\mathbf{A},\boldsymbol{\varphi})=\mathcal{C}(\mathbf{p},\mathbf{q}) \tag{2.1}$$

then the conservation of the Poisson brackets under completely canonical maps (see for instance Ref. [1],p.237) implies that the system admits on W ℓ constants of motion, namely $A_1,\ldots,A_\ell$

2.2 $$\begin{aligned}&\{A_i,A_j\}\stackrel{def}{=}\sum_{k=1}^{\ell}\big(\partial_{p_k}A_i\partial_{q_k}A_j-\partial_{q_k}A_i\partial_{p_k}A_j\big)=0,\quad\text{or}\\&\{\mathbf{A},\mathbf{A}\}=(\partial_{\mathbf{p}}\mathbf{A})\cdot(\partial_{\mathbf{q}}\mathbf{A})^T-(\partial_{\mathbf{q}}\mathbf{A})\cdot(\partial_{\mathbf{p}}\mathbf{A})^T.\end{aligned} \tag{2.2}$$

where $\{\text{cot},\cdot\}$ denotes here and in the folloeing the *Poisson bracket.* Thus a necessary condition for canonical integrability is the existence of ℓ constants of motion *in involution.*

p.s2.1 **2.1.** *Does existence of ℓ constants of motion in involution imply integrability ? (generically yes).* It is therefore interesting to examine what happens when a system (H,W) admits ℓ constants of motion $I_1,\ldots,I_\ell$ independent,in involution and such that $H=I_1$ or, more generally, $H=h(I_1,\ldots,I_\ell)$. This is the situatin which arises in the one and two degrees of freedom systems admitting one constant of motion besides the energy (like two–body problem,or the geodesic motion on the ellipsoid or the one–dimensional Schrödinger equation) because the condition on a function for being a constant of motion is just that of being in involution with the Hamiltonian. This is also the situation that sometimes arises in systems with more than two degrees of freedom like the solid with a fixed point where three involutory constants of motion can be taken $E=$ energy,

A = angular momentum, δ = angle between the z-axis and the angular momentum (this is so because $A, A_2 = A\cos\delta$ are, as is well known, from the angular momentum commutation rules, in involution).

Then assuming that the system admits ℓ involutry integrals $I_1, \ldots, I_\ell$ among whch the energy, and assuming that the surfaces $\mathbf{I} = const$ are boundaryless and compact for$\mathbf{I}$ near some $\mathbf{I}_0$ and if some technical invertibility dondition (to be specified later) is satisfied, then the system is canonically integrable in the vicinity of the surface $\mathbf{I} = \mathbf{I}_0$. In fact, as we shall see, the condition that H is one of the I_j can be relaxed to H being a function of the I_j.

The thorem is surprisingly simple if one does not try to avoid the use of "technical assumptions" and its proof can be divided in two distinct parts, [3].

(i) One proves that the surfaces

2.3
$$\sigma(\mathbf{i}) = \big\{\mathbf{p},\mathbf{q}\,|\,(\mathbf{p},\mathbf{q}) \in W,\, I(\mathbf{p},\mathbf{q}) = \mathbf{i}\big\} \tag{2.3}$$

for $\mathbf{i}$ close to $\mathbf{I}_0$ are ℓ–dimensional tori. In most applications such a property will however be known *a priori* as in all the examples of Sec. 1. In fact in most applications one will be able to integrate the motion by showing that it is quasi periodic, hence that it takes place on invariant tori of dimension ℓ and one will wish to apply the above Arnold–Liouville theorem only to prove the canonical integrability (starting from simple integrability).

(ii) One shows that the $\sigma(\mathbf{i})$ are *Lagrangian manifolds* in phase space, *i.e.* they are surfaces on which the differential form $\mathbf{p}\cdot d\mathbf{q}$ is locally integrable. Once $\sigma(\mathbf{i})$ is represented parametrically as

2.4
$$\mathbf{p} = \boldsymbol{\alpha}(\mathbf{i},\boldsymbol{\psi}), \qquad \mathbf{q} = \boldsymbol{\beta}(\mathbf{i},\boldsymbol{\psi}), \quad \boldsymbol{\psi} \in \mathbb{T}^\ell \tag{2.4}$$

(possible because $\sigma(\mathbf{i})$ is a torus) this allows one to define the action variables as

2.5
$$A_k(\mathbf{i}) = \frac{1}{2\pi}\oint_{\gamma_k(\mathbf{i})} \mathbf{p}\cdot d\mathbf{q} \tag{2.5}$$

where $\gamma_k(\mathbf{i})$ is the k-th circle on $\sigma(\mathbf{i})$ in the chosen coordinates, *i.e.* $\gamma_k(\mathbf{i}) = \big\{\boldsymbol{\psi}\,|\,\psi_j = 0.\, j \neq k\big\}$. The angle variables $\boldsymbol{\varphi}$ will be defined so that locally the function

2.6
$$S(\mathbf{q},\mathbf{A}) = \int_{\boldsymbol{\psi}}^{(\mathbf{p},\mathbf{q})} \mathbf{p}\cdot d\mathbf{q} \tag{2.6}$$

is the generating function of the completely canonical map $(\mathbf{p},\mathbf{q})\longleftrightarrow(\mathbf{A},\boldsymbol{\varphi})$; the details are given below; in (2.6) the integral is along any curve on $\sigma(\mathbf{i})$ joining the point $\boldsymbol{\psi} = \mathbf{0}$ (say) to the point with position coordinate $\mathbf{q}$.

The proof of (i) is based on abstract considerations of differential topology. If on an ℓ–dimensional boundaryless compact surface $\sigma(\mathbf{i})$ one can define ℓ independent vector fields commuting with each other, then $\sigma(\mathbf{i})$ must be a ℓ–dimansional torus, *i.e.* it must be representable in the form (2.4). In our case of course the ℓ vector fields are ℓ Hamiltonian fields with Hamiltonians $I_1, \ldots, I_\ell$. The condition of involution $\big\{I_i, I_j\big\} = 0$ simply means that the ℓ Hamiltonian fields commute and, therefore, must be tangent to $\sigma(\mathbf{i})$ (because the Hamiltonian is a constant of motion in all motions generated by the corresponding Hamiltonian vector fields) and can be regarded as commuting vector field defined on $\sigma(\mathbf{i})$.[1]

Part (ii) of the theorem is much more interesting. To study the form $\mathbf{p}\cdot d\mathbf{q}$ on $\sigma(\mathbf{i})$ one expresses $\mathbf{p}$ in terms of $\mathbf{q}$ and $\mathbf{i}$ locally, by inverting the relation $\mathbf{i}=\mathbf{I}(\mathbf{p},\mathbf{q})$ with respect to $\mathbf{p}$: we assume that this is possible: in general the independence of the $\mathbf{I}$'s does not necessarily imply that this is possible because it only means that $\partial_{\mathbf{p}\,\mathbf{q}}\mathbf{I}(\mathbf{p},\mathbf{q})$ has rank 2ℓ, but it might be that $\partial_{\mathbf{p}}\mathbf{I}$ has rank lower than ℓ, see the example in the figure below. The general case, however, could be treated by suitably changing the role of some coordinates.

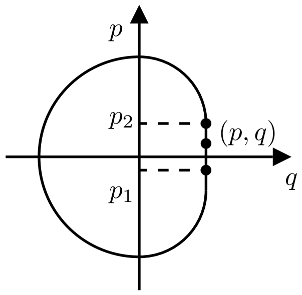


Fig.6 $\ell=1$ *and* $H=I$ *such that for* $p\in\big[p_1,p_2\big]$ *the graph of* H *is vertical: one sees that, if* $\big(p,q\big)$ *is like the point marked in the figure,* $\pi\big(i,q\big)$ *cannot be defined.*

Let $\mathbf{p}=\boldsymbol{\pi}(\mathbf{i},\mathbf{q})$ be the inverse function of $\mathbf{i}=\mathbf{I}(\mathbf{p},\mathbf{q})$: then $\mathbf{p}\cdot d\mathbf{q}$ becomes on $\sigma(\mathbf{i})$

2.7 $$\boldsymbol{\pi}(\mathbf{i},\mathbf{q})\cdot d\mathbf{q} \tag{2.7}$$

Saying that $\sigma(\mathbf{i})$ is a Lagrangian manifold just means that

2.8 $$\partial_{q_h}\pi_k=\partial_{q_k}\pi_h \qquad \text{for all } k,h \tag{2.8}$$

which can be checked immediately to be just a different way of writing the involution conditions $\big\{I_i,I_j\big\}=0$ for all i,j.

In fact from $\mathbf{i}=\mathbf{I}(\boldsymbol{\pi}(\mathbf{i},\mathbf{q}),\mathbf{q})$ it follows that

2.9 $$\mathbf{0}=\partial_{\mathbf{p}}\mathbf{I}\,\partial_{\mathbf{q}}\boldsymbol{\pi}+\partial_{\mathbf{q}}\mathbf{I} \qquad\Rightarrow\qquad \partial_{\mathbf{q}}=-(\partial_{\mathbf{p}}\mathbf{I})^{-1}\partial_{\mathbf{q}} \tag{2.9}$$

which allows us to write (2.8) as: $-\big(\partial_{\mathbf{p}}\mathbf{I}\big)^{-1}\big(\partial_{\mathbf{q}}\mathbf{I}\big)=-\big(\partial_{\mathbf{q}}\mathbf{I}\big)^T\big((\partial_{\mathbf{p}}\mathbf{I})^T\big)^{-1}$ or

[1] The construction of the parametric representation (2.4) requires solving the Hamiltonian equations with Hamiltonians $I_1,\ldots,I_\ell$ near $\boldsymbol{I}=\boldsymbol{I}_0$. Fix one point $\mathbf{o}\in\sigma(\boldsymbol{I})$ and define for $\boldsymbol{t}\in\mathbb{R}^\ell$, $\big(\boldsymbol{p}(\boldsymbol{t}),\boldsymbol{q}(\boldsymbol{t})\big)=S^{t_1}_{I_1}\ldots S^{t_\ell}_{I_\ell}\mathbf{o}$, where $S^{t_j}_{I_j}$ is the evolution generated by I_j. This realizes a map of $\mathbb{R}^\ell$ on a subset of $\sigma(\boldsymbol{i})$ which is open (by the independence of $I_1,\ldots,I_\ell$) and which must also have an open complement: hence it coincides with $\sigma(\boldsymbol{i})$. Furthermore $\boldsymbol{t}\to\boldsymbol{t}+\boldsymbol{\tau}$ induces an action of $\mathbb{R}^\ell$ on $\sigma(\boldsymbol{i})$ via $\big(\boldsymbol{p}(\boldsymbol{t}),\boldsymbol{q}(\boldsymbol{t})\big)\to\big(\boldsymbol{p}(\boldsymbol{t}+\boldsymbol{\tau}),\boldsymbol{q}(\boldsymbol{t}+\boldsymbol{\tau})\big)$. Therefore $\sigma(\boldsymbol{i})$ is a *homogeneous* compact space for the group $\mathbb{R}^\ell$: hence there must exist ℓ real positive numbers $\big(T_1,\ldots,T_\ell\big)$ such that $\boldsymbol{t}$ and $\boldsymbol{t}+\boldsymbol{T}$ represent the same point on $\sigma(\boldsymbol{i})$ and for all $\boldsymbol{t}'$ with $0<t'_1<T_1,\ldots,0<t'_\ell<T_\ell$ the $\boldsymbol{t}$ and $\boldsymbol{t}'$ *do not* represent the same point. So the coordinates ψ can be defined as $\psi_j=2\pi t_jT_j^{-1}$, $j=1,\ldots,\ell$ and $\big(\boldsymbol{\alpha}(\boldsymbol{i},\boldsymbol{\psi}),\boldsymbol{\beta}(\boldsymbol{i},\boldsymbol{\psi})\big)=\big(\boldsymbol{p}(\frac{T_1}{2\pi}\psi_1,\ldots,\frac{T_\ell}{2\pi}\psi_\ell),\boldsymbol{q}(\frac{T_1}{2\pi}\psi_1,\ldots,\frac{T_\ell}{2\pi}\psi_\ell)\big)$.

2.10 $$\big(\partial_{\mathbf{q}}\mathbf{I}\big)\big(\partial_{\mathbf{p}}\mathbf{I}\big)^T=\big(\partial_{\mathbf{p}}\mathbf{I}\big)\big(\partial_{\mathbf{q}}\mathbf{I}\big)^T\qquad\longleftrightarrow\qquad\big\{\mathbf{I},\mathbf{I}\big\}=\mathbf{0}\boldsymbol{\tau}_{\boldsymbol{\alpha}_{\boldsymbol{\tau}}}\tag{2.10}$$

It is now possible to define ℓ new constants of mption, $\mathbf{A}(\mathbf{I})$, via (2.5) and to define the multivalued function (2.6). The different determinations of $S(\mathbf{a},\mathbf{q})$ on $\sigma(\mathbf{i})$ differ by $\sum_{j=1}^{\ell}2\pi A_j m_j$ where $\mathbf{m}=(m_1,\ldots,m)\ell)\in\mathbb{Z}^\ell$ are integers; this follows immediately by the local integrability of the form $\mathbf{p}\cdot d\mathbf{q}$ and by the defintion of the A_j.

Defining for $(\mathbf{p},\mathbf{q})\in\sigma(\mathbf{i})$, $\mathbf{A}=\mathbf{A}(\mathbf{i})$

2.11 $$\boldsymbol{\varphi}=\partial_{\mathbf{A}}S(\mathbf{q},\mathbf{A})\tag{2.11}$$

one realizes that the $\boldsymbol{\varphi}$'s are defined modulo 2π (because $S(\boldsymbol{\varphi},\mathbf{A})$ is defined modulo $2\pi\mathbf{m}\cdot\mathbf{A}$).

We now make the technical assumption that the map $\mathcal{C}:(\mathbf{A},\varphi)\longleftrightarrow(\mathbf{p},\mathbf{q})$ defined by (2.5), (2.11) is invertible in the vicinity of $\sigma(\mathbf{I}_0)$.

By construction this map is then canonical and, in fact, completely canonical action preserving, because setting $\Phi=S-\mathbf{A}\cdot\varphi$, F is single valued on the grpah of the map and

2.12 $$\mathbf{p}\cdot d\mathbf{q}=\partial_{\mathbf{q}}S\cdot d\mathbf{q}=dS-\partial_{\mathbf{A}}S\cdot d\mathbf{A}=dS-\boldsymbol{\varphi}\cdot d\mathbf{A}=d(S-\boldsymbol{\varphi}\cdot\mathbf{A})+\mathbf{A}\cdot d\boldsymbol{\varphi}=\mathbf{A}\cdot d\boldsymbol{\varphi}+d\Phi.\tag{2.12}$$

Of course $\mathcal{C}$ integrates canonically the system because $H(\mathbf{p},\mathbf{q})=h(\mathbf{A})$ by construction.

Note that the example in Fig.6 shows us in Fig.6 shows us that the "technical assumption" is not trivial in general; in that case (2.11) does not have the required invertibility properties. However this same example shows that the theorem holds under more general assumptions, as it it an easy exercise to check that the system considered in Fig.6 is in fact canonically integrable.

[Prove that the action–angle variables can be defined, if $(\partial_p H,\partial_q H)\neq\mathbf{0}$ by

2.13 $$A=\oint p\frac{dq}{2\pi}=\oint\widetilde{T}(E)\frac{dE}{2\pi},\qquad\varphi=\frac{2\pi}{T(A)}t(p,q)\tag{2.13}$$

where $\widetilde{T}(E)=T(A)$ is the period of the motion with energy E or corresponding action A and $t(p,q)$ is the time required by a reference motion of action A to reach the initial datum. Show that $\frac{dE}{dA}=\frac{2\pi}{T(A)}$. In fact the "technical assumption" made above is not necessary; for more general constructions which do not make the assumption see [4]].

The above formulae provide explicit expressions for the action–angle variables once the tori $\sigma(\mathbf{i})$ are known, *i.e.* once their parametric quations (2.5) have been found: however as explained above their determination requires the preliminary integration of the equations of motion by "standard methods".

p.s2.2 **2.2.** *Generically there cannot be more than ℓ independent constants of motion.*

A simple corollary of the proof of the above theorem is the well–known statement that a Hamiltonian system cannot in general admit more than ℓ independent constants of motion in involution. Let in fact $I_1,\ldots,I_\ell$ be ℓ independent constants of motion and let $I_{\ell+1}$ be another constant of motion. Suppose that $\big\{I_i,I_j\big\}=0$, $i,j=1,\ldots,\ell+1$. We *do not* suppose that the surfaces $(I_1,\ldots,I_\ell)=const$ are compact or boundaryless because the above statement has a local nature; we suppose, however, the "general" property that the $\mathbf{I}(\mathbf{p},\mathbf{q})=\mathbf{i}$ can be locally inverted with respect to $\mathbf{q}$ and $\mathbf{i}$ as $\mathbf{p}=\boldsymbol{\pi}(\mathbf{i},\mathbf{q})$. Then one defines $S(\mathbf{i},\mathbf{q}0$ locally by

2.14
$$S(\mathbf{i},\mathbf{q})=\int_{\mathbf{q}_0}^{\mathbf{q}}\mathbf{p}\cdot d\mathbf{q} \tag{2.14}$$

along any line on $\sigma(\mathbf{i})$ joining $\mathbf{q}_0$ to $\mathbf{q}$ (with $\mathbf{p}=\boldsymbol{\pi}(\mathbf{i},\mathbf{q})$); the calculations (2.9), (2.10) show that this is possible because $\mathbf{p}\cdot d\mathbf{q}$ is locally integrable at fixed $\mathbf{I}$. Then we define

2.15
$$\mathbf{p}=\partial_{\mathbf{q}}S(\mathbf{i},\mathbf{q}),\qquad \tau=\partial_{\mathbf{i}}S(\mathbf{i},\mathbf{q}) \tag{2.15}$$

and use these expressions to deine near a point $(\mathbf{p}_0,\mathbf{q}_0)$ a completely canonical map $(\mathbf{p},\mathbf{q})\longleftrightarrow(\mathbf{i},\boldsymbol{\tau})$. Then, since the canonical maps preseve the Poisson brackets

2.16
$$\{I_{\ell+1},I_j\}=0\qquad \Rightarrow \partial_{\tau_j}I_{\ell+1}=0,\quad j=1,\ldots,\ell \tag{2.16}$$

i.e. $I_{\ell+1}$ is locally a function of $I_1,\ldots,I_\ell$.

This illustrates the meaning of the involution property. Alternatively and more simply, one can just say that once $\mathbf{I}(\mathbf{p},\mathbf{q})=\mathbf{i}$ is assumed to define an ℓ–dimensional surface the involution condition means that on this surface there are ℓ pointwise independent vector fields; hence any other vector field must be dependent on them.

Other interesting consequences of the above idea can be deduced for *anisochronus systems.*

p.s2.3 **2.3.** *Existence of ℓ involutory constants of motion is necessary for canonical integrability if the system is anisochronous.*

A system which is integrable (not necessarily canonically integrable) is called *anisochronus* (sometimes one says *satisfyng the twist condition* or *twisting*) if, calliong $(\mathbf{I},\boldsymbol{\varphi})$ the integrating variables (called $(\mathbf{A},\boldsymbol{\varphi})$ in the definition at the beginning of Sect,1),

2.17
$$\det\partial_{\mathbf{I}}(\mathbf{I})\neq 0 \tag{2.17}$$

This means that at least locally in the I_i variables one can take as constants of motion the ω_i themselves and the integrating map can be given the form

2.18
$$\mathcal{I}(\mathbf{p},\mathbf{q})=(\boldsymbol{\omega},\boldsymbol{\varphi}),S_t\mathcal{I}^{-1}(\boldsymbol{\omega},\boldsymbol{\varphi})=\mathcal{I}^{-1}(\boldsymbol{\omega},\boldsymbol{\varphi}+\boldsymbol{\omega}t) \tag{2.18}$$

Then it is remarkable that, if a system is integrable and anisochronus, it can be canonically integrable only if

2.19
$$\{I_i,I_j\}=0 \tag{2.19}$$

In fact if the system is canoncally integrable, denoting $(\mathbf{A},\boldsymbol{\varphi})$ the action–angle variables:

2.20
$$\boldsymbol{\omega}=\boldsymbol{\omega}(\mathbf{A})=\partial_{\mathbf{A}}h(\mathbf{A}) \tag{2.20}$$

so that, locally, $\mathbf{I}=\mathbf{i}(\boldsymbol{\omega})=\widehat{\mathbf{i}}(\mathbf{A})$ and $\{A_i,A_j\}=0$ implies $\{I_i,I_j\}=\sum_{h,k=1}^{\ell}\partial_{A_h}\widehat{\mathbf{i}}_i\partial_{A_k}\widehat{\mathbf{i}}_j\{A_h,A_k\}$ $=0$. One can also interpret the above as saying that in anisochronous canonically integrable systems all constants of motion are in involution.

The anisochrony is essential: taking $H=0$ and $\ell\ge 3$ one can easily find a foliation into ℓ–dimensional (trivially invariant) tori of a subset $W\subset\mathbb{R}^{2\ell}$ whose ℓ parameters I_j depend on $\mathbf{p},\mathbf{q}$ so that $\{I_i,I_j\}\ne 0$.

p.s2.4 **2.4.** *Integrability and anisochrony imply canonical integrability*

A final question one can ask is whether integrability implies canonical integrability or, more generally, if quasi–periodicity of all motions implies integrability. *This question can be easily answered in the case in which the quasi–periodic motions have ℓ anisochronous periods. The few following paragraphs of this section differ from the original published text, and replace it, because the analysis was presented as an open problem (at least for me): however afterwards I found a simple solution to at least part of the problem and therefore I think that it is better to present it together with a new reference.*

Suppose that the system admits ℓ integrals of motion $\mathbf{I}$ and that $\mathbf{I}=\mathbf{i}$ determines a torus with frequencies $\boldsymbol{\omega}(\mathbf{I})$; suppose that $H=I_1$ and that the system is nanisochronous $\det\partial_{\mathbf{I}}\boldsymbol{\omega}(\mathbf{I})\ne 0$.

The case $\ell=2$ is reduced to the above discussed Arnold–Liouville construction. The cases $\ell\ge 3$ are more interesting but their analysis is essentially the same and we discuss only $\ell=3$ for simplicity.

To fall back on Arnold–Liouville construction we only need to show that $\{I_2,I_3\}=0$ as we already know that $\{H,I_2\}=0$ and $\{H,I_3\}=0$. We shall show that $\{I_2,I_3\}\ne 0$ leads to a contradiction.

Indeed let $J=\{I_2,I_3\}$: then $\{H,J\}=0$ by the Poisson brackets Jacobi identity. Let $S^{I_2}_\varepsilon(\mathbf{I},\boldsymbol{\varphi})$ be the Hamiltonian evolution with Hamiltonian I_2 and $S^H_t\boldsymbol{\varphi}$ be the Hamiltonian evolution with Hamiltonian H. Then $S^H_tS^J_\varepsilon\equiv S^{I_2}_\varepsilon S^H_t$ because $\{H,I_2\}=0$. Since $J\ne0$ it will be $S^{I_2}_\varepsilon(\mathbf{I},\mathbf{0})=(\mathbf{I}_\varepsilon,\boldsymbol{\varphi}_\varepsilon)$ and $\mathbf{I}_\varepsilon\ne\mathbf{I}$) because I_3 change to $O(\varepsilon)$ for ε small and $\det\partial_{\mathbf{I}}\boldsymbol{\omega}(\mathbf{I})\ne0$. The equality of $S^H_tS^{I_2}_\varepsilon$ and $S^{I_2}_\varepsilon S^H_t$ implies

2.22
$$\boldsymbol{\varphi}_\varepsilon+\boldsymbol{\omega}(\mathbf{I}_\varepsilon)t=S^{I_2}_\varepsilon\left(\boldsymbol{\omega}(\mathbf{I})t+\boldsymbol{\varphi}\right)\tag{2.21}$$

which cannot be true for all t if $\boldsymbol{\omega}(\mathbf{I})\ne\boldsymbol{\omega}(\mathbf{I}_\varepsilon)$ because the *r.h.s.* differs from $\boldsymbol{\omega}(\mathbf{I})t+\boldsymbol{\varphi}$ by at most $O(\varepsilon)$. Hence $J=0$ and $I_1=H,I_2,I_3$ are independent and in involution so that we can apply the Arnold–Liouville theorem.

Of course one can discuss the more involved cases in which the system is not anisochronous: a deeper analysis including most such cases can be found in Ref. [5].

sec.3

3. Classical perturbation theory.

We now consider a canonically integrable Hamiltonian system described directly in action–angle coordinates by

3.1
$$H_0(\mathbf{A},\boldsymbol{\varphi})=h(\mathbf{A}),\quad(\mathbf{A},\boldsymbol{\varphi})\in V\times\mathbb{T}^\ell\tag{3.1}$$

Let f be an analytic funtion on $V\times\mathbb{T}^\ell$ and consider for $\varepsilon\in\mathbb{R}$

3.2
$$H_\varepsilon(\mathbf{A},\boldsymbol{\varphi})=h(\mathbf{A})+\varepsilon\,f(\mathbf{A},\boldsymbol{\varphi})\tag{3.2}$$

To be more quantitative we shall regard f as a function on $\mathbb{C}^{2\ell}$ by setting

3.3 $$f(\mathbf{a},\mathbf{z}) \equiv f(\mathbf{A},\boldsymbol{\varphi}), \quad \text{if} \quad \mathbf{z}=(e^{i\varphi_1},\ldots,e^{i\varphi_\ell}) \tag{3.3}$$

and real analyticity will be imposed by requiring that f has a holomorphic extension to $\mathbb{C}^{2\ell}$ to the neighborhood

3.4 $$W(\rho,\xi,;V)=\big\{\mathbf{A},\mathbf{z}\,\big|\,(\mathbf{A},\mathbf{z})\in\mathbb{C}^{2\ell},\ \exists\mathbf{A}_0\in V \text{ such that } |a_i-A_{0i}|<\rho, \text{ and } e^{-\xi}<|z_i|<e^{\xi}\cdot\big\}, \tag{3.4}$$

and we shall denote

3.5 $$E=\sup_{W(\rho,\xi,;V)}|\partial_{\mathbf{A}}h(\mathbf{A})|, \qquad \|f\|_{\rho,\xi}=\sup_{W(\rho,\xi,;V)}\big(|\partial_{\mathbf{A}}f|+\rho^{-1}|\partial_{\boldsymbol{\varphi}}f|\big) \tag{3.5}$$

where $|\mathbf{w}|=\sum_{j=1}^{\ell}|w_i|$.

Clearly E^{-1} is a time scale measuring the characteristic time of variation of the unperturbed system, while $(\varepsilon\|f\|_{\rho,\xi})^{-1}E^{-1}$ is the time scale over which the perturbation becomes effective compared to the unperturbed evolution time scale.

The problem of perturbation theory can be stated as follows: find for ε small a one parameter family $\mathcal{C}_\varepsilon$ of completely canonical maps defined on $V\times\mathbb{T}^\ell$ such that $\mathcal{C}_\varepsilon\to$ identity as the parameter $\varepsilon\to0$ and if $(\mathbf{A}',\boldsymbol{\varphi}')=\mathcal{C}_\varepsilon(\mathbf{A},\boldsymbol{\varphi})$

3.6 $$H_\varepsilon(\mathbf{A},\boldsymbol{\varphi})=h_\varepsilon(\mathbf{A}')+o(\varepsilon) \tag{3.6}$$

where Zh_ε is a suitable analytic function of $\mathbf{A}'\in V$. The ideal goal would be to have $o(\varepsilon)=0$: but this can happen only if the perurbed system is canonically integrable by a one-paprameter family of completely canonical maps depending analytically on ε, a rare event as we shall see soon.

To study the above problem one can imagine that $\mathcal{C}_\varepsilon$ is defined via a generating function Φ by

3.7 $$\mathbf{A}=\mathbf{A}'+\partial_{\boldsymbol{\varphi}}\Phi(\mathbf{A}',\boldsymbol{\varphi}), \qquad \boldsymbol{\varphi}'=\boldsymbol{\varphi}+\partial_{\mathbf{A}'}\Phi(\mathbf{A}',\boldsymbol{\varphi}) \tag{3.7}$$

as one can check that any action preserving completely canonical map analytically close to the identity has the form (3.7). Then the perturbation theory solution for the "*Hamilton–Jacobi*" equation, *i.e.* (3.6) with $o(\varepsilon)=0$, is built by setting

3.8 $$\Phi_\varepsilon=\varepsilon\Phi_1+\varepsilon^2\Phi_2+\ldots, \qquad h_\varepsilon=h+\varepsilon h_1+\varepsilon^2h_2+\ldots \tag{3.8}$$

and writing

3.9 $$H_\varepsilon(\mathbf{A}'+\partial_{\boldsymbol{\varphi}}\Phi_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}),\boldsymbol{\varphi})=h_\varepsilon(\mathbf{A}') \tag{3.9}$$

Note that (3.9) is a differential equation for Φ_ε at fixed $\mathbf{A}'$ and this is the reason why the Hamilton–Jacobi equation (which in principle is an equation in which the unknown is a map) is usually thought of as a differential equation.

One then substitutes (3.8) into (3.9) and, after developing both sides in powers of ε, one equates the coefficient of ε^k to 0. If $\boldsymbol{\omega}(\mathbf{A})\stackrel{def}{=}\partial_{\mathbf{A}}h(\mathbf{A})$ one finds

3.10 $$h(\mathbf{A}')+\varepsilon\boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\mathbf{A}'}\Phi_1(\mathbf{A}',\boldsymbol{\varphi})+\varepsilon f(\mathbf{A}',\boldsymbol{\varphi})+O(\varepsilon^2)=h(\mathbf{A}')+\varepsilon h_1(\mathbf{A}')+O(\varepsilon^2) \tag{3.10}$$

leading to the *first order* or *linearized Hamilton–Jacobi equation*:

3.11
$$\boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\mathbf{A}'}\Phi_1(\mathbf{A}',\boldsymbol{\varphi})+f(\mathbf{A}',\boldsymbol{\varphi})=h_1(\mathbf{A}') \tag{3.11}$$

which implies immediately that h_1 is the average over $\boldsymbol{\varphi}$ of f (by integrating bith sides over $\boldsymbol{\varphi}\in\mathbb{T}^\ell$), so that (3.11) can in fact be written as

3.12
$$\begin{aligned} h_1(\mathbf{A}') =&\overline{f}(\mathbf{A}')\stackrel{def}{=}\frac{1}{(2\pi)^\ell}\int_{\mathbb{T}^\ell} f(\mathbf{A}',\boldsymbol{\varphi})\,d\boldsymbol{\varphi},\\ \boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\mathbf{A}'}\Phi_1(\mathbf{A}',\boldsymbol{\varphi}) =&-\big[f(\mathbf{A}',\boldsymbol{\varphi})-\overline{f}(\mathbf{A}')\big]\end{aligned} \tag{3.12}$$

If the second equation of (3.12) admits a solution (which we can determine up to a function of $\mathbf{A}'$, an ambiguity which by convention I shall resolve by requiring $\overline{\Phi}_1=0$) then one can write down the second order equation for Φ, *i.e.* the equation for Φ_2, and in terms of its solution the third order equation, *etc*

Assuming that all such equations admit a solution one finds that the n-th order equation has the form

3.13
$$\boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\mathbf{A}'}\Phi_n(\mathbf{A}',\boldsymbol{\varphi})+N_n(\mathbf{A}',\boldsymbol{\varphi})=h_n(\mathbf{A}') \tag{3.13}$$

where

3.14
$$N_n(\mathbf{A}',\boldsymbol{\varphi})=\Big\{\text{polynomial in }\partial_{\boldsymbol{\varphi}}\Phi_j(\mathbf{A}',\boldsymbol{\varphi}),\ j=1,\ldots,n-1\Big\} \tag{3.14}$$

with coefficients which are monomials in $\partial^{|\mathbf{a}|}_{\mathbf{A}'^{\mathbf{a}}}f(\mathbf{A}',\boldsymbol{\varphi})$ with $|\mathbf{a}|<n$ and $\partial^{|\mathbf{a}|}_{\mathbf{A}'^{\mathbf{a}}}h(\mathbf{A}')$ with $|\mathbf{a}|\le n$.

One checks that for $n\ge2$

3.15
$$\begin{aligned} N^f_n(\mathbf{A}',\boldsymbol{\varphi})=&\sum_{i\le|\mathbf{a}|\le n}\frac{1}{\mathbf{a}!}\partial^{|\mathbf{a}|}_{A'^{\mathbf{a}}}h(\mathbf{A}')\sum_{\substack{k^{(1)}_1,\ldots,k^{(1)}_{a_1}\ge1\\ \ldots\\ k^{(\ell)}_1,\ldots,k^{(\ell)}_{a_\ell}\ge1,\ \sum k^{(i)}_j=n}}\prod_{i=1}^{\ell}\prod_{j=1}^{a_\ell}\partial_{\varphi_j}\Phi_{k^{(i)}_j}(\mathbf{A}',\boldsymbol{\varphi})+\\ &+\sum_{i\le|\mathbf{a}|\le n}\frac{1}{\mathbf{a}!}\partial^{|\mathbf{a}|}_{A'^{\mathbf{a}}}f(\mathbf{A}',\boldsymbol{\varphi})\sum_{\substack{k^{(1)}_1,\ldots,k^{(1)}_{a_1}\ge1\\ \ldots\\ k^{(\ell)}_1,\ldots,k^{(\ell)}_{a_\ell}\ge1,\ \sum k^{(i)}_j=n-1}}\prod_{i=1}^{\ell}\prod_{j=1}^{a_\ell}\partial_{\varphi_j}\Phi_{k^{(i)}_j}(\mathbf{A}',\boldsymbol{\varphi})\end{aligned} \tag{3.15}$$

where $\mathbf{a}=(a_1,\ldots,a_\ell)$, $\mathbf{a}!=a_1!\ldots a_\ell!$, $|\mathbf{a}|=\sum_i a_i$.

Eq. (3.13) yields $h_n(\mathbf{A}')=\overline{N^f_n}(\mathbf{A}')$ and we see that the determination of any order in perturbation heory is equivalent to the study of an equation of the type

3.16
$$\boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\mathbf{A}'}\Phi(\mathbf{A}',\boldsymbol{\varphi})=-\big[N^f(\mathbf{A}',\boldsymbol{\varphi})-\overline{N^f(\mathbf{A}')}\big] \tag{3.16}$$

Before explaining some applications of the above general considerations it is convenient to formulate a more general perturbation theory problem. Namely one can ask whether there is a family $\mathcal{C}_\varepsilon$ of analytic completely canonical maps, analytically dependent on ε for small ε, such that

3.17
$$H_\varepsilon(\mathcal{C}^{-1}_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}'))=h_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}') \tag{3.17}$$

where h_ε has a prescribed $\boldsymbol{\varphi}'$–dependence, *e.g.*

3.18 $$h_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}') = \sum_{\boldsymbol{\nu}\in\pi\subset\mathbb{Z}^\ell} h_{\varepsilon\boldsymbol{\nu}}(\mathbf{A}')\, e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}'} \tag{3.18}$$

with π being a fixed subset of $\mathbb{Z}^\ell$. Ordinary perturbation theory is obtained when $\pi = \mathbf{0}$. Using the notation (3.15) and writing

3.19 $$h_\varepsilon = h(\mathbf{A}') + \varepsilon\, h_1(\mathbf{A}',\boldsymbol{\varphi}') + \varepsilon^2\, h_2(\mathbf{A}',\boldsymbol{\varphi}') + \dots \tag{3.19}$$

one finds the equation for Φ_ε:

3.20 $$\begin{gathered}\boldsymbol{\omega}(\mathbf{A}')\cdot\partial_{\boldsymbol{\varphi}} + N^f_n(\mathbf{A}',\boldsymbol{\varphi}) = \sum_{k=1}^{n-1}\widehat{N^h}_{n-k}(\mathbf{A}',\boldsymbol{\varphi}) + h_n(\mathbf{A}',\boldsymbol{\varphi}) \qquad \text{with}\\ \widehat{N^h}_{n-p} = \sum_{0<|\mathbf{a}|<\le p}\partial^{|\mathbf{a}|}_{\boldsymbol{\varphi}^{\mathbf{a}}} h_{n-p}(\mathbf{A}',\boldsymbol{\varphi}) \sum_{\substack{\sum k^{(i)}_j = p\\ k^{(i)}_j\ge 1}} \prod_{i=1}^{\ell}\prod_{j=1}^{a_\ell}\partial_{A'_j}\Phi_{k^{(i)}_j}(\mathbf{A}',\boldsymbol{\varphi})\end{gathered} \tag{3.20}$$

where $N^f_1 = f$, and $\widehat{N}^h_1 \equiv 0$; or writing $M^f_n = N^f_n = \sum_{k=1}^{n-1}\widehat{N^h_{n-k}}$ and denoting by $P_\pi M^f_n$ the orthogonal projection of M^f_n on the Fourier modes with $\boldsymbol{\nu}\in\pi$:

3.21 $$h_n(\mathbf{A}',\boldsymbol{\varphi}) = P_\pi M^f_n(\mathbf{A}',\boldsymbol{\varphi}), \qquad \boldsymbol{\omega}(\mathbf{A})\cdot\partial_{\boldsymbol{\varphi}}\Phi_n + M^f_n - P_\pi M^f_n = 0 \tag{3.21}$$

where Φ_n is chose, if it exists, suct the $P_\pi\Phi_n = 0$.

We shall now consider two general applications and one concrete example. First consider the case in which f is a trigonometric polynomial of degree N analytic in $W(\rho,\xi;V)$ and:

3.22 $$\begin{gathered} f(\mathbf{A},\boldsymbol{\varphi}) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^\ell,\ |\boldsymbol{\nu}|\le N} f_{\boldsymbol{\nu}}(\mathbf{A})\, e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\\ |\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}|^{-1} \le C\,|\boldsymbol{\nu}|^\alpha, \quad \forall\,|\boldsymbol{\nu}|\le N,\ \boldsymbol{\nu}\ne\mathbf{0}\end{gathered} \tag{3.22}$$

for some $C>0, \alpha>0$. Then given $\rho'<\rho$, $\xi'<\xi$ it is possible to define *for ε small enough* a completely canonical transformation $\mathcal{C}_\varepsilon$ analytic in $W(\rho',\xi';V)$ and in ε such that

3.23 $$\begin{aligned} &(i)\ \mathcal{C}_\varepsilon : W(\rho',\xi';V)\to W(\rho,\xi;V), \qquad \mathcal{C}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \text{identity}\\ &(ii)\ H_\varepsilon(\mathcal{C}_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}')) = h_\varepsilon(\mathbf{A}') + \varepsilon^2 f_1(\mathbf{A}',\boldsymbol{\varphi}',\varepsilon)\\ &(iii)\ h_\varepsilon(\mathbf{A}') = h(\mathbf{A}') + \varepsilon\overline{f}(\mathbf{A}')\end{aligned} \tag{3.23}$$

with f_1 analytic in ε and $(\mathbf{A}',\boldsymbol{\varphi}')\in W(\rho',\xi';V)$.

The above theorem says that in the variables $\mathbf{A}',\boldsymbol{\varphi}'$ the motion is approximated by

3.24 $$\mathbf{A}' = \text{const}, \qquad \boldsymbol{\varphi}' \to \boldsymbol{\varphi}' + \partial_{\mathbf{A}'} h_\varepsilon(\mathbf{A}')\, t \tag{3.24}$$

up to times of order $(O(\varepsilon^{-1})$ for small ε; *i.e.* (3.24) provides a description of the motion valid for a much longer time than could be estimated by neglecting the perturbation directly in the original coordinates (which would give a time of order $O(\varepsilon^{-\frac12}))$.

The proof of the above theorem is simple. One starts from (3.10) and (3.11) and soves it by

3.25
$$\Phi_1(\mathbf{A}',\boldsymbol{\varphi})=\sum_{\boldsymbol{\nu}\ne\mathbf{0},\ |\boldsymbol{\nu}|\le N}\frac{f_{\boldsymbol{\nu}}(\mathbf{A}')}{-i\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}}\,e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\tag{3.25}$$

This is well defined and holomorphic in $W(\rho,\xi;V)$ because $|\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}|^{-1}<C<\infty$ on $W(\rho,\xi;V)$ by hypothesis (see (3.22)).

Given $0<\delta<\xi$ the size of Φ_1 can be estimated in a domain $W(\rho e^{-\delta},\xi-\delta;V)$ by *dimensional analysis*, *i.e.* by Cauchy's theorem. In fact the analyticity asumptions (3.5) imply

3.26
$$\begin{aligned}\big|\partial_{\mathbf{A}}f_{\boldsymbol{\nu}}(\mathbf{A})\big|\;&\le\big(\sup_{W(\rho,\xi;V)}\big|\partial_{\mathbf{A}}f\big|\big)\,e^{-\xi|\boldsymbol{\nu}|}\le\|f\|_{\rho,\xi}e^{-\xi|\boldsymbol{\nu}|},\\ \big|\,|\boldsymbol{\nu}|\,f_{\boldsymbol{\nu}}(\mathbf{A})\big|\;&\le\big(\sup_{W(\rho,\xi;V)}\big|\partial_{\boldsymbol{\varphi}}f\big|\big)\,e^{-\xi|\boldsymbol{\nu}|}\le\;\rho\,\|f\|_{\rho,\xi}e^{-\xi|\boldsymbol{\nu}|}\end{aligned}\tag{3.26}$$

Therefore there is $B_1>0$, a numerical constant dependent on α,ℓ such that

3.27
$$\begin{aligned}&\sup_{W(\rho e^{-\delta},\xi-\delta;V)}|\Phi_1|\le\;C\sum_{\boldsymbol{\nu}\ne\mathbf{0},|\boldsymbol{\nu}|\le N}\frac{|\boldsymbol{\nu}|^{\alpha}\,|\boldsymbol{\nu}|\,|f_{\boldsymbol{\nu}}(\mathbf{A})|}{|\boldsymbol{\nu}|}e^{(\xi-\delta)|\boldsymbol{\nu}|}\le\\ &\le C\,\rho\,\|f\|_{\rho,\xi}\sum_{\boldsymbol{\nu}\ne\mathbf{0}}|\boldsymbol{\nu}|^{\alpha-1}e^{-\delta|\boldsymbol{\nu}|}\le B_1C\,\rho\,\|f\|_{\rho,\xi}\delta^{\ell-\alpha+}\end{aligned}\tag{3.27}$$

and again by dimensional analysis, supposing $0<\delta<\xi<1$

3.28
$$\begin{aligned}&\sup_{W(\rho e^{-2\delta},\xi-2\delta;V)}\Big(|\partial_{\mathbf{A}}\Phi_1|+\frac{1}{e^{-\delta}\rho}\,|\partial_{\boldsymbol{\varphi}}\Phi_1|\Big)\le\;B_2\,C\,\delta^{-\ell-\alpha}\,\|f\|_{\rho,\xi},\\ &\sup_{W(\rho e^{-2\delta},\xi-2\delta;V)}\Big|\partial^2_{\mathbf{A}\boldsymbol{\varphi}}\Phi_1\Big|\le B_3\,C\,\delta^{-\ell-\alpha-1}\,\|f\|_{\rho,\xi}\end{aligned}\tag{3.28}$$

The inequalities (3.28) allow us to discuss in a quantitative way the implicit functions problems necessary to define $\mathcal{C}_\varepsilon,\mathcal{C}_\varepsilon^{-1}$ from the relations

3.29
$$\begin{aligned}\mathbf{A}=&\mathbf{A}'+\varepsilon\,\partial_{\boldsymbol{\varphi}}\Phi_1(\mathbf{A}',\boldsymbol{\varphi}),\qquad\boldsymbol{\varphi}'=\boldsymbol{\varphi}+\varepsilon\,\partial_{\mathbf{A}'}\Phi_1(\mathbf{A}',\boldsymbol{\varphi}),\qquad\text{or}\\ \mathbf{A}=&\mathbf{A}'+\varepsilon\,\partial_{\boldsymbol{\varphi}}\Phi_1(\mathbf{A}',\boldsymbol{\varphi}),\qquad\mathbf{z}'=\mathbf{z}\,\exp\big(i\varepsilon\,\partial_{\mathbf{A}'}\Phi_1(\mathbf{A}',\boldsymbol{\varphi})\big)\end{aligned}\tag{3.29}$$

where $\mathbf{z}\overset{def}{=}(z_1e^{i\boldsymbol{\varphi}_1},\ldots,z_\ell e^{i\boldsymbol{\varphi}_\ell})$.

First it is clear that (3.28) and the analyticity of Φ_1 allow us to solve (3.29) form small ε as

3.30
$$\begin{aligned}&\mathcal{C}_\varepsilon:\mathbf{A}=\mathbf{A}'+\boldsymbol{\mathit{\Xi}}(\mathbf{A}',\mathbf{z}'),\qquad\mathbf{z}=\mathbf{z}'\exp\big(i\boldsymbol{\mathit{\Delta}}(\mathbf{A}',\mathbf{z}')\big),\qquad\text{or}\\ &\mathcal{C}'_\varepsilon:\mathbf{A}'=\mathbf{A}+\boldsymbol{\mathit{\Xi}}'(\mathbf{A},\mathbf{z}),\qquad\mathbf{z}'=\mathbf{z}\exp\big(i\boldsymbol{\mathit{\Delta}}'(\mathbf{A},\mathbf{z})\big),\end{aligned}\tag{3.30}$$

where $\boldsymbol{\mathit{\Xi}},\boldsymbol{\mathit{\Xi}}',\boldsymbol{\mathit{\Delta}},\boldsymbol{\mathit{\Delta}}'$ holomorphic in $W(\rho e^{-2\delta},\xi-2\delta;V)$ and $\mathcal{C}_\varepsilon,\mathcal{C}'_\varepsilon$ mapings such that

3.31
$$\begin{aligned}&W(\rho e^{-3\delta},\xi-3\delta;V)\subset\mathcal{C}_\varepsilon W(\rho e^{-2\delta},\xi-2\delta;V)\subset W(\rho e^{-\delta},\xi-\delta;V),\\ &W(\rho e^{-3\delta},\xi-3\delta;V)\subset\mathcal{C}'_\varepsilon W(\rho e^{-2\delta},\xi-2\delta;V)\subset W(\rho e^{-\delta},\xi-\delta;V),\end{aligned}\tag{3.31}$$

for small ε and with $\mathcal{C}_\varepsilon,\mathcal{C}'_\varepsilon$ analytic in ε and in $W(\rho e^{-2\delta},\xi-2\delta;V)$.

By construction it will also be true that (3.23) holds on $W(\rho e^{-2\delta},\xi-2\delta;V)$ and our result is proved by setting $\rho'=\rho e^{-2\delta},\xi'=\xi-2\delta)$ and recalling that δ was arbitrary.

For future purposes it is convenient to determine explicitly how small ε has to be chosen if δ is given (*i.e.* if $\rho' < \rho, \xi' < \xi$ are given). Once $\boldsymbol{\Xi}, \boldsymbol{\Xi}', \boldsymbol{\Delta}, \boldsymbol{\Delta}'$ exist they must satisfy, on their domain and if $\mathcal{C}_\varepsilon(\mathbf{A}', \mathbf{z}') = (\mathbf{A}, \mathbf{z})$:

3.32
$$\begin{aligned}\boldsymbol{\Xi}_\varepsilon(\mathbf{A}', \mathbf{z}') = \varepsilon \partial_{\boldsymbol{\varphi}} \Phi_1(\mathbf{A}', \mathbf{z}) = \boldsymbol{\Xi}'_\varepsilon(\mathbf{A}, \mathbf{z})\\ \boldsymbol{\Delta}_\varepsilon(\mathbf{A}', \mathbf{z}') = -\varepsilon \partial_{\mathbf{A}'} \Phi_1(\mathbf{A}', \mathbf{z}) = -\boldsymbol{\Delta}'_\varepsilon(\mathbf{A}, \mathbf{z}),\end{aligned} \tag{3.32}$$

so that by the simensional estimates (3.28)

3.33
$$\begin{aligned}|\boldsymbol{\Xi}_\varepsilon|_{\rho e^{-2\delta}, \xi-2\delta},\ |\boldsymbol{\Xi}'_\varepsilon|_{\rho e^{-2\delta}, \xi-2\delta} \le B_2 C \|f\|_{\rho,\xi} \delta^{-\ell-\alpha}\\ |\boldsymbol{\Delta}_\varepsilon|_{\rho e^{-2\delta}, \xi-2\delta},\ |\boldsymbol{\Delta}'_\varepsilon|_{\rho e^{-2\delta}, \xi-2\delta} \le B_2 C\, \rho\, C \|f\|_{\rho,\xi} \delta^{-\ell-\alpha}.\end{aligned} \tag{3.33}$$

Also by simple applications of implicit functions theorems (see for instance Ref. [1], p.490-491) the existence of $\mathcal{C}_\varepsilon, \mathcal{C}'_\varepsilon$) can be guaranteed by requiring the following conditions on the Jacobians:

3.34
$$|\varepsilon \partial^2_{\mathbf{A}\boldsymbol{\varphi}} \Phi_1| < \mu \tag{3.34}$$

where μ is a small absolute constant, *e.g.* 2^{-8}. So using (3.28) we see that there exists $B_4 > 0$ such that if

3.35
$$\varepsilon B4 B_3 C \|f\|_{\rho,\xi} \delta^{-\ell-\alpha-1} < 1 \tag{3.35}$$

the maps $\mathcal{C}_\varepsilon, \mathcal{C}'_\varepsilon$ exist and satisfy (3.33),(3.31); this gives ε in terms of δ.

sec.4
4. Birkhoff theorems on harmonic oscillators.

Another very interesting application of perturbation theory is the following theorem which I shall call Birkhoff theorem (although in the literature this name usually refers to a slightly different theorem, proved in the same way and consequence of the theorem discussed here), see Rfs. [1],p.475. Consider the system

4.1
$$H_\varepsilon(\mathbf{A}, \boldsymbol{\varphi}) = \boldsymbol{\omega} \cdot \mathbf{A} + \varepsilon\, f(\mathbf{A}, \boldsymbol{\varphi}), \quad (\mathbf{A}, \boldsymbol{\varphi}) \in V \times \mathbb{T}^\ell \tag{4.1}$$

with $\boldsymbol{\omega}$ satisfying, for some $C, \alpha > 0$, the *non resonance* or *Dophantine* inequality

4.2
$$|\boldsymbol{\omega} \cdot \boldsymbol{\nu}|^{-1} \le C\, |\boldsymbol{\nu}|^\alpha, \qquad \mathbf{0} \ne \boldsymbol{\nu} \in \mathbb{Z}^\ell. \tag{4.2}$$

Such $\boldsymbol{\omega}$'s exist because the measure of the $\boldsymbol{\omega}$'s in $\mathbb{R}^\ell$ satisfying (4.2) and contained in the set $\{|\boldsymbol{\omega}| < R\}$ can be estimated as

4.3
$$\begin{aligned}&\mathrm{vol}\,(|\boldsymbol{\omega}| \le R) - \sum_{\boldsymbol{\nu}\ne\mathbf{0}} \mathrm{vol}\Big\{\boldsymbol{\omega}\,|\,|\boldsymbol{\omega}| \le R,\, |\boldsymbol{\omega}\cdot\boldsymbol{\nu}| < C^{-1}|\boldsymbol{\nu}|^{-\alpha}\Big\} \ge\\ &\ge \gamma_\ell R^\ell - \gamma_{\ell-1} R^{\ell-1} \sum_{\boldsymbol{\nu}\ne\mathbf{0}} \frac{2}{C\,\boldsymbol{\nu}|^\alpha\,||\boldsymbol{\nu}||} \ge \big(\mathrm{vol}\,(\{|\boldsymbol{\omega}| \le R)\big) \cdot \big(1 - \frac{\overline{\gamma}_\alpha}{CR}\big)\end{aligned} \tag{4.3}$$

where $\|\boldsymbol{\nu}\| = (\sum \nu_i^2)^{\frac12}$ and $\overline{\gamma}_\alpha = 2\frac{\gamma_{\ell-1}}{\gamma_\ell} \sum_{0<|\boldsymbol{\nu}|} |\boldsymbol{\nu}|^{-\alpha} \|\boldsymbol{\nu}\|^{-1} < \infty$, if $\alpha > \ell - 1$. In fact (4.3) shows that the set of the $\boldsymbol{\omega}$'s satisfying (4.2) for some $C > 0$ and $\sigma = \ell$ (say) has full measure in $\mathbb{R}^\ell$.

Eq. (4.1) has the interpretation of a harmonic oscillators with "non resonating" frequencies perturbed by εf and described in the action–angle coordinates of the unperturbed harmonic oscillators. In this case the Birkhoff theorem states that perturbation theory is well defined to *all orders* $n \geq 1$ and the functions $h_n(\mathbf{A}'), \Phi_n(\mathbf{A}', \boldsymbol{\varphi})$ introduced in the preceding section can be simoultaneously defined, for all n, as holomorphic functions in $W(\rho, \xi, V)$.

The reason is simply that if $N(\mathbf{A}', \boldsymbol{\varphi})$ (see (3.16) and(3.13)) is holomorphic in $W(\rho, \xi; V)$ then so is its average $\overline{N(\mathbf{A}')}$ and the solution to (3.16) can be taken to be

4.4
$$\Phi(\mathbf{A}', \boldsymbol{\varphi}) = \sum_{\mathbf{0}\neq\boldsymbol{\nu}\in\mathbb{Z}^\ell} \frac{N_{\boldsymbol{\nu}}(\mathbf{A}')}{-i\boldsymbol{\omega}\cdot\boldsymbol{\nu}} e^{i\boldsymbol{\omega}\cdot\boldsymbol{\nu}} \tag{4.4}$$

the behavior at $\boldsymbol{\nu} \to \infty$ of $N_{\boldsymbol{\nu}}(\mathbf{A}')/\boldsymbol{\omega}\cdot\boldsymbol{\nu}$ does not differ more than polynomially from that of $N_{\boldsymbol{\nu}}(\mathbf{A}')$ which, by the holomorphy assumption is of order $O(e^{-\xi|\boldsymbol{\nu}|})$. Hence φ is also holomorphic in $W(\rho, \xi; V)$ and in fact given $\delta > 0$ satisfies

4.5
$$\begin{aligned} &\max_{W(\rho e^{-\delta}, \xi-\delta; V)} |\Phi| \leq B_\alpha C \delta^{-\ell-\alpha} \max_{W(\rho,\xi;V)} |N|, \\ &\max_{W(\rho e^{-\delta}, \xi-\delta; V)} |\partial_{\mathbf{A}}\Phi| + \frac{1}{\rho e^{-\delta}} |\partial_{\boldsymbol{\varphi}}\Phi| \leq B_\alpha C \delta^{-\ell-\alpha-1}\rho^{-1} \max_{W(\rho,\xi;V)} |N|, \\ &\max_{W(\rho e^{-\delta}, \xi-\delta; V)} |\partial^2_{\mathbf{A}\boldsymbol{\varphi}}\Phi| \leq B_\alpha C \delta^{-\ell-\alpha-2}\rho^{-1} \max_{W(\rho,\xi;V)} |N|, \end{aligned} \tag{4.5}$$

for suitable ℓ–dependent universal constant B_α; (4.5) is proved exactly as (3.27),(3.28), *i.e.* by dimensional estimates. Hence Birkhoff theorem is proved.

A corollary of the quantitative bounds (4.5) is the following: given $n > 0, \delta$, $0 < \delta < \xi$, there exists a one parameter family of analytic completely canonical maps $\mathcal{C}_\varepsilon, \mathcal{C}'_\varepsilon$ defined and analytic also in ε for ε small enough such that $\mathcal{C}'_\varepsilon, \mathcal{C}_\varepsilon \xrightarrow[\varepsilon\to 0]{} 1$ and satisfying

4.6
$$\begin{aligned} (1)\ &\mathcal{C}_\varepsilon, \mathcal{C}'_\varepsilon W(\rho e^{-\delta}, \xi-\delta; V) \supset W(\rho e^{-2\delta}, \xi - 2\delta; V) \\ &\mathcal{C}_\varepsilon, \mathcal{C}'_\varepsilon W(\rho e^{-\delta}, \xi-\delta; V) \subset W(\rho e^{-\frac{1}{2}\delta}, \xi - \frac{1}{2}\delta; V) \\ &\mathcal{C}_\varepsilon\mathcal{C}'_\varepsilon = \mathcal{C}'_\varepsilon\mathcal{C}_\varepsilon = \text{ identity on } W(\rho e^{-2\delta}, \xi-2\delta; V) \end{aligned} \tag{4.6}$$

4.7
$$\begin{aligned} (2)\ &\text{If } (\mathbf{A}, \mathbf{z}) = \mathcal{C}'_\varepsilon(\mathbf{A}', \mathbf{z}'),\ (\mathbf{A}', \mathbf{z}') \in W(\rho e^{-2\delta}, \xi-2\delta; V) \text{ then} \\ &\mathbf{A} = \mathbf{A}' + \partial_{\boldsymbol{\varphi}}\big(\varepsilon\Phi_1 + \ldots + \varepsilon^n\Phi_n)(\mathbf{A}', \boldsymbol{\varphi}) \\ &\boldsymbol{\varphi} = \boldsymbol{\varphi}' + \partial_{\mathbf{A}'}\big(\varepsilon\Phi_1 + \ldots + \varepsilon^n\Phi_n)(\mathbf{A}', \boldsymbol{\varphi}) \end{aligned} \tag{4.7}$$

4.8
$$(3)\ H_\varepsilon(\mathcal{C}_\varepsilon(\mathbf{A}', \boldsymbol{\varphi})) = h(\mathbf{A}') e h_1(\mathbf{A}') + \ldots + \varepsilon^n h_n(\mathbf{A}') + \varepsilon^{n+1} f_{n+1}(\mathbf{A}', \boldsymbol{\varphi}, \varepsilon) \tag{4.8}$$

with f_{n+1} analytic in ε for ε small enough and in $(\mathbf{A}', \boldsymbol{\varphi}) \in W(\rho e^{-2\delta}, \xi - 2\delta; V)$.

This means that a system like (4.1),(4.2) can be approximated by an integrable system up to acertain time scale. This corollary has a simple proof follows from the analysis of the similar theorem in Sect. 3.

If the series, that we call "Birkhoff series",

4.9
$$h_\varepsilon(\mathbf{A}') = \boldsymbol{\omega}\cdot\mathbf{A}' + \sum_{k=1}^\infty \varepsilon^k h_k(\mathbf{A}'), \qquad \Phi_\varepsilon(\mathbf{A}', \boldsymbol{\varphi}) = \sum_{k=1}^\infty \varepsilon^k \Phi_k(\mathbf{A}') \tag{4.9}$$

converge in the domain $W(\rho',\xi';V)$ for some $|r',\xi'>0$ and for ε small enough then the systemwould clearly be integrable for ε small. However such series are not in general convergent. The simplest example is

4.10 $$A_1+\sqrt{2}A_2+\varepsilon\,\big(A_2+f(\varphi_1,\varphi_2)\big) \tag{4.10}$$

In fact one can heck, by the algorithm of perturbation theory explicitly owrked out above, that

4.11 $$\begin{aligned}&\Phi_n(\mathbf{A}',\boldsymbol{\varphi})=\sum_{\boldsymbol{\zeta}\neq\mathbf{0}}f_{\boldsymbol{\nu}}\,\frac{e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{i\boldsymbol{\omega}\cdot\boldsymbol{\nu}}\,\Big(\frac{\nu_2}{\omega_1\nu_1+\omega_2\nu_2}\Big)^{n-1},\\&h_n(\mathbf{A}')=0\quad\text{for } n>1,\qquad h_1(\mathbf{A}')=A_2'\end{aligned} \tag{4.11}$$

where $f_{\boldsymbol{\nu}}$ are the Fourier coefficients of f, which we take to be $f_{\boldsymbol{\nu}}>0$ for all $\boldsymbol{\nu}\neq\mathbf{0}$ and $f_{\mathbf{0}}=0$.

One sees that the series in (4.9) with Φ_n given by (4.11) do not converge: because, if they did, the system would be integrable and therefore its trajectories would be bounded. But the equations of motion of (4.9) are so trivial that they can be easily integrated explicitly and one can see that for many ε's (as small as we please) the Hamiltonian (4.10) generates unbounded motions.

Anintersting extensio of Birkhoff theorem is the "resonant Birkhoff theorem" which deals with systems like (4.1) with $\boldsymbol{\omega}$ satisfying

4.12 $$|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^{-1}<C\,|\boldsymbol{\nu}|^a,\qquad\text{for all } \boldsymbol{\nu}\in\mathbb{Z}^\ell,\ \boldsymbol{\nu}\notin\boldsymbol{\pi} \tag{4.12}$$

where $\boldsymbol{\pi}$ is an hyperplane spanned by $s<\ell$ linearly independent vectors $\boldsymbol{\nu}_1,\boldsymbol{\nu}_2,\ldots,\boldsymbol{\nu}_s\in\mathbb{Z}^\ell$.

In this case one can apply the algorithm developed in Sect. 3 and extendvit to the analysis of the equation

4.13 $$\begin{aligned}&\boldsymbol{\omega}\cdot\big(\mathbf{A}'+\partial_{\boldsymbol{\varphi}}\Phi(\mathbf{A}',\boldsymbol{\varphi})\big)+\varepsilon f\big(\mathbf{A}'+\partial_{\boldsymbol{\varphi}}\Phi(\mathbf{A}',\boldsymbol{\varphi}),\boldsymbol{\varphi}\big)=h_\varepsilon(\mathbf{A}'+\partial_{\mathbf{A}'}\Phi(\mathbf{A}',\boldsymbol{\varphi}))\\&(1-P_{\boldsymbol{\pi}})h_\varepsilon=0,\qquad P_{\boldsymbol{\pi}}\Phi=0\end{aligned} \tag{4.13}$$

where $P_{\boldsymbol{\pi}}$ is the projection of the function $h_\varepsilon(\mathbf{A}',\boldsymbol{\varphi})$ over the linear span of $\{e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\}$, and one can prove that the (4.13) admits a perturbative solution

4.14 $$\Phi=\varepsilon\Phi_1+\varepsilon^2\Phi_2+\ldots,\qquad h_\varepsilon=\boldsymbol{\omega}\cdot\mathbf{A}'+\varepsilon h_1(\mathbf{A}',\boldsymbol{\varphi}')+\varepsilon^2h_2(\mathbf{A}',\boldsymbol{\varphi}')+\ldots, \tag{4.14}$$

with

4.15 $$P_{\boldsymbol{\pi}}\Phi_k=0,\quad (1-P_{\boldsymbol{\pi}})h_k=0,\quad k=1,2,\ldots \tag{4.15}$$

Since the algorithm for a recursive construction of h_n,Φ_n has been explained in the preceding section one just has to check that the n-th order equations for Φ_n,h_n can be solved and this follows from (4.12), see (3.20),(3.21).

Of course the resonant Birkhoff theorem has a corollary similar to the one drawn above from the "ordinary" Birkhoff theorem giving that for a long time of order ε^{-n+1} one can use the Hamiltonian h_ε in (4.14) to approximate the motions.

sec.5 5. Some applications of perturbation theory.

p.5.1 **5.1.** *Precession of Mercury*

The theorms on perturbation theory described in Sect. 3,4 are essentially the sharpest possible in the great generality in which they have been formulated.

It would however be naive to believe that the only thing which is left is to apply them to cases of interest. They only provide a general framework within which the the ideas and techniques to solve real problems can be developed.

As an example of the gap between the above conceptual theorems and some real applications I shall work out the calculation of the precesion of Mercury under the influence of Jupiter unde some simplifying assumptions (needed to avoid devoting too much time to this example). The assmpitons are: Mercury does not influence the motion of the Sun–Jupiter system, the three heavenly bodies are supposed to be on a fixed plane and, finally, the eccentritcity of the motion of Jupiter is zero (circular orbit for Jupiter). This is a rtather restricted but non trivial and interesting three body porblem.

We refer the motion of Mercury to a reference system centered at the Su and with axes fixed in an inertial frame. Denote by $|k$ the gravitational constant and by S, G, M the masses of the Sun, Jupiter and Mercury.

By our hypotheses the aceleration of the SUn is known

5.1
$$\mathbf{a}_S = \kappa G \frac{\boldsymbol{\varrho}_G|}{|\boldsymbol{\varrho}_G|} \tag{5.1}$$

where $\boldsymbol{\varrho}_G$ is the vector joining the Sun to Jupiter. The inertial acceleration of Mercury is $\mathbf{a} = \mathbf{a}_M + \mathbf{a}_S$ where $\mathbf{a}_M$ is the acceleration of Mercury in the Sun–centered reference fram: $\mathbf{a}_M = \ddot{\boldsymbol{\varrho}}_M$, with obvious notations. Hence the equations of motion for $\boldsymbol{\varrho}_M$ are, if $\varepsilon \stackrel{def}{=} \frac{G}{S} \simeq 10^{-3}$,

5.2
$$\begin{aligned} M\,(\ddot{\boldsymbol{\varrho}}_M + \mathbf{a}_M) &= -\kappa S M \frac{\boldsymbol{\varrho}_M}{|\boldsymbol{\varrho}_M|^3} - \kappa G M \frac{\boldsymbol{\varrho}_M - \boldsymbol{\varrho}_G}{|\boldsymbol{\varrho}_M - \boldsymbol{\varrho}_G|^3} \qquad \text{or} \\ \ddot{\boldsymbol{\varrho}}_M &= -\kappa S \frac{\boldsymbol{\varrho}_M}{|\boldsymbol{\varrho}_M|^3} - \varepsilon\kappa S \frac{\boldsymbol{\varrho}_M - \boldsymbol{\varrho}_G}{|\boldsymbol{\varrho}_M - \boldsymbol{\varrho}_G|^3} - \varepsilon \kappa S \frac{\boldsymbol{\varrho}_G|}{|\boldsymbol{\varrho}_G|} \end{aligned} \tag{5.2}$$

which, setting $K = \kappa S$, is recognized to be the Hamilton equation for the Hamiltonian

5.3
$$H(\mathbf{p}, \boldsymbol{\varrho}, T, \varphi) = \frac{\mathbf{p}^2}{2} - \frac{K}{|\boldsymbol{\varrho}|} - \frac{\varepsilon K}{|\boldsymbol{\varrho} - \widehat{\boldsymbol{\varrho}}_G(\varphi)|} + \varepsilon K \frac{\widehat{\boldsymbol{\varrho}}_G(\varphi) \cdot \boldsymbol{\varrho}}{|\widehat{\boldsymbol{\varrho}}_G(\varphi)|^3} + \omega_G T \tag{5.3}$$

where ω_G is the mean angular velocity of Jupiter and $\widehat{\boldsymbol{\varrho}}_G(\varphi)$ is the function expressing the position of Jupiter in terms of its average anomaly, so that the actual position of Jupiter at time t is given by

5.4
$$\boldsymbol{\varrho}_G(t) = \widehat{\boldsymbol{\varrho}}_G(\varphi_0 + \omega_G t) \tag{5.4}$$

for some φ_0. The quantity $\omega_G T$ is the "missing energy"; its variations measure the work done on Mercury by the sustem Jupiter–Sun. One can also call T the "action of the forcing force". If one wanted to avoid the use of the unfamiliar variable T one should use a time–dependent Hamiltonian. However the use of the forcing action variable T andof its conjugate anomaly φ is very useful to treat on the same footing autonomous and periodically forced non autonomous systems.

For $\varepsilon = 0$ the system (5.3) is integrable and has three degrees of freedom (two fo Mercury and one for the forcing term). The first act in applying perturbation theory is to write the Hamiltonian in the action–angle variables of the unperturbed system. In this case we call L, G, ℓ, g the action angle coordinates of the two body problem and we recall that such variables are defined as follows.

$g =$	angle between the major semiaxis, oriented towards the perihelion, and the fixed x–axis
$L =$	$\frac{K}{\sqrt{-2E}}$, with $E =$ energy: $E = -\frac{K^2}{2L^2}$
$G =$	angular momentum $= \rho\dot{\theta}^2 = L\sqrt{1-e^2}$ with
$e =$	eccentricity of the ellipse of Mercury
$\ell =$	$\xi + e\sin\xi$ where ξ is the "eccentric anomaly" of the ellipse if its equation is $\rho = a\,(1 - e\cos\xi)$ with $a =$ major semiaxis
$T_M =$	period, with $\omega_M = \frac{2\pi}{T_M} = \frac{dE}{dL} = \frac{K^2}{L^2}$.

This means that the cartesian coordinates of $\boldsymbol{\varrho}_M$ are

$$\begin{aligned} x =& a\,(1 - e\cos\xi)\,\cos(\theta + g) = \rho\,\cos(\theta + g)\\ y =& a\,(1 - e\cos\xi)\,\sin(\theta + g) = \rho\,\sin(\theta + g)\end{aligned}$$

if ρ, θ are the polar coordinates referred to the perihelion and θ is related to ξ by

5.5 $$a\,(1 - e\cos\xi) = \frac{p}{1 + e\cos\theta}, \qquad p \stackrel{def}{=} \frac{G^2}{K} = a\,(1 - e^2) \tag{5.5}$$

or

5.6 $$(1 - e\cos\xi)(1 + e\cos\theta) = 1 - e^2 \tag{5.6}$$

This shows that

5.7 $$\begin{aligned} \xi =& \ell - e\,\sin\ell + e^2\,c_2(\ell) + e^3\,c_3(\ell) + \ldots\\ \theta =& \ell - e\,\sin\ell + e^2\,b_2(\ell) + e^3\,b_3(\ell) + \ldots\end{aligned} \tag{5.7}$$

Therefore the perturbed Hamiltonian is

5.8 $$H_\varepsilon = -\frac{K^2}{2L^2} + \omega_G T - \varepsilon\,\frac{K}{\big(\rho_G^2 + \rho^2 - 2\rho\rho_G\cos(\theta + g - \varphi)\big)^{\frac{1}{2}}} + \varepsilon\frac{K\,\rho\,\rho_G}{\rho_G^3}\cos(\theta + g - \varphi) \tag{5.8}$$

which can be expanded in terms of Legendre polynomials using

5.9 $$\frac{1}{\big(1 + x^2 - 2xz\big)^{\frac{1}{2}}} = \sum_{k=0}^{\infty} x^n P_n(z) \tag{5.9}$$

yielding

5.10 $$H_\varepsilon = \frac{K^2}{2L^2} + \omega_G T - \varepsilon\,\frac{K}{\rho_G} - \varepsilon\frac{K}{\rho_G}\sum_{n=2}^{\infty}\big(\frac{\rho}{\rho_G}\big)^n P_n(\cos(\theta + g - \varphi)) \tag{5.10}$$

We are first going to neglect all the terms in the series with $n \geq 3$ for the sake of a simple calculation. This causes an error in the acceleration of the prder of $\frac{\varepsilon K}{\rho_G}\big(\frac{\rho}{\rho_G}\big)^3\frac{1}{\rho}$ and hterfore one can estimate the time t_0 over which the error in the position becomes of the order of ρ itself by

5.11 $$\frac{\varepsilon K}{\rho_G}\big(\frac{\rho}{\rho_G}\big)^3\frac{1}{\rho}t_0^2 \simeq \rho \qquad (5.11)$$

Comparing t_0 to $T_M = 2\pi a^{\frac{3}{2}} K$ one finds

5.12 $$\frac{t_0}{T_M} = \Big(\frac{a^3\rho_G}{\varepsilon K}\big(\frac{\rho_G}{a}\big)^3\Big)^{\frac{1}{2}}\frac{K^{\frac{1}{2}}}{2\pi a^{\frac{3}{2}}} = \frac{1}{2\pi}\frac{1}{\sqrt{\varepsilon}}\big(\frac{\rho_G}{a}\big)^2 \simeq 894 \qquad (5.12)$$

which is not really large: but one could easily improve this by taking higher orders in the expansion (5.10). For simplicity, however, we shall take

5.13 $$H_\varepsilon = \frac{K^2}{2L^2} + \omega_G T - \varepsilon\frac{K}{\rho_G}\left(\frac{\rho_G}{a}\right)^2\frac{3}{2}(\cos^2(\theta+g-\varphi)-1) \qquad (5.13)$$

Since $\rho = a\,(1-e\cos\xi)$ (5.13) is still rather complicated because $|x,\theta$ have to be expressed in terms of e,ℓ. We are going to neglect all the terms of the]rhs of(5.13) of order higher than, or equal to, e^5 ($e \simeq .206$ and $e^5 \simeq .3\,10^{-3}$). This will cause an error of order ρ not earlier than t_1, where t_1 is evaluated as t_0 and

5.14 $$\frac{t_1}{t_0} = \Big(\frac{a}{\rho_G}\,\frac{1}{e^5}\Big)^{\frac{1}{2}} \gg 1 \qquad (5.14)$$

So we see that this approximation is acceptable on a time scale $\ll t_0$.

We do not perform the detailed (an unnecessary) calculation and write the result as

5.15 $$H_\varepsilon = \frac{K^2}{2L^2} + \omega_G T - \varepsilon\frac{K}{\rho_G}\left(\frac{\rho_G}{a}\right)^2\big(\pi_0(e) + \gamma_e(\ell,\varphi-g)\big) \qquad (5.15)$$

where π_0, γ_e are fourth order polynomials in e, and γ_e has zero average over ℓ, φ, g:

5.16 $$\gamma_e(\ell,\varphi-g) = \sum_{\nu_1,\nu_2\neq 0}\gamma_e(\nu_1,\nu_2)e^{i(\nu_1\ell+\nu_2(\varphi-g)}, \qquad (5.16)$$

and in fact $|\nu_1|,|\nu_2| \le 4$.

Since for L, e near the unperturbed value we have, for $|\nu_1|,|\nu_2| \le 4$,

5.17 $$\nu_1\omega(L) + \nu_2\omega_G \neq 0, \qquad \omega(L) = \frac{dE}{dL}, \qquad (5.17)$$

because $\omega_g/\omega_M \simeq 49$ we can push perturbation theory to order $\Omega(\varepsilon^2)$ and cAST h_ε into the form (up to m$O(\varepsilon^2)$)

5.18 $$h_\varepsilon(L',G',T') = \frac{K^2}{2L'^2} + \omega_G T - \varepsilon\frac{K}{\rho_G}\left(\frac{a'}{\rho_G}\right)^2\pi_0(e') \qquad (5.18)$$

and in the new coordinates

5.19 $$g' = \partial_{G'}\,h_\varepsilon = -\varepsilon\,\frac{K}{\rho_G}\,\left(\frac{a'}{\rho_G}\right)^2\,\partial_{e'^2}\pi_0(e')\;\partial_{G'}e'^2 \qquad (5.19)$$

Since $g' - g = O(\varepsilon)$ we see that the precession angular velocity is, up to $O(\varepsilon^2)$ and using $e^2 = 1 - \frac{G^2}{L^2} \Rightarrow \partial_G e^2 = -2\frac{1}{L}\frac{G}{L}$, given by

5.20
$$\omega_p=\varepsilon\frac{K}{\rho_G}\left(\frac{a}{\rho_G}\right)^2\partial_{e^2}\pi_0(e)\frac{2G}{L^2} \tag{5.20}$$

So to compute the precession of the perihelion of Mercury within the approximations considered here one only needs to know π_0.

The easiest way to compute π_0 is to take first the average over ℓ, g, φ of the relevant part of (5.10):

5.21
$$\begin{aligned}&-\varepsilon\frac{K}{\rho_G}\sum_{n=2}^{\infty}\big(\frac{\rho}{\rho_G}\big)^n\int(1-e\cos\xi)^n P_n(\cos(\theta+g-\varphi))\frac{d\ell\, d\varphi\, dg}{(2\pi)^3}=\\&=-\varepsilon\frac{K}{\rho_G}\sum_{n=2}^{\infty}\big(\frac{\rho}{\rho_G}\big)^n p(n)\int_0^{2\pi}(1-e\cos\xi)^n\frac{d\ell}{2\pi}\end{aligned} \tag{5.21}$$

where $p(n)=\int_0^{2\pi}P_n(\cos\varphi)\frac{d\varphi}{2\pi}$ and then to neglect terms of order $O(\varepsilon^5)$ after using the relation $\ell=\xi+e\sin\xi$ to compute the integral

5.22
$$\int_0^{2\pi}(1-e\cos\xi)^n\frac{d\ell}{2\pi}=\int_0^{2\pi}(1-e\cos\xi)^n(1+e\cos\xi)\frac{d\xi}{2\pi} \tag{5.22}$$

In this way one avoids the need of expressing ξ,θ in terms of ℓ. The result is, if $[\cdot]^{(4)}$ denotes truncation of the series in ε to fourth order:

5.23
$$\pi_0(e)=\left[\int_0^{2\pi}\frac{d\xi}{2\pi}(1+e\cos\xi)(1-e\cos\xi)^2\right]^{(4)}p(2)=\frac{1}{4}(1-\frac{1}{2}e^2) \tag{5.23}$$

[In fact $\pi_0(e)=\frac{1}{4}(1-\frac{1}{2}e^2)$ exactly because in (5.10) we only consider the $n=2$ term so that the truncation here is not necessary]. So the final result is (to order $\varepsilon^2+e^5\big(\frac{a}{\rho_G}\big)^2+\big(\frac{a}{\rho_G}\big)^3$)

5.24
$$\omega_p=\varepsilon\frac{K}{\rho_G}\left(\frac{a}{\rho_G}\right)^2\frac{3}{4L}\frac{G}{L}=\varepsilon\frac{\sqrt{K}}{\rho_G}\left(\frac{a}{\rho_G}\right)^2\frac{1}{4\sqrt{a}}\sqrt{1-e^2}=\varepsilon\,\frac{\omega_G^2}{\omega_M}\frac{2\sqrt{1-e^2}}{4}=154.67\frac{\text{arcsec}}{\text{century}} \tag{5.24}$$

It is remarkable that in fact one can get easily the expression for ω_p to "any order" in e and $\frac{a}{\rho_G}$ and to order $O(\varepsilon^2)$":

5.25
$$\begin{aligned}\omega_p=&-\varepsilon\frac{K}{\rho_G}\frac{1}{L}\frac{2G}{L}\partial_{e^2}\int_0^{2\pi}\frac{1-e\cos\xi}{\Big(1+\frac{a^2}{\rho_G^2}(1-e\cos\xi)^2-2\frac{a}{\rho_G}(1-e\cos\xi)\cos\varphi\Big)^{\frac{1}{2}}}\frac{d\xi d\varphi}{2\pi}=\\=&-\varepsilon\frac{K}{\rho_G}\frac{1}{L}\frac{2G}{L}\sum_{k=2}^{\infty}\left(\frac{a}{\rho_G}\right)^k p(k)\sum_{h\ge1}e^{2(h-1)}\left[2h\binom{k-1}{2h}-(2h-1)\binom{k-1}{2h-1}\right]\end{aligned} \tag{5.25}$$

however the above expression cannot be taken too seriously because if k becomes large the function $\gamma_e(\ell,\varphi-g)$, see (5.15),will have many more non vanishing Fourier coefficients and the corresponding (5.17) will fail or become too close to 0; in other words one will not gain forever by increasing the precision (*i.e.* the order) of the eccentricity series.

The numerical value of (5.24) justifies the name of the above computation as the "computation of the secular perturbations".[4]

The above analysis shows one of the main difficulties in perturbation theory: small divisors prevent reaching arbitrary precision even if the expansion parameters are small.

I wish to discuss two more corollaries of the ideas involved in perturbation theory as further examples of its applications.

p.5.2 **5.2.** *Generic non integrability*

The first is the following theorem of Poincaré: *"Generically" a Hamiltonian system with Hamiltonian $H_\varepsilon = h(\mathbf{A}) + \varepsilon f(\mathbf{A}, \boldsymbol{\varphi})$ analytic in $(V \times \mathbb{T}^\ell)$ and with $\partial^2_{\mathbf{A}\mathbf{A}} h(\mathbf{A})$ a matrix of rank ≥ 2 is not integrable by a completely canonical map $\mathcal{C}_\varepsilon$ analytic in $\varepsilon, \mathbf{A}, \boldsymbol{\varphi}$ for small ε and tending to the identity as $\varepsilon \to 0$.*

The reason is simply that if integrability is assumed the generating function of the integrating map would have the form $\mathbf{A}' \cdot \boldsymbol{\varphi} + \Phi_\varepsilon(\mathbf{A}', \boldsymbol{\varphi})$ which would have to satisfy

5.26 $$\boldsymbol{\omega}(\mathbf{A}') \cdot \partial_{\boldsymbol{\varphi}} \Phi_1(\mathbf{A}', \boldsymbol{\varphi}) + f(\mathbf{A}', \boldsymbol{\varphi}) = \overline{f}(\mathbf{A}') \tag{5.26}$$

which implies

5.27 $$f_{\boldsymbol{\nu}}(\mathbf{A}') = 0 \qquad \text{if } \boldsymbol{\omega}(\mathbf{A}') \cdot \boldsymbol{\nu} = 0 \tag{5.27}$$

but since there is no relation between $\boldsymbol{\omega}(\mathbf{A})$ and $f(\mathbf{A}, \boldsymbol{\varphi})$ this property will generally not hold so that Φ_1, hence Φ_ε does not exist.

p.5.3 **5.3.** *Non existence of regular constants of motion: Poincaré triviality*

Another application involving roughly the same ideas is that *generically the Hamiltonian in Sect. 5.2 does not admit constants of motion which depend analytically on $(\varepsilon, \mathbf{A}, \boldsymbol{\varphi})$ other than the energy H_ε itself.* I discuss the proof of this well known theorem of Poincaré, [6], in the case in which the matrix $\partial^2_{\mathbf{A}\mathbf{A}} h(\mathbf{A})$ has maximal rank ℓ, *i.e.* $\det \partial^2_{\mathbf{A}\mathbf{A}} h(\mathbf{A}) \neq 0$.

Let $B(\varepsilon, \mathbf{A}, \boldsymbol{\varphi})$ be an analytic constant of motion (or, as Poincaré calls it, a "uniform" constant of motion, meaning that it is single–valued in the non simply connected region $V \times \mathbb{T}^\ell$; analyticity is implicit in Poincaré's terminology in this context):

5.28 $$B(\varepsilon, \mathbf{A}, \boldsymbol{\varphi}) = B_0(\mathbf{A}, \boldsymbol{\varphi}) + \varepsilon B_1(\mathbf{A}, \boldsymbol{\varphi}) + \varepsilon^2 B_2(\mathbf{A}, \boldsymbol{\varphi}) + \ldots \tag{5.28}$$

Since B is a constant of motion $\{B, H_\varepsilon\} = 0$, *i.e.*

5.29 $$\{h(\mathbf{A}), B_0(\mathbf{A}, \boldsymbol{\varphi})\} \equiv \boldsymbol{\omega}(\mathbf{A}) \cdot \partial_{\boldsymbol{\varphi}} B_0(\mathbf{A}, \boldsymbol{\varphi}) \;=\; 0, \qquad \{f(\mathbf{A}, \boldsymbol{\varphi}), B_0(\mathbf{A}, \boldsymbol{\varphi})\} + \{h, B_1\} = 0 \tag{5.29}$$

are the conditions of the zero-th and first order in ε.

[4] A byproduct of the calculation is the "zero eccentricity limit" of the precession given by (5.23) with $e = 0$; to first order in ε one obtains, if $\eta \overset{def}{=} \frac{a}{\rho_G}$,

$$\omega_p = -\varepsilon \frac{K}{\rho_G} \frac{1}{L} \big(\eta^2 \partial_{\eta^2} - \eta \partial_\eta\big) \int \frac{d\varphi}{2\pi} \frac{1}{\sqrt{1 + \eta^2 - 2\eta \cos\varphi}} = -\varepsilon \frac{\omega_G^2}{\omega_M} \frac{1}{\eta} \partial_\eta 4 \int \frac{d\varphi}{2\pi} \frac{1}{\sqrt{1 + \eta^2 - 2\eta\cos\varphi}}$$

valid if a is such that the (analytic) function of a given by $\frac{\omega_G}{\omega_M}$ is such that $\frac{\omega_G}{\omega_M}$ is a Diophantine number.

The first equation implies that B_0 is $\boldsymbol{\varphi}$–independent: $B_0 = B_0(\mathbf{A})$: in fact it must be $\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}\, B_{\boldsymbol{\nu}}(\mathbf{A}) = 0$ but the hypothesis $\det\partial_{|VA}\boldsymbol{\nu}(\mathbf{A}) \ne 0$ implies that, if $\boldsymbol{\nu}\ne\mathbf{0}$ then on a dense set it is $\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}\ne 0$ so that $B_{\boldsymbol{\nu}}(\mathbf{A})\equiv 0$ for $\boldsymbol{\nu}\ne\mathbf{0}$.

Hence the second equation (5.29) yields

5.30 $$\partial_{\mathbf{A}} B_0(\mathbf{A})\cdot\partial_{\boldsymbol{\varphi}} f(\mathbf{A},\boldsymbol{\varphi}) - \boldsymbol{\omega}(\mathbf{A})\cdot\partial_{\boldsymbol{\varphi}} B_1 = 0 \tag{5.30}$$

which implies that $\boldsymbol{\nu}\cdot\partial_{\mathbf{A}} B_0(\mathbf{A})$ and $\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}$ vanish simoultaneously *umless f is very special*, namely unless $f_{\boldsymbol{\nu}}(\mathbf{A}) = \boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}\,\widetilde{f}_{\boldsymbol{\nu}}(\mathbf{A})$. This means that generically $\partial_{\mathbf{A}}(B_0(\mathbf{A})$ and $\boldsymbol{\omega}(\mathbf{A})$ are parallel, *i.e.* for some $\lambda(\mathbf{A})$ it is:

5.31 $$\partial_{\mathbf{A}} B_0(\mathbf{A}) = \lambda(\mathbf{A})\partial_{\mathbf{A}} h(\mathbf{A}) = \lambda(\mathbf{A})\boldsymbol{\omega}(\mathbf{A}) \tag{5.31}$$

hence $\lambda(A) = F'(h(\mathbf{A}))$ and $B_0(\mathbf{A}) = F(h(\mathbf{A}))$ for some F. Then (5.30) implies

5.32 $$B_1(\mathbf{A},\boldsymbol{\varphi}) = f(\mathbf{A},\boldsymbol{\varphi})\,F'(h(\mathbf{A})) + C_1(\mathbf{A}) \tag{5.32}$$

Summarizing: B has the form

5.33 $$\begin{aligned} B(\mathbf{A},\boldsymbol{\varphi}) =& F(h(\mathbf{A})) + \varepsilon F'(h(\mathbf{A}))\,f(\mathbf{A},\boldsymbol{\varphi}) + \varepsilon C_1(\mathbf{A}) + \varepsilon^2 B_1 + \ldots = \\ =& F(h+\varepsilon f) + \varepsilon C_1(\mathbf{A}) + O(\varepsilon^2) = F(H_\varepsilon) + \varepsilon(B_0' + \varepsilon B_1' + \ldots) \end{aligned} \tag{5.33}$$

which implies that $B_0' + \varepsilon B_1' + \ldots$ is another analytic constant of motion.

Repeating the argument we see that also $B_0' + \varepsilon B_1' + \ldots$ must have the form $F_1(H_\varepsilon) + \varepsilon\,(B_0'' + \varepsilon B_1'' + \ldots)$; conclusion

5.34 $$B = F(H_\varepsilon) + \varepsilon\, F_1(H_\varepsilon) + \varepsilon^2 F_2(H_\varepsilon) + \ldots + \varepsilon^n F_n(H_\varepsilon) + O(\varepsilon^{n+1}) \tag{5.34}$$

By analyticity $B = F_\varepsilon(H_\varepsilon(\mathbf{A},\boldsymbol{\varphi}))$ for some F_ε and this completes the proof of Poincaré's triviality theorem (*i.e.* that generically all constants of motion are trivial in the case considered).

sec.6 6. Bounds on time scales of Arnold's diffusion. Nekhorossev theorem.

The unperturbed motions of an integrable system are all confined to tori on which the action variables stay constant. It is natural to ask if a perturbation is such that, even though the perturbed system is not integrable, the motions stay confined near the original unperturbed tori after the perturbation is turned on. There is a remarkable theorem which deals with the above question, "Nekhorossev theorem", [7]:

Suppose that $H_\varepsilon = h(\mathbf{A}) + \varepsilon f(\mathbf{A},\boldsymbol{\varphi})$ and

6.1 $$\partial^2_{\mathbf{A}\mathbf{A}} h(\mathbf{A}) > 0 \qquad \text{for} \mathbf{A}\in V \tag{6.1}$$

where h, f are holomorphic in $W(\mathbf{A},\boldsymbol{\varphi};V)$. Then there exist $A, B, \theta, a, b, \delta > 0$ such that

6.2 $$|\mathbf{A}(t) - \mathbf{A}(0)| \le A\varepsilon^a, \qquad \text{for all } |t| < \theta\, e^{B\varepsilon^{-b}} \tag{6.2}$$

The theorem shows that diffusion in phase space can only take place *very slowly* for small ε; so slowly to be practically unnoticeable for ε small enough.

The assumptions on h are far from optimal as one can realize from the following variant of the theorem (also due to Nekhorossev):

The same conclusions (6.2) hold also in the case

6.3
$$h(\mathbf{A})=\boldsymbol{\omega}\cdot\mathbf{A}\quad\text{and }|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|^{-1}<C|\boldsymbol{\nu}|^{\alpha}\qquad\text{for }\boldsymbol{\nu}\neq\mathbf{0}\tag{6.3}$$

with $C,\alpha>0$.

p.6.1 **6.1.** *Isochronous Nekhorossev estimate*

One of the basic ideas of the theorem emerges clearly in the easier of the above two cases, namely the second. Therefore I start by proving the statement relative to (6.3).

The set V shall be fixed to be the region $\{|\mathbf{A}|\le R\}$. The proof is based on the two Birkhoff theorems of Sect. 4 and, essentially, can be reduced to them.

Let n be an integer and define $\Phi_1,\Phi_2,\ldots,\Phi_n$ and $h_1,h_2,\ldots,h_n$ as described in Birkhoff's algorithm, see (3.15), (3.16), (4.4). Let

6.4
$$\Phi^{(n)}=\varepsilon\Phi_1+\varepsilon^2\Phi_2+\ldots+\Phi_n,\qquad h^{(n)}=\varepsilon h_1+\varepsilon^2h_2+\ldots+h_n,\tag{6.4}$$

see (3.15). Then Φ_k satisfies for some N_ε:

$$\boldsymbol{\omega}\cdot\partial_{\boldsymbol{\varphi}}\Phi_k(\mathbf{A}',\boldsymbol{\varphi})+N_k(\mathbf{A}',\boldsymbol{\varphi})=\overline{N}_k(\mathbf{A}')$$

and by the dimensional estimates of Sect. 3, given $\delta<\xi$ there is $B_0>0$ so that

6.5
$$\Big|\partial_{\boldsymbol{\varphi}}\Phi\Big|_{\rho,\xi-\delta}\le\frac{B_0C}{\delta^{\ell+\alpha+1}}\max_{W(\mathbf{A},\boldsymbol{\varphi};V)}|N|,\qquad|\overline{N}|_\rho\le\max_{W(\mathbf{A},\boldsymbol{\varphi};V)}|N|\tag{6.5}$$

Therefore if we assume, inductively, that for $k=1,\ldots,n$, it is

6.6
$$\Big|\partial_{\boldsymbol{\varphi}}\Phi_k\Big|_{\rho e^{-k\delta},\xi-k\delta}\le A\,B^k\,\delta^{-\beta k}\,k!,\qquad\beta=\ell+\alpha+1\tag{6.6}$$

we shall show that we can derive (6.6) for expression $n+1$ starting from the (3.15).

If $\mathbf{a}!^{-1}\partial^{\mathbf{a}}_{\mathbf{A}'^{\mathbf{a}}}f(\mathbf{A}',\boldsymbol{\varphi})$ is dimensionally bounded by

6.7
$$\Big|\frac{1}{\mathbf{a}!}\partial^{\mathbf{a}}_{\mathbf{A}'^{\mathbf{a}}}f(\mathbf{A}',\boldsymbol{\varphi})\Big|\le\frac{1}{(\rho-\rho e^{-\delta})^{|\mathbf{a}|-1}}\max_{W(\rho,\xi;V)}|\partial_{\mathbf{A}}f|=\frac{1}{(\rho\widehat{\delta})^{|\mathbf{a}|-1}}\|f\|_{\rho,\xi}\tag{6.7}$$

where $\widehat{\delta}=1-e^{-\delta}$, one finds from (3.15), (6.6), (6.5)

6.8
$$\begin{aligned}&\Big|\partial_{\boldsymbol{\varphi}}\Phi_{m+1}\Big|_{\rho e^{-(m+1)\delta},\xi-(m+1)\delta}\le\\&\le\sum_{1\le|\mathbf{a}|\le m}\|f\|_{\rho,\xi}\frac{1}{(\rho\widehat{\delta})^{|\mathbf{a}|-1}}A^{|\mathbf{a}|}B^m\delta^{-m\beta}\cdot\sum_{\substack{k_s^{(r)}\ge1\\ \sum k_j^{(i)}=m}}\big(\prod k_s^{(r)}!\big)B_0C\delta^{-\beta}\le\end{aligned}\tag{6.8}$$

$$\le \|f\|_{\rho,\xi} B_0 C \sum_{p=1}^{m} \delta^{-\beta} 2^{\ell+p} A^p B^m \delta^{-(m+1)\beta} \frac{m!}{(\rho\widehat{\delta})^{p-1}} \le$$
$$\le \|f\|_{\rho,\xi} B_0 C \delta^{-(m+1)\beta} \rho\,\widehat{\delta}\,(m+1)! 2^\ell\, B^m \sum_{p=1}^{m} \left(\frac{2A}{\rho\widehat{\delta}}\right)^p$$

where the step from the first to the second inequality makes use of the relation

$$\sum_{\{j_\sigma\}} \prod_\sigma j_\sigma! \le J! \qquad \text{if } j_\sigma \ge 1 \text{ and } \sum_\sigma j_\sigma = J$$

Hence if $A = \frac{1}{4}\rho\widehat{\delta}$, $B > 4\|f\|_{\rho,\xi} 2^\ell B_0 C$, (6.8) implies

6.9 $$\Big|\partial_{\boldsymbol{\varphi}} \Phi_{m+1}\Big|_{\rho e^{-(m+1)\delta}, \xi-(m+1)\delta} \le A B^{m+1} (m+1)!\, \delta^{-(m+1)\beta} \tag{6.9}$$

Hence (6.6) holds for all k such that $\xi - k\delta > \frac{1}{2}\xi$, provided it holds for $m=0$: the latter relation says

6.10 $$\Big|\partial_{\boldsymbol{\varphi}} \Phi_1\Big|_{\rho e^{-\delta}, \xi-\delta} \le \frac{1}{4}\widehat{\delta}\,\rho\, B\, \delta^{-\beta} \tag{6.10}$$

which if B is taken as $B = 4 2^{\ell-1} B_0 C \widehat{\delta}^{-1} \|f\|_{\rho,\xi}$ is a consequence of (3.28) possibly readjusting B_0.

The calculation]equ(6.8) is in practice a bound on N_{n+1} and therefore on its average over the angles $\overline{N}_{n+1}$': it implies

6.11 $$|\overline{N}_k| \le k! B^k \rho \widehat{\delta}^{-\beta k}\, \widehat{\delta} 2^\ell \|f\| \tag{6.11}$$

We now choose $n = N(\varepsilon)$, $\delta = \frac{\xi}{2n}$ and we realize that in the system of coordinates $(\mathbf{A}', \boldsymbol{\varphi})$ associated with the canonical map generated by $\Phi^{N(\varepsilon)}$ the Hamiltonian is $h^{(N(\varepsilon)}$ with an error of the order of

6.12 $$\varepsilon^{N(\varepsilon)} N(\varepsilon)! \left(\frac{\xi}{2N(\varepsilon)}\right)^{-(\ell+\alpha)N(\varepsilon)} \simeq \left(\frac{1}{2}\xi\varepsilon N(\varepsilon)^{\ell+\alpha+1}\right)^{N(\varepsilon)} \tag{6.12}$$

so that the $\mathbf{A}'$ are constants within a time scale of the order of the reciprocal of the *r.h.s.* of (6.12). And the variables $\mathbf{A}'$ are

6.13 $$\mathbf{A} = \mathbf{A}' + \partial_{\,\boldsymbol{\varphi}} \Phi^{(N(\varepsilon))}(\mathbf{A}', \boldsymbol{\varphi}) \tag{6.13}$$

and, therefore, differ from $\mathbf{A}$ by at most $\max|\partial_{\,\boldsymbol{\varphi}} \Phi^{(N(\varepsilon))}(\mathbf{A}', \boldsymbol{\varphi})|$ which by the preceding bounds is

6.14 $$\max|\partial_{\,\boldsymbol{\varphi}} \Phi^{(N(\varepsilon))}(\mathbf{A}', \boldsymbol{\varphi})| \le A \sum_{k=1}^{N(\varepsilon)} k! (B\widehat{\delta}^{-\beta}\varepsilon)^k \le$$
$$\le \frac{1}{4}\widehat{\delta}\rho \sum_{k=1}^{N(\varepsilon)} \Big[4 2^{\ell-1} B_0 \|f\|_{\rho,\xi} C \big(\frac{\xi}{2}\big)^{-(\ell+\alpha+1)} N(\varepsilon)^{\ell+\alpha+1} \varepsilon N(\varepsilon)\Big]^k \le \tag{6.14}$$
$$\le \frac{\xi\rho}{4N(\varepsilon)} 4\, 2^{\ell-1} B_0 \|f\|_{\rho,\xi} C \big(\frac{2}{\xi}\big)^{\ell+\alpha+1} \varepsilon N(\varepsilon)^{\ell+\alpha+2}$$

provided

6.15 $$4\,N(\varepsilon)2^{\ell-1}B_0\|f\|_{\rho,\xi}C\Big(\frac{2}{\xi}\Big)^{\ell+\alpha+1}\varepsilon N(\varepsilon)^{\ell+\alpha+2}<\frac{1}{2} \tag{6.15}$$

and with the choice $N(\varepsilon)$ such that $\varepsilon N(\varepsilon)^{\ell+\alpha+1}=\sqrt{\varepsilon}$ the (6.2) follows and one can take

6.16 $$a=\frac{1}{2},\qquad b<\frac{1}{2(\ell+\alpha+1)} \tag{6.16}$$

p.6.2 **6.2.** *The anisochronous Nekhorossev theorem*

The case $h(\mathbf{A})=\frac{1}{2}\mathbf{A}^2$ is more interesting because one needs an extra idea which will also clarify the notion of "resonance". Its treatment illustrates also hw to discuss the more general case $\partial^2_{\mathbf{A}\mathbf{A}}h(\mathbf{A})>0$ which is only trivially different. A *resonance of order* $\boldsymbol{\nu}$, $\boldsymbol{\nu}\in\mathbb{Z}^\ell$, for the Hamiltonian $h(\mathbf{A})$ is the surface

6.17 $$\Sigma(\boldsymbol{\nu})=\{\,\mathbf{A}\,|\,|\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}=0\} \tag{6.17}$$

In our paradigmatic case, $h(\mathbf{A})=\frac{1}{2}\mathbf{A}^2$, the surface $\Sigma(\boldsymbol{\nu})$ is simply an hyperplane orthogonal to $\boldsymbol{\nu}$ in $\mathbb{C}^\ell$ because $\boldsymbol{\nu}\omega(\mathbf{A})\equiv\mathbf{A}$. In the following discussion we shall take $V=\{\,\mathbf{A}\,|\,\frac{1}{2}R\le|\mathbf{A}|<R\}$, again for simplicity. Let $N(\varepsilon)=\varepsilon^{-\gamma}$ for some $\gamma>0$ to be fixed later. Let the projection of f on the Fourier modes of order $|\boldsymbol{\nu}|\le N(\varepsilon)$ be

6.18 $$\widehat{f}(\mathbf{A},\boldsymbol{\varphi})=\sum_{|\boldsymbol{\nu}|\le N(\varepsilon)}f_{\boldsymbol{\nu}}(\mathbf{A},\boldsymbol{\varphi})\,e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}} \tag{6.18}$$

We also assume for simplicity that $\xi<1$.

The first remark is that up to a time scale of the order of $e^{-\frac{1}{2}\xi N(\varepsilon)}$ one can replace H_ε by

6.19 $$\widehat{H}_\varepsilon=h(\mathbf{A})+\varepsilon\widehat{f}(\mathbf{A},\boldsymbol{\varphi}) \tag{6.19}$$

because if B is such that $B\xi^{-\ell}\ge\sum_{\boldsymbol{\nu}}e^{-\frac{1}{2}\xi|\boldsymbol{\nu}|}$

6.20 $$\|f-\widehat{f}\|_{\rho,\frac{1}{2}\xi}\le(\max_{W(\rho,\xi;V)}|f|)\sum_{|\boldsymbol{\nu}|>N(\varepsilon)}e^{-\xi|\boldsymbol{\nu}|}\le B\xi^{-\ell}e^{-\frac{1}{2}\xi N(\varepsilon)} \tag{6.20}$$

Given a parameter $\sigma>0$ to be determined later, around each $\Sigma(\boldsymbol{\nu})$, $|\boldsymbol{\nu}|\le N(\varepsilon)$ we consider a "resonant layer" $\Sigma_\varepsilon(\boldsymbol{\nu})$ of width $\varepsilon^\sigma|\boldsymbol{\nu}|^{-\ell-1}$, *i.e.* if $|\boldsymbol{\nu}|=\max_i|\nu_i|$ and $\|\boldsymbol{\nu}\|=(\sum\nu_i^2)^{\frac{1}{2}}$,

6.21 $$\Sigma_\varepsilon(\boldsymbol{\nu})=\Big\{\mathbf{A}\,|\,|\mathbf{A}\cdot\boldsymbol{\nu}|\le\frac{\varepsilon^\sigma\|\boldsymbol{\nu}\|}{|\boldsymbol{\nu}|^{\ell+1}}\Big\} \tag{6.21}$$

It is convenient to regard $\boldsymbol{\nu}=\mathbf{0}$ as a resonance but to define $\Sigma(\mathbf{0})=\Sigma_\varepsilon(\mathbf{0})=\mathbf{0}$. Also we fix $\ell=3$; the case $\ell=2$ is "too easy" while the cases $\ell>3$ can be discussed as the $\ell=3$ case up to minor modifications.

Then the different resonant layers $\Sigma_\varepsilon(\boldsymbol{\nu})$ can only overlap in pairs if σ is large and ε is small and $\boldsymbol{\nu},\boldsymbol{\nu}'$ are not parallel (as our choice of the domain V excludes a ball around the origin, see comment followin (6.17)). In fact the planes $\Sigma(\boldsymbol{\nu})$ and $\Sigma(\boldsymbol{\nu}')$ have only one line in common if $\boldsymbol{\nu}$ is not parallel to $\boldsymbol{\nu}'$; such line will be denoted $\Sigma(\boldsymbol{\nu},\boldsymbol{\nu}')$ and called a "double resonance".

The regions $\Sigma_\varepsilon(\boldsymbol{\nu},\boldsymbol{\nu}')=\Sigma_\varepsilon(\boldsymbol{\nu})\cap\Sigma_\varepsilon(\boldsymbol{\nu})$ are tubes around the double resonance; the tube is cut by the plane generated by $\boldsymbol{\nu},\boldsymbol{\nu}'$ in a small parallelepipedal cross section whose diameter is of the order of ε^{ξ_1} for some ξ_1. This can be seen by a simple geometric argument evaluating the angle between non parallel vactors $\boldsymbol{\nu},\boldsymbol{\nu}'$; since the lengths os such vecr=tors are bounded by $N(\varepsilon)$ and they are on a lattice of mesh 1 their angle θ can be bounded below by $\sim (N(\varepsilon)-1)^{-1}-N(\varepsilon)^{-1}\sim N(\varepsilon)^{-2}$ so that θ can be bounded below by a quantity of order $\varepsilon^{-2\gamma}$. The cross section diameter is then estimated, from Fig.7, as the distance between the points P,P' as $\varepsilon^\sigma/\theta\sim\varepsilon^{\sigma-2\gamma}$ and we can take $\xi_1=\sigma-2\gamma$ if σ (so far arbitrary) is restricted by

6.22
$$\sigma>2\gamma \tag{6.22}$$

Note that the region $\Sigma(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ is *unbounded* if $s<\ell$ while its intersection with the plane $\boldsymbol{\pi}$ is not only bounded but it has a small diameter ($\varepsilon^{\sigma-2\gamma}$).

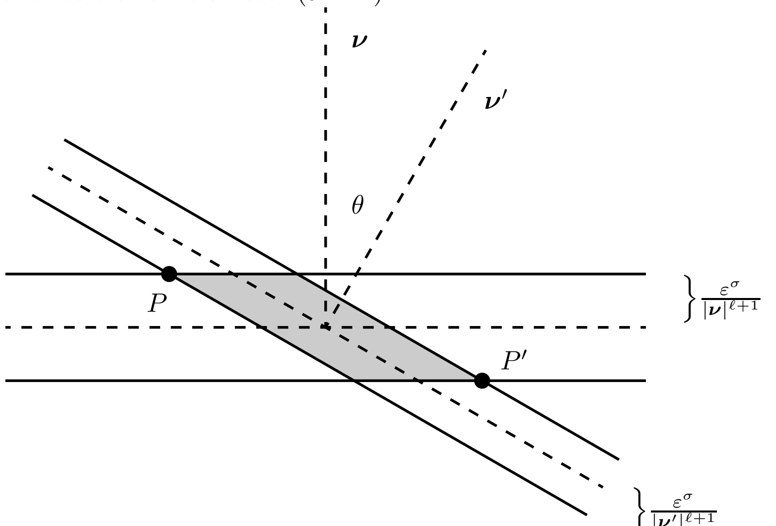


Fig.7 *The cross section of the double resonance between $\boldsymbol{\nu}$ and $\boldsymbol{\nu}'$.*

The distance between any pair of double resonance lines is, since we require $\frac{1}{2}R\le|\mathbf{A}|$, of the order of θ, *i.e.* of order ε^{ξ_2} with $\xi_2=2\gamma$: hence if $\sigma=4\gamma$ and if ε is small enough the double resonance lines (and the corresponding parallelepipedal regions) are disjoint.

We now proceed by considering a region $\Sigma_\varepsilon(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ where the resonance is of order s: $s=0,1,2$ as we are considering the case $\ell=3$. We call $\boldsymbol{\pi}\equiv\boldsymbol{\pi}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\pi}_s)$ the plane spanned by the vectors $\boldsymbol{\nu}_1,\ldots,\boldsymbol{\pi}_s$ which determine the order of the resonance: if $s=0$ we define $\boldsymbol{\pi}=\{\mathbf{0}\}$. In the region $\mathbf{A}\in\Sigma_\varepsilon(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ we shall apply the resonant perturbation theory developed in Sect. 3.

This time we need estimates of $\partial_{\mathbf{A}}\Phi_k$ as well as of h_k which is now defined by (see (3.21))

6.23
$$h_\varepsilon(\mathbf{A}',\boldsymbol{\varphi})=\sum_{\boldsymbol{\nu}\in\boldsymbol{\pi},|\boldsymbol{\nu}|\le N(\varepsilon)}e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}\int\frac{d\boldsymbol{\psi}}{(2\pi)^s}e^{-i\boldsymbol{\nu}\cdot\boldsymbol{\psi}}M_k^f(\mathbf{A}',\boldsymbol{\psi}). \tag{6.23}$$

We assume inductively, for $\rho e^{-2\delta}>\frac{1}{2}\rho$, $]x-k\delta>\frac{1}{2}\xi$ and $\delta N(\varepsilon)^\ell<1$ and $\beta,A,B,D>0$ suitably chosen, that

6.24
$$\Big|\partial_{\boldsymbol{\varphi}}\Phi_k\Big|_{\rho e^{-k\delta},\xi-k\delta}\le AB^k k!\delta^{-\beta k},\qquad \Big|\partial_{\mathbf{A}'}\Phi_k\Big|_{\rho e^{-k\delta},\xi-k\delta}\le \frac{1}{\rho}AB^k k!\delta^{-\beta k},\qquad |h_k|_{\rho e^{-k\delta},\xi-k\delta}|\le Dk!B^k\delta^{-\beta k}\tag{6.24}$$

Proceeding as before one proves the validity of (6.24) for $k=1,\ldots,N(\varepsilon)$ if δ is so small that $N(\varepsilon)\delta$ does ot exceed $\log_2 e, \frac{1}{2}\xi$ and $\delta N(\varepsilon)^\ell<1$ (which still allows us to choose $N(\varepsilon)$ as $\varepsilon^{-\gamma}$ with $\gamma>0$) and one can take

6.25
$$\begin{aligned}A&=const\,\rho\,\delta^{-const}\varepsilon^{-const\,(\sigma+\gamma)}\\ B&=const\,\|f\|_{\rho,\xi}\varepsilon^{-const\,(\sigma+\gamma)}\\ D&=const\,\rho,\qquad \beta=const\end{aligned}\tag{6.25}$$

for suitably chosen constants depending only on the dimension ℓ if the system and on the parameter R (in the presnt case $\ell=3$ and R is kept fixed).

The conclusion is that choosing σ,γ small enough compared to 1 (and proceeding as in the former case) in the region $\Sigma_\varepsilon(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ one can find a change of variables $(\mathbf{A},\boldsymbol{\varphi})\longleftrightarrow(\mathbf{A}',\boldsymbol{\varphi}')$ completely canonical and generated by $\varepsilon\Phi_1+\varepsilon^2\Phi_2+\ldots+\varepsilon^{N(\varepsilon)}\Phi_{N(\varepsilon)}$ which puts the Hamiltonian in the form

6.26
$$h^\Sigma_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}')+O(e^{-const\,\xi})\tag{6.26}$$

and which differs from the identity by a quantity of order $O(\varepsilon^{\xi'})$ with $\xi,\xi'>0$ dependent only on ℓ. A detailed determination of the constants is left to the reader.

To roceed to checking (6.2) as an application of the above analytic considerations the idea is to take, for $(\mathbf{A}',Bf)\in\Sigma_\varepsilon(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ the system as a Hamiltonian system with Hamiltonian given exactly by $h^\Sigma_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}')$, at least as far as the evolution up to times $T=O(e^{+\varepsilon^{-\xi}})$ of the action variables $\mathbf{A}'$ is concerned.

Since the gradient with repect to $\boldsymbol{\varphi}'$ of h^Σ_ε lies, by construction in the plane $\boldsymbol{\pi}$ spanned by $\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s$ we see that as long as $(\mathbf{A}',\boldsymbol{\varphi}')$ stays in Σ_ε the variables $\mathbf{A}'$ can vary only by moving within the cross section cut on $\Sigma(\boldsymbol{\nu}_1)\cap\Sigma(\boldsymbol{\nu}_2)\cap\ldots\cap\Sigma(\boldsymbol{\nu}_s)$ by the plane $\boldsymbol{\pi}$ which by the above discussion has been shown to have a small diameter $O(\varepsilon^{\sigma-2\gamma}$.

It may happen that after a while, earlier than T, the point $(\mathbf{A}',\boldsymbol{\varphi}')$ gets out of the resonant region $\Sigma_\varepsilon(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$. Since the double resonance regions (we now consider the case $\ell=3$) are pairwise disjoint the point has to enter either a single resonance region ($s=1$) or a no resonance region.

In the second case we can use the new form (6.26) with $\Sigma=\mathbf{0}$; in this case $h^\Sigma_\varepsilon(\mathbf{A}',\boldsymbol{\varphi}')$ is $\boldsymbol{\varphi}'$ independent and the point will no longer move before a time which again has length of the order of $T=O(e^{+\varepsilon^{-\xi}})$.

In the first case it will enter a single resonance region, say $\Sigma_\varepsilon(\boldsymbol{\nu})$; and again we could use (6.26) and repeat the argument: either the pair $(\mathbf{A}',\boldsymbol{\varphi}')$ is blocked inside the $\Sigma_\varepsilon(\boldsymbol{\nu})$ with the $\mathbf{A}'$ variables varying at most by $O(\varepsilon^{\sigma-2\gamma}$ (note the in this case the parallelpipedal region is just a one dimensional parallelepiped, *i.e.* a small sement) or, if it gets out of this region it must enter a non resonance region (because the resonances are disjoint) where it will remain until a time of the order of $O(3^{\varepsilon^{-\xi}})$ for some $\xi>0$. To see that the point cannot reenter the double resonance just look at Fig.7 and use that the actions of any point in $\Sigma_\varepsilon(\boldsymbol{\nu})$ moves essentially parallel to $\boldsymbol{\nu}$.

For a less heuristic argument one should use the fact that one can do the above analysis playing with two choices of σ to discuss the "overlap" of the resonances; also one should not confuse $\mathbf{A}$

and $\mathbf{A}'$: however the discussion is basically sraightforward and the reader can consult Ref. [8] (or Ref. [9] for a rigorous discussion along the above lines).

As we shall see implicitly in the discussion, below, about motion near a resonance, see (7.13) below, further assertions can be made. If a point starts in a resonant region of order s then it stays there for a time of order $O(e^{const\varepsilon^{-\sigma}})$ without moving to lower order resonance regions at all: this holds, see Sect. 7, under rather mild extra assumptions.

So the variables $\mathbf{A}'$ can only change by at most $O(\varepsilon^{\sigma-2\gamma})$. Simnce the maps putting the Hamiltonian in the form (6.26) are close to the identity within $O(\varepsilon^{\overline{\xi}})$ with $\overline{\xi}>0$ if γ,σ are small enough compared to 1 (and $\overline{\xi}$ can be evaluated in terms of ℓ,γ,σ) we see that the original $\mathbf{A}$ cannot vary by more than $O(\varepsilon^a)$ for some $a>0$ up to a time $O(e^{const\varepsilon^{-b}})$, for some $b>0$, and the (6.2) is proved.

Therefore we can draw the following picture on the qualitative features of quasi integrable systems. For small ε the system behaves as we should expect: with the $\mathbf{A}$ variables almost constant up to a time of order $O(e^{const\,\varepsilon-b})$; during this time the variables $\mathbf{A}$ may move from resonating regions to other resonating regions of lower order [but as remarked above it could even happen that they stay inside a resonating region of order s during the whole time $O(e^{const\,\varepsilon-b})$]. Staying inside a resonating region of order $s<\ell$ means that the system in suitable coordinates moves as if it was driven by a Hamiltonian like $h^\Sigma_\varepsilon($
$VA',\boldsymbol{\varphi}')$ having only Fourier components in the plane $\boldsymbol{\pi}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_s)$ spanned by the resoooonance vectors. By a linear change of coordinates $(\mathbf{A}',\boldsymbol{\varphi}')\longleftrightarrow(\mathbf{B},\boldsymbol{\psi})$ one can put $h^\Sigma_\varepsilon($
$VA',\boldsymbol{\varphi}')$ into the form

6.27
$$h^\Sigma_\varepsilon(\mathbf{B})+\varepsilon f^\Sigma(\mathbf{B},\psi_1,\ldots,\psi_s), \tag{6.27}$$

i.e. the system moves as if it has $\ell-s$ constants of motion, [9].

After a time of order $O(e^{const\,\varepsilon-b})$ not much can be said about the motion and the point representing the state of the system may wander over much larger distances in phase space. This phenomenon, really occurring in some examples, is called “Arnold diffusion”: for an example see [10].

We shall see as a simple corollary of the KAM theorem (in Sect. 8) that the case $\ell=2$, also covered by the above discussion, has in fact the property that no Arnold diffusion can take place for ε small.

sec.7

7. Resonances and chaos.

We have seen in Sect. 4 that in general one cannot expect that a perturbation of an integrable system (*i.e.* a quasi integrable system in our terminology) cannot be integrable.

Non integrability is closely related to the existence and unavoidability of resonances.

This is clear if $\partial^2_{\mathbf{AA}}$ is non degenerate (anisochronous systems); it is slightly less clear in the case when $h(\mathbf{A})=\boldsymbol{\omega}\cdot\mathbf{A}$ with $\boldsymbol{\omega}$ nonresonating and Diophantine. It is slightly less clear in the case when $h(\mathbf{A})=\boldsymbol{\omega}\cdot\mathbf{A}$ with $\boldsymbol{\omega}$ nonresonating and Diophantine. However in this case it is easy to give explicit counterexamples, see Sect. 3, their analysis shows that perturbation theory may develop resonances which cause non integrability. In fact a formal summation of the divergent series (4.9), (4.11) yields

7.1 $$\Phi_\varepsilon(\mathbf{A}',\boldsymbol{\varphi})=\varepsilon\sum_{\boldsymbol{\nu}\ne\mathbf{0}}\frac{f_{\boldsymbol{\nu}}e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}}{-i(\nu_1\sqrt{2}+(1+\varepsilon)\nu_2)},\qquad h_\varepsilon(\mathbf{A}')=A_1\sqrt{2}+(1+\varepsilon)A_2 \tag{7.1}$$

which is the correct answer only if $\boldsymbol{\omega}_2=\sqrt{2},1+\varepsilon$ is nonresonant and, say, Diophantine; we see that the perturbed system develops resonances for ε as close as wished to 0.

p.7.1 **7.1.** *The resonance confining role of energy conservation.*

The proper way to look at the resonances is to remark that if $\boldsymbol{\omega}\cdot\boldsymbol{\nu}=0$ there are $s<ell$ rationally independent numbers that can be extracted from the components of $\boldsymbol{\omega}$; one can then find a linear change of coordinates (see [9], Sec. 9, lemmas 5,6)

7.2 $$\boldsymbol{\varphi}'=R\,\boldsymbol{\varphi} \tag{7.2}$$

where R has *integer* components such that if $\mathcal{C}$ is

7.3 $$\mathbf{A}'=R^{T^{-1}}\mathbf{A},\qquad \boldsymbol{\varphi}'=R\,\boldsymbol{\varphi} \tag{7.3}$$

in the new coordinates the Hamiltonian has the form $h'(\mathbf{A})$ such that

7.4 $$\partial_{\mathbf{A}'}h(\mathbf{A}')=(\overline{\omega}_1,\ldots,\overline{\omega}_s,0,\ldots,0) \tag{7.4}$$

at the image points of those $\mathbf{A}$'s which satisfy $\boldsymbol{\omega}(\mathbf{A})=\boldsymbol{\omega}$. In other words one can always suppose that a given resonance $\boldsymbol{\omega}(\mathbf{A}_0)$ has the form $(\overline{\omega}_1,\ldots,\overline{\omega}_s,0,\ldots,0)$ with $\overline{\omega}_1,\ldots,\overline{\omega}_s$ rationally independent.

We shall consider here the simple case $s=1$. In this case if in $\mathbf{A}_0$ there is a resonance (*i.e.* $\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}=0$ for some $\boldsymbol{\nu}$ it is not restrictive to suppose

7.5 $$\boldsymbol{\omega}_0(\mathbf{A}_0)=(\omega,\mathbf{0}),\quad \omega\in\mathbb{R},\quad \mathbf{0}\in\mathbb{R}^{\ell-1} \tag{7.5}$$

Near such a resonance the motion of the coordinates $\varphi_2,\ldots,\varphi_\ell$ will be entirely dependent on the perturbation and we cannot expect that the motions which start near the resonance resemble closely, over a long time scale, those of the independent system.

In fact, near the resonances chaotic motion tend to appear (as well as *other* types of ordered motions *i.e.* still quasi periodic but with periods unrelated to those of the unperturbed motion possibly aside from one of them close to $2\pi/\omega$). I will illustrate this phenomenon in detail.

Consider for definiteness a system like

7.6 $$\frac{1}{2}A^2+\frac{1}{2}\mathbf{B}^2+\varepsilon\,f(\varphi_1,\boldsymbol{\varphi}_2,A,\mathbf{B}),\qquad A,\boldsymbol{\varphi}_1\in\mathbb{R}\times\mathbb{T}^1,\quad \mathbf{B},\boldsymbol{\varphi}_2\in\mathbb{R}^{\ell-1}\times\mathbb{T}^{\ell-1} \tag{7.6}$$

and the motions starting near the resonance $(1,\mathbf{0})$. If $\omega=1$ we write these motions as

7.7 $$\begin{aligned}A(t)&=1+\sqrt{\varepsilon}\,a(\sqrt{\varepsilon}\,t),\qquad \varphi_1(t)=\delta(\sqrt{\varepsilon}\,t)\\ \mathbf{B}(t)&=\sqrt{\varepsilon}\,\mathbf{b}(\sqrt{\varepsilon}\,t),\qquad \boldsymbol{\varphi}_2(t)=\boldsymbol{\gamma}(\sqrt{\varepsilon}\,t).\end{aligned} \tag{7.7}$$

where $a(0)=a_0,\mathbf{b}(0)=\mathbf{b}_0,\delta(0)=\delta_0,\boldsymbol{\gamma}(0)=\boldsymbol{\gamma}_0$ are the initial data which we take to be ε–independent. This is a convenient way of looking at motions evolving from data close within $\sqrt{\varepsilon}$ to the resonance.

Substituting (7.7) into the equations of motion one finds

7.8
$$\begin{aligned}
&\varepsilon\,\dot a(\sqrt{\varepsilon}\,t)=-e\,\partial_{\varphi_1}f(\delta(\sqrt{\varepsilon}\,t),\boldsymbol{\gamma}(\sqrt{\varepsilon}\,t),1+\sqrt{\varepsilon}\,a(\sqrt{\varepsilon}\,t),\sqrt{\varepsilon}\,\mathbf{b}(\sqrt{\varepsilon}\,t)),\\
&\varepsilon\,\dot{\mathbf{b}}(\sqrt{\varepsilon}\,t)=-e\,\partial_{\boldsymbol{\varphi}_2}f(\delta(\sqrt{\varepsilon}\,t),\boldsymbol{\gamma}(\sqrt{\varepsilon}\,t),1+\sqrt{\varepsilon}\,a(\sqrt{\varepsilon}\,t),\sqrt{\varepsilon}\,\mathbf{b}(\sqrt{\varepsilon}\,t)),\\
&\sqrt{\varepsilon}\,\dot\delta(\sqrt{\varepsilon}\,t)+\varepsilon\,\partial_A f(\ldots)\\
&\varepsilon\,\dot{\boldsymbol{\gamma}}(\sqrt{\varepsilon}\,t)=\sqrt{\varepsilon}\,\mathbf{b}(\sqrt{\varepsilon}\,t)+\varepsilon\,\partial_{\boldsymbol{\gamma}}(\ldots)
\end{aligned}\tag{7.8}$$

i.e. , writing $\tau=\sqrt{\varepsilon}\,t$,

7.9
$$\begin{aligned}
&\dot a(\tau)=-\partial_{\varphi_1}f(\delta(\tau),\boldsymbol{\gamma}(\tau),1+\sqrt{\varepsilon}\,a(\tau),\sqrt{\varepsilon}\,\mathbf{b}(\tau)),\\
&\dot{\mathbf{b}}(\tau)=-\partial_{\boldsymbol{\varphi}_2}f(\delta(\tau),\boldsymbol{\gamma}(\tau),1+\sqrt{\varepsilon}\,a(\tau),\sqrt{\varepsilon}\,\mathbf{b}(\tau)),\\
&\dot\delta(\tau)=\frac{1}{\sqrt{\varepsilon}}+a(\tau)+\sqrt{\varepsilon}\partial_A f(\delta(\tau),\boldsymbol{\gamma}(\tau),1+\sqrt{\varepsilon}\,a(\tau),\sqrt{\varepsilon}\,\mathbf{b}(\tau)),\\
&\dot{\boldsymbol{\gamma}}(\tau)=\mathbf{b}(\tau)+\sqrt{\varepsilon}\partial_{\mathbf{B}}f(\delta(\tau),\boldsymbol{\gamma}(\tau),1+\sqrt{\varepsilon}\,a(\tau),\sqrt{\varepsilon}\,\mathbf{b}(\tau)),
\end{aligned}\tag{7.9}$$

showing that for ε small δ rotates very quickly (compared to $\boldsymbol{\gamma}$) and implying that the limits for $\varepsilon\to0$ of $a,\mathbf{b},\boldsymbol{\gamma}$ exist at fixed τ and obey, if a overline denotes the average over φ_1, *i.e.* over the "fast angle",

7.10
$$\dot a=-\overline{\partial_{\varphi_1}f}(\boldsymbol{\gamma},1,\mathbf{0})=0,\qquad \dot{\mathbf{b}}=-\overline{\partial_{\boldsymbol{\varphi}_2}f}(\boldsymbol{\gamma},1,\mathbf{0}),\qquad \dot{\boldsymbol{\gamma}}=\mathbf{b}\tag{7.10}$$

this shows that on time scales larger than $1/\sqrt{\varepsilon}$ the $\mathbf{b}$ and $\boldsymbol{\gamma}$ variable nove very differently from the unperturbed case(where $\mathbf{b}=const$ and $\boldsymbol{\gamma}\to\boldsymbol{\gamma}+\sqrt{\varepsilon}\,\mathbf{b}\,\tau$ because their motion is described by the Hamiltonian

7.11
$$\frac12\mathbf{b}^2+\overline{f}(\boldsymbol{\gamma}),\qquad \mathbf{b},\boldsymbol{\gamma}\in\mathbb{R}^{\ell-1}\times\mathbb{T}^{\ell-1}\tag{7.11}$$

where $\overline{f}(\boldsymbol{\gamma})$ is the average of $f(\varphi_1,\boldsymbol{\gamma},1,\mathbf{0})$ over φ_1. Clearly for $\ell=2$ eq. (7.11) describes a double pendulum which has a much richer phase space motion than th unperturbed motion with topologically different types of periodic motions as well as some aperiodic motions). Eq. (7.11) does not describe the motion correctly on time scales long compared to $1/\sqrt{\varepsilon}$: to go futher in time one can no longer eliminate completely the ε dependence of the preturbation.

The Nekhorossev therem allows us to describe the motions up to time scales of arbitrarily high order for data which are within $O(\varepsilon^\sigma)$ of the resonance, where σ cab be taken as small as we please. The time scale over which such motions can be described is $O(e^{const\,\varepsilon^{-b}})$ with b possibly very small (see [9], appendix A, eqs. (A.18) and (A.23) for a more detailed analysis of the latter statement). The motion is described by a Hamiltonian of the form, see Sect. 5)

7.12
$$\frac12A'^2+\frac12\mathbf{B}'^2+\varepsilon\,\widetilde f(\boldsymbol{\varphi}_2',A',\mathbf{B}')+\varepsilon^2f_\varepsilon(\boldsymbol{\varphi}_2',A',\mathbf{B}')+O(\varepsilon^{-const\,\varepsilon^{-b}})\tag{7.12}$$

where $\widetilde f=P_{\boldsymbol{\pi}}f$, $\boldsymbol{\pi}$ being the resonance plane. This means that A' is still a constant of motion, while $A',\mathbf{B}',\varphi_1',\boldsymbol{\varphi}_2'$ are new canonical coordinates related to the priginal ones by a completely canonical map differing from the identity by a quantity of order $O(\varepsilon^{1-\theta})$, where θ can be taken small.

This means that if the initial datum is closer to the resonance than ε^σ, $\sigma < \frac{1}{2}$, *i.e.* it is $A = 1+\sqrt{\varepsilon}\,a_0, \mathbf{B}_0 = \sqrt{\varepsilon}\mathbf{b}_0, \varphi_0 = \delta_0, \boldsymbol{\varphi}_{20} = \boldsymbol{\gamma}_0$ then as long as the motion remains closer to the resonance than $O(\varepsilon^\sigma)$ it will be described by (7.12) or, looking at the motion on scale $1/\sqrt{\varepsilon}$ ($t = \tau/\sqrt{\varepsilon}$) by

7.13
$$\frac{1}{2} + \widetilde{f}(\boldsymbol{\gamma}, \sqrt{\varepsilon}\,a_0, \sqrt{\varepsilon}\,\mathbf{b}) + \varepsilon\, f_\varepsilon(\boldsymbol{\gamma}, \sqrt{\varepsilon}, \sqrt{\mathbf{b}}) + O(\varepsilon^{-const\,\varepsilon^{-b}}) \tag{7.13}$$

Eq. (7.13) allows us to obtain via energy conservation of the motions with Hamiltonian (7.13) an a priori bound on $|\mathbf{b}|$ *in erms of* $\mathbf{b}_0$*: hence we see that (7.13) itself implies that it describes the motion up to times* $O(e^{const\varepsilon^{-b}})$*; i.e. the point does not leave the region of resonance where the motion is described in the coordinates* $A', \mathbf{B}'$ *by (7.12),(7.13).*

The (7.13) immediately explains how chaotic motions can arise near a resonance. In fact $\widetilde{f}(\boldsymbol{\gamma}, \sqrt{\varepsilon}\,a_0, \sqrt{\varepsilon}\mathbf{b})$ can be chosen, for the purpose of producing example, essentially arbitrarily and therefore the motions of the Hamiltonian $\frac{1}{2}\mathbf{b}^2 + \widetilde{f}(\boldsymbol{\gamma})$ can be as complicated as we wish if $\ell - 1 > 1$ on a time scale of $O(1)$ in the variable τ (*i.e.* on atime scale of order $O(1/\sqrt{\varepsilon})$ in the original time units) at least for a time span of $O(e^{const\,\varepsilon^{-b}})$.

p.7.2 **7.2.** *The case* $\ell = 2$ *as an illustration of homoclinic chaos.*

Also if $\ell = 2$ one can understand from (7.13) that chaotic motions arise near a resonance. This time the Hamiltonian describing the motion of the one dimensional variables b, γ is a "pendulum Hamiltonian" (*i.e.* one point mass on a circle subject to a conservative force; therefore its ordered motions are well known. Nevertheless as it becomes clear from (7.1) the pendulum equations change ove a time scale of $O(e^{const\,\varepsilon^{-b}})$ because a_0 changes on this time scale.

The simplest way in which a_0 can change is via a periodic motion of very long period. On can show that a pendulum subject to a periodic force no matter how small exhibits chaotic motions; therefore this can be used to provide an heuristic explanation of how chaos can arise in systems with two degrees of freedom.

The mechamism giving rise to chaotic motions in a forced pendulum is the mechanism of "homoclinic intersections". We describr it in the simple case

7.14
$$\frac{1}{2}b^2 - \cos\gamma - \varepsilon\, b\, \cos\omega\, t, \tag{7.14}$$

with a constant and fixed which represents a forced pendulum:

7.15
$$\ddot{\gamma} + \sin\gamma = \varepsilon\cos\tau\, t \tag{7.15}$$

The dicussion above would indicate that rather than (7.14) one should consider an equation like (7.15) with ω replaced by $\omega_\varepsilon = O(e^{-const\,\varepsilon^{-b}})$. This is a more difficult problem and it will not be considered here; hence the explanation of the onset of chaos for $\varepsilon > 0$ through the discussion of the existence of homoclinic points will have only a heuristic character.

The phase space for (7.15) is three-dimensional but the t–variable appears in a trivial way in (7.15); therefore we shall study this equation, as it is usually done in cases like this, via its "*Poincaré map*" S_ε. This map is defined in the plane (v, q) $v = \dot{\gamma}, \dot{q} = \gamma$) amps (v, q) into the point in which it evolves in the time $t_0 = \frac{2\pi}{\omega}$

7.16
$$S_e(v, q) = (v', q') \tag{7.16}$$

It is easy to descrie the action of S_0. If one draws in the plane (q,v) the solutions of the equations $\ddot q + \sin q, \quad \dot q = v$ the map S_0 maps a point on such a curve into another point on the same curve, see Fig.8.

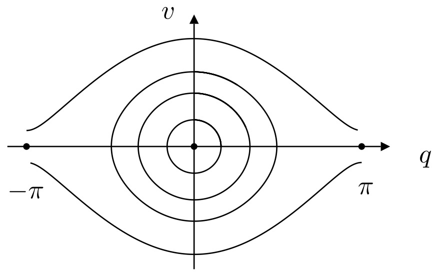


Fig.8 *The invariant curves for the map S_0. The fixed point is $(\pm\pi, 0)$.*

The forced pendulum develops chaotic motions near the separatrices:

7.17
$$v(q) = \xi_\pm = \pm\sqrt{2(1+\cos q)} \tag{7.17}$$

Drawn diffrently we see that Fig.8 days that the point $p_0 = (\pi, 0)$ is a fixed point for S_0 and its stable and unstable manifolds can be drawn more appropriately as in Fig.9 where the q–axis is drawn as a circle (as it really is).

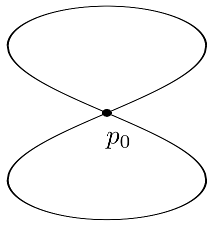


Fig.9 *A representation of the separatrix of p_0.*

In other words the fixed point has a stable and unstable manifolds which merge into each other, symbolically drawn in Fig.10.

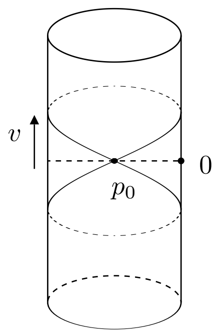


Fig.10 *The phase space of the pendulum as a cylinder with p_0 and the separatrices.*

Such a istuation is very unstable and we shall now show that for $\varepsilon > 0$ the map S_ε still has a fixed point p_ε close to p_0 (by the implicit functions theorem) but that it i s no longer true that the stable and unstable manifolds of p_ε coincide (see Fig.11).

Rather they cross each other at a nonzero angle. Of course, if they cross once they cross infinitely many times. And it would be easy to prove that this implies the existence of chaotic motions near the separatrix for $\varepsilon \geq 0$. Since it is well known that existence of homoclinic points leads to chaotic motions I will not repeat here the arguments.

I now illustrate the technique that can be used to prove existence of one homoclinic point: it is once more an illustration of perturbation theory. It is however of a different kind as it is a perturbation around the separatrix, *i.e.* around a non periodic orbit.

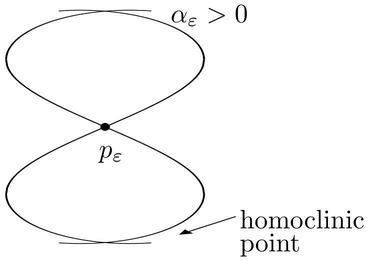


Fig.11 *The splitting pf the stable and unstable manifolds of p_ε for small $\varepsilon \neq 0$.*

The point p_0 "survives" the perturbation by the inplicit functions theorm: the only thing to be checked is that the Jacobian of S_0 in p_0 does not have 1 among its eigenvalues. One way to see this, which in fact gives more information, is that the pendulum motion on the separatrix can be explicitly computed (from the relation $\dot q = \pm\big(2(1+\cos q)\big)^{\frac12}$). Let $Q(t)$ be the motion of the unperturbed pendulum with initial data $q(0)=0, \dot q(0)=2$:

7.18
$$Q(t) = \pi - 4\,\mathrm{arctg}\, e^{-t}. \tag{7.18}$$

and the curve $(\xi_+(q), q)$ that are described by $(\dot Q(t), Q(t))$ is the "upper part" of the (coinciding) stable and unstable manifolds of p_0; the "lower part" is the curve $(\xi_-(q), q)$ described by $(-\dot Q(t), Q(t))$.

Then $S_0(\xi_+(q), q) = (\xi_+(q'), q')$ where $q' = \pi - 4\,\mathrm{arctg}\, e^{-t+2\pi/\omega}$ if $q = \pi - 4\,\mathrm{arctg}\, e^{-t}$. This proves that the two eigenvalues of ∂S_0 in p_0 are

7.19
$$\lambda_\pm = e^{\pm 2\pi/\omega} \tag{7.19}$$

This implies that S_ε will have a hyperbolic fixed point p_ε. Physically this means that the penulum admits a motion in which it stays always near p_e isiting it exactly evert $2\pi/\omega$ units of time.

We shall now study the stable and unstable manifolds of p_ε. On general grounds one can show that the stable (and unstable) manifolds $\xi_\pm(q)$ depend analytically on ε, [6], so that we can write, for the "upper part of the manifold"

7.20
$$\xi_\varepsilon(q) = \xi_+(q) + \varepsilon\,\xi_1(q) + \varepsilon^2\,\xi_2(q) + \ldots, \tag{7.20}$$

where $\xi_1, \xi_2, \ldots$ have to be determined.

The conditions to determine $\xi_1, \xi_2, \ldots$ are:

7.21
$$\begin{aligned} &\text{(i) The set of the points } \{(q, \xi_\varepsilon(q))\} \text{ must be invariant under } S_\varepsilon. \\ &\text{(ii)} \qquad \lim_{n\to\infty} S^n_\varepsilon(q, \xi_\varepsilon(q)) = (q_\varepsilon, v_\varepsilon) \stackrel{def}{=} p_\varepsilon \end{aligned} \tag{7.21}$$

The first condition is imposed by requiring that $(q, \xi_\varepsilon(q))$ and $S^n_\varepsilon(q, \xi_\varepsilon(q))$ lay on the same orbit, solution of

7.22
$$\begin{aligned}\dot v(t) =& -\sin q(t) + e\cos\omega t, \quad q(0) = q\\ \dot q(t) =& v(t), \qquad v(0) = \xi_\varepsilon(q)\end{aligned} \tag{7.22}$$

We write the solution to (7.22) as

7.23
$$q(t) = q_0(t) + \varepsilon\, q_1(t) + \varepsilon^2\, q_2(t) + \ldots \tag{7.23}$$

and we find, by substitution of (7.23) in (7.22):

7.24
$$\begin{aligned}\dot v_0(t) = -\sin q_0(t), \qquad q_0(0) = q\\ \dot q_0(t) = v_0(t), \qquad v_0(0) = \xi_+(q)\end{aligned} \tag{7.24}$$

and

7.25
$$\begin{pmatrix}\dot v_j\\ \dot q_j\end{pmatrix} = \begin{pmatrix}0 & -\cos q_0(t)\\ 1 & 0\end{pmatrix}\begin{pmatrix}v_j\\ q_j\end{pmatrix} + \begin{pmatrix}f_j(t)\\ 0\end{pmatrix}, \qquad j \ge 1 \tag{7.25}$$

where $f_j(t)$ is a function of $q_0(t), \ldots, q_{j-1}(t)$. For instance

7.26
$$f_1(t) = \cos\,\omega\, t, \qquad f_2(t) = \frac{1}{2} q_1(t)^2\, \sin q_0(t) \tag{7.26}$$

an the initial conditions for (7.25) are

7.27
$$v_j(0) = \xi_j(0), \qquad q_j(0) = 0,\ j \ge 1 \tag{7.27}$$

where the ξ_j are unknown. We write (7.25) as

7.28
$$\dot x = L)t)\, x + f \tag{7.28}$$

and we note that $L(t)$ can be written as $L(t) = L_0(t + t_q)$ where t_q is the time needed for the solution of (7.24), *i.e.* (7.18), with initial data $q = 0$ $v = 2$ to reach $\xi_+(q), q)$ and

7.29
$$L_0(t) = \begin{pmatrix}0 & -\cos Q(t)\\ 1 & 0\end{pmatrix} \tag{7.29}$$

where $Q(t)$ is the solution to (7.24) with initial data $q = 0, v = 2$, *i.e.* (7.18).

The solution (7.28) is expressed in terms of the Wronskian, *i.e.* of the solution of the matrix equation

7.30
$$\dot W_0 = L_0(t)\, W_0, \qquad W(0) = 1 \tag{7.30}$$

which allows us to write the Wronskian for (7.28) as

7.31
$$W(t) = W_0(t + t_q) W_0(t_q)^{-1} \tag{7.31}$$

and the solution of (7.25) as

7.32
$$\begin{pmatrix}v_j(t)\\ q_j)t)\end{pmatrix} = W_0(t + t_q) \cdot \left(W_0(t_q)^1 \begin{pmatrix}\xi_j(q)\\ 0\end{pmatrix} + \int_0^t W_0(\tau + t_q)^{-1} \begin{pmatrix}f_j(\tau)\\ 0\end{pmatrix} d\tau \right) \tag{7.32}$$

We are interested in (7.32) computed at $t = nt_0 \equiv mn \cdot 2\pi/\omega$. The further condition (see (ii) above) is to impose that

7.33 $$\begin{pmatrix} v_j(nt_0) \\ q_j(nt_0) \end{pmatrix} \xrightarrow[n\to\infty]{} \begin{pmatrix} v_{\varepsilon,j} \\ q_{\varepsilon,j} \end{pmatrix} \tag{7.33}$$

which means that $(q, \xi_\varepsilon(q))$ is on the stable manifold of p_ε.

To impose the condition in a meaningful way it is necessary to remark that

7.34 $$W_0(t) = \begin{pmatrix} \pi_1(t) & \pi_2(t) \\ \kappa_1(t) & \kappa_2(t) \end{pmatrix} \tag{7.34}$$

where $x_j = \begin{pmatrix} \pi_1 \\ \kappa_1 \end{pmatrix}$ and $x_2 = \begin{pmatrix} \pi_2 \\ \kappa_2 \end{pmatrix}$ solbe $\dot{x} = L_0(t)x$ with initial data $\binom{1}{0}$ and $\binom{0}{1}$, respectively.

Furthermore there is a linear combination x of x_1 and x_2 such that

7.35 $$x(t) = \alpha\, x_1(t) + \beta\, x_2(t) \xrightarrow[t\to\pm\infty]{} 0 \tag{7.35}$$

The latter property might seem hard to believe beause it means that the Schrödinger operator $q \to H\,q$:

7.36 $$H\,q = -\ddot{q} - \cos Q(t)\,q \tag{7.36}$$

admits an eigenvalue E *exactly* $= 0$. . The "potential" $Q(t)$ in (7.36) however has the form shown in Fig.12,

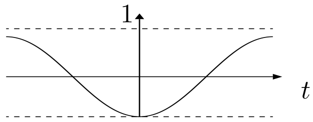


Fig.12 *The graph of* $\dot{Q}(t)$.

and $E = 0$ is an eigenvalue because $Q(t)$ satisfies $\ddot{Q} + \sin Q = 0$, so that $\dot{Q}(t)$ satisfies

7.37 $$\frac{d^2}{dt^2}\dot{Q}(t) + [\cos Q(t)]\,\dot{Q}(t) = 0 \tag{7.37}$$

and form (7.18) one sees that $\dot{Q}(t) = -\frac{2}{\cosh t} \in L_2$. Note that this is not a "miracle" but it is a very general property. If $Q(t)$ satisfies $\ddot{Q} = -\partial_q V(Q)$, then the Schrödinger operator $-\frac{d^2}{dt^2} - \partial^2_{qq} V(Q(t))$ admits $E = 0$ as an eigenvalue, if $Q(t) \in L_2$.

Since $\dot{Q}(t) \neq 0$, $\ddot{Q}(t) \neq 0$, the constants α, β in (7.35) are both non zero. Furhermore $\det W_0 \equiv 1$ because $\mathrm{Tr}\, L_0(t) \equiv 0$ so that

7.38 $$W_0(t)^{-1} = \begin{pmatrix} \kappa_2(t) & -\pi_2(t) \\ -\kappa_1(t) & \pi_1(t) \end{pmatrix} \tag{7.38}$$

Multiply (7.32) written fot $t = n\,t_0 = -n \cdot 2\pi\,\omega^{-1}$ to the left by the vector

3.39 $$(\beta, -\alpha) W_(t + t_q)^{-1} \equiv \begin{pmatrix} \dot{Q}(t + t_q) \\ -\ddot{Q}(t + t_q) \end{pmatrix} \tag{7.39}$$

where α, β are chosen so that (7.35) holds. Then let $n \to \infty$: one finds, using (7.31), (7.32) and (7.33)

7.40
$$0=(\beta,-\alpha)\left(W_0(t_q)^{-1}\begin{pmatrix}\xi_j(q)\\0\end{pmatrix}+\int_0^\infty W_0(\tau+t_q)^{-1}\begin{pmatrix}f_j(\tau)\\0\end{pmatrix}d\tau\right)= \\ =\dot Q(t_q)\xi_j(q)+\int_0^\infty \dot Q(\tau+t_q)f_j(\tau)\,d\tau \tag{7.40}$$

which determines $\xi_j(q)$:

7.41
$$\xi_j(q)=\int_0^\infty \frac{\dot Q(t_q+\tau)}{\dot Q(t_q(\tau)}\,f_j(\tau)\,d\tau \tag{7.41}$$

(note that $\dot Q(t_q)\ne 0$ from (7.18)).

A similar argument alllows us to determine the equations for the unstable manifold of p_ε:

7.42
$$\widetilde\xi_j(q)=-\int_0^{-\infty}\frac{\dot Q(t_q+\tau)}{\dot Q(t_q)}\,f_j(\tau)\,d\tau \tag{7.42}$$

We now study $\xi_1,\widetilde\xi_1$, *i.e.* the stable and unstable maifolds to first order in ε. If they cross to first order in ε and in a transverse way, *i.e.* if

7.43
$$\xi_1(q)=\widetilde\xi_1(q) \tag{7.43}$$

admits a solution q with

7.44
$$\xi_1'(q)\ne\widetilde\xi_1'(q) \tag{7.44}$$

where the prime denotes q–differentiation. It is clear, by the implicit functions theorem, that there is a homoclinic point whose location is determined to first order in ε by (7.43). It is, in fact,

7.45
$$\begin{aligned}&\xi_1(q)-\widetilde\xi_1(q)=\int_{-\infty}^\infty\frac{\dot Q(\tau+t_q)}{\dot Q(t_q)}\cos\omega\tau\,d\tau=\int_{-\infty}^\infty\frac{\dot Q(\tau)}{\dot Q(t_q)}\cos\omega(\tau+t_q)\,d\tau=\\&=\cos\tau t_q\int_{-\infty}^\infty\frac{\dot Q(\tau)}{\dot Q(t_q)}\cos\omega\tau\,d\tau, \qquad\text{(a)}\\&\frac{d}{dq}(\xi_1(q)-\widetilde\xi_1(q))=-t_q'\frac{\ddot Q(t_q)}{\dot Q(t_q)^2}\int_{-\infty}^\infty\cos\omega t_q\,\cos\omega\tau\,\dot Q(\tau)\,d\tau-\\&-t_q'\,\omega\,\sin\omega t_q\int_{-\infty}^\infty\frac{\dot Q(\tau)}{\dot Q(t_q)}\cos\omega\,\tau\,d\tau \qquad\text{(b)}\end{aligned} \tag{7.45}$$

where in the last step in (7.45) we used that $\dot Q_0(\tau)=-\frac{2}{\cosh\tau}$ is even in τ. Since

7.46
$$\int_{-\infty}^\infty\dot Q(t)\,\cos\omega t\,dt=-2\int_{-\infty}^\infty\frac{\cos\omega t}{\cosh t}\,dt\;=\;-\frac{2\pi}{\cosh\frac{\pi}{2}\omega}\ne 0 \tag{7.46}$$

we see that the equation for the homoclinic points, to frst order in ε, is

7.47
$$\cos\omega t_q=0\qquad i.e.\;\;\omega^{-1}\frac{(2k+1)\pi}{2}=t_q \tag{7.47}$$

and in such points

7.48
$$\xi_1'(q) - \widetilde{\xi}_1'(q) = \pm t_q'\,\omega \int_{-\infty}^{\infty} \frac{\dot Q(\tau)}{\dot Q(t_q)}\,\cos\tau\, d\tau \neq 0 \tag{7.48}$$

because $t_q \neq 0$ implies $t_q' \neq 0$.

This completes the proof of the existence of homoclinic points in the forced pendulum; the technique used is of course much more general and can be applied to a variety of similar problems.

sec.8

8. Non resonant invariant tori. The KAM theorem.

In the preceding sections we have seen how integrability is a rare event in quasi integrable systems and how this fact is related to the resonances which inevitably are present in the unperturbed system or are generated by the perturbation itself.

The theory of resonance in Sect. 6,7 sheds some light on the onset of chaotic motions near resonances, but it remains to understand in some deeper way what happens "away" from the resonances.

A typical question isthe following: given an unperturbed system described in action–angle variables by $h_0(\mathbf{A})$, $\mathbf{A} \in V$, let $\mathbf{A}_0$ be such that $\boldsymbol{\omega}_0(\mathbf{A}_0) = \frac{\partial}{\partial \mathbf{A}} h_0(]VA_0)$ has the property

8.1
$$|\boldsymbol{\omega}(\mathbf{A}_0)\cdot\boldsymbol{\nu}|^{-1} < C_0\,|\boldsymbol{\nu}|^{\ell}, \qquad \forall\,\boldsymbol{\nu} \neq \mathbf{0} \tag{8.1}$$

i.e. let $\{\mathbf{A}_0\} \times \mathbb{T}^\ell$ be a torus in phase space which is traversed quasi–periodically and in a nonresonant way by the unperturbed motion: then one asks whether the perturbed Hamiltonian system

8.2
$$H(\mathbf{A},\boldsymbol{\varphi}) = h_0(\mathbf{A}) + \varepsilon\, f(\mathbf{A},\boldsymbol{\varphi}) \tag{8.2}$$

admits an invariant torus still traversed quasi–periodically with the same frequencies and close to the unperturbed torus.

This means asking whether there are two functions $\boldsymbol{\alpha}_\varepsilon, \boldsymbol{\beta}_\varepsilon$ on $\mathbb{T}^\ell$ such that the set

8.3
$$\mathbf{A} = \mathbf{A}_0 + \boldsymbol{\alpha}_\varepsilon(\boldsymbol{\psi}), \qquad \boldsymbol{\varphi} = \boldsymbol{\psi} + \boldsymbol{\beta}_\varepsilon(\boldsymbol{\psi}), \qquad \boldsymbol{\psi} \in \mathbb{T}^\ell \tag{8.3}$$

is a torus on which motion is represented by $\boldsymbol{\psi} \to \boldsymbol{\psi} + \boldsymbol{\omega}_0(\mathbf{A}_0)\,t$ and furthermore $\|\boldsymbol{\alpha}_\varepsilon\|, \|\boldsymbol{\beta}_\varepsilon\| \xrightarrow[\varepsilon\to 0]{} 0$.

Of course one needs some assumptions on h_0; *e.g.* if $h_0(\mathbf{A}) = \boldsymbol{\omega}_0 \cdot \mathbf{A}$ it is clear that one cannot expect the above properties to hold; the unperturbed syste having only one set of frequencies may well change them under perturbation. For instance the perturbation $\varepsilon\,\widetilde{\boldsymbol{\omega}} \cdot \mathbf{A}$ leaves the system integrable but changes its frequencies everywhere in phase space from $\boldsymbol{\omega}_0$ to $\boldsymbol{\omega}_0 + \varepsilon\widetilde{\boldsymbol{\omega}}$ so that no perturbed mtion will have the unperturbed frequencies although all motions will be trivially quasi periodic.

A naturalcondition is therefore the anisochrony condition, also called *twist condition*:

8.4
$$\det \partial^{\,2}_{\mathbf{A}\,\mathbf{A}} h_0(\mathbf{A}) \neq 0 \tag{8.4}$$

guaranteeing that the unperturbed system has an ℓ–dimensional continuum of frequencies, *i.e.* that if $\boldsymbol{\omega}$ is a set of frequencies for the unperturbed system (equivalently if there is a $\mathbf{A}_0 \in V$ such that $\boldsymbol{\omega}(\mathbf{A}_0) = \boldsymbol{\omega}$) then any $\boldsymbol{\omega}'$ is, if $|\boldsymbol{\omega} - \boldsymbol{\omega}'|$ is small enough, also a set of frequencies for the unperturbed system.

The theorem of Kolmogorov-Arnold-Moser states the following

Let h_0, f_0 be holomorphic in $W(\rho_0, \xi_0; \{\mathbf{0}\})$ and let

8.5
$$\begin{aligned} E_0 &\geq \sup_{W(\rho_0,\xi_0;\{\mathbf{0}\})} \big|\partial_{\mathbf{A}} h_0(\mathbf{A})\big|, \\ \eta_0 &\geq \sup_{W(\rho_0,\xi_0;\{\mathbf{0}\})} \big|\partial^2_{\mathbf{A}\,\mathbf{A}} h_0(\mathbf{A})^{-1}\big|, \\ \varepsilon_0 &\geq \sup_{W(\rho_0,\xi_0;\{\mathbf{0}\})} \big|\partial_{\mathbf{A}} f(\mathbf{A},\boldsymbol{\varphi})\big| + \frac{1}{\rho_0}\big|\partial_{\boldsymbol{\varphi}} f(\mathbf{A},\boldsymbol{\varphi})\big| \end{aligned} \tag{8.5}$$

and, for

8.6
$$\boldsymbol{\omega}_0(\mathbf{0})\cdot\boldsymbol{\nu}|^{-1} \leq C_0\,|\boldsymbol{\nu}|^{\ell}, \qquad \forall\,|\boldsymbol{\nu}|>0 \qquad (8.6).$$

Then the exist $\boldsymbol{\alpha},\boldsymbol{\beta}$ analytic on $\mathbb{T}^\ell$ and with values in $\mathbb{R}^\ell$ or, respectively, $\mathbb{R}^\ell$ such that the torus $\mathcal{T}_{\boldsymbol{\omega}_0}$ defined by

8.7
$$\mathbf{A} = \boldsymbol{\alpha}(\boldsymbol{\psi}), \quad \boldsymbol{\varphi} = \boldsymbol{\psi} + \boldsymbol{\beta}(\boldsymbol{\psi}), \qquad \boldsymbol{\psi}\in\mathbb{T}^\ell \tag{8.7}$$

is invariant under the motion generated by $H_0(\mathbf{A}) + f_0(\mathbf{A},\boldsymbol{\varphi})$ and the motion on $\mathcal{T}_{\boldsymbol{\omega}_0}$ is simply, if $\boldsymbol{\omega}_0 \stackrel{def}{=} \boldsymbol{\omega}_0(\mathbf{0})$,

8.8
$$\boldsymbol{\psi} \,\to\, \boldsymbol{\omega}_0\, t \tag{8.8}$$

provided ε_0 is small enoguh; more precisely if

8.9
$$B\,(\varepsilon_0 C_0)\,(E_0\eta_0\rho_0^{-1})^{\alpha_1}\,(E_0 C_0)^{\alpha_2}\xi_0^{-\alpha_3} \,<\, 1 \tag{8.9}$$

where $\xi_0 < 1$ is assumed (for simplicity) and $B, \alpha_1, \alpha_2, \alpha_3$ are constants depending only on the dimension ℓ.

The condition (8.9) is easy to interpret: it says that $(\varepsilon_0 C_0)$ has to be small compared to the quantities $(E_0\rho_0^{-1}\eta_0)^{-1}, \xi_0$ and $(E_0C_0)^{-1}$; the latter four numbers are the only dimensionless quantities that can be formed starting from the only parameters introduced to formulate the theorem, namley $E_0, \eta_0, \rho_0, \xi_0, \varepsilon_0$. Therefore it is clear that *if a theorem like the above holds for ε_0 small, the condition must necessarily take the form* (8.9) (leaving aside the problem of finding an optimal condition). In fact it is easy to see, from the holomorphy assumptiion, that $(E_0\rho_0^{-1}\eta_0)^{-1} > 1$ and $E_0C_0 > 1$ (see (8.6) with $\boldsymbol{\nu}$ a unit vector) and that the theorem cannot be true if $(E_0\eta_0\rho_0^{-1}) = \infty$ or if $E_0C_0 = \infty$ or if $\xi_0 = 0$.

The proof of the above theorem was provided by Kolmogorov in 1955 at the same time when a paper by Fermi, Pata and Ulam appeared in which the gave numerical evidence that perturbing integrable systems one would in general obtain systems which in some sense behaved as if they had several constants of motion, against a rather widespread belief anmong physiscists (inherited from certain views in the fiundations of Statistical Mechanics).

It has been questioned whether the Kolmogorov proof was really mathematically complete: it is very short and sketchy on some points. However one can fill in the details as it has been explicitly shown in Ref. [11], and one is left to pecultate whether Kolmogorov had understood such details or not. In my opinion there is no question that Kolmogorov's proof is complete in the analytic case he considered.

Later Arnold presented a new proof of the theorem by a rather different method; this method is the one by which I let myself be inspired in this section and, in modern laanguage, it is a *renormalization group method*. Arnold was able to use his method to modify the theorem so that it could be applied to the three body problem proving for the first time that gravitation is in principle compatible with stable planetary systems. AT the same time Moser discovered a third method of proof which allowed him to treat the (harder) differentiable case. This method has also the advantage of providing much better numerical estimates (as shown by Rüssmann, [12],[13],[14]. For more recent developments see also the works of Hermann and De la Llave, [15],[16].

It should be clear from the discussion below that the proof of the theorem is just a clever organization of all the results and ideas discussed in the preceding sections on perturbation theory. It really seems that the hardest part of the theorem, at least in the analytic cases, and its originality is in its formulation; but this is not meant to diminish the importance pf Kolmogorov's work. On the contrary it shows that sometimes deep results are not technically hard if placed in the right cultural context; however they require very deep and daring intuition.

It also shows that one should not be too much influenced by *no go* theorems: in Mechanics the theorem of triviality by Poincaré of Sect. §5 seems to have been in some respects an obstacle, for half a century, to the understanding of the nonperturbative meaning of perturbation theory in Mechanics. In fact it turns out that the KAM theorem is a nnperturbative result which is proved by using perturbation theory whose divergence was proved by Poincaré. Here the analogy with constructive field theory in superrinormalizable and renormalizable asymptotically free theories is manifest (I believe): the renormalization group approach allows us to use (divergent) perturbation analysis to construct the theories (see the review paper [17], and the lectures by Gawedski-Kupiainen in this volume). The proof is as follows (from Ref. [1], p.494, and Ref. [18]).

p.8.1 **8.1.** *Definition of a renormalization transformation $\mathcal{T}$ on the Hamiltonians.*

One starts by changing coordinates $(\mathbf{A},\boldsymbol{\varphi})\to(\mathbf{A}',\boldsymbol{\varphi}')$ in a completely canonical way via a canonical map $\mathcal{C}_0$ defined in the vicinity of the unperturbed torus $\mathbb{T}_{\boldsymbol{\omega}_0}=\{\mathbf{0}\}\times\mathbb{T}^\ell$ in such a way that in the new coordinates H_0 has the form

8.10
$$\widetilde{h}_1(\mathbf{A}')+\widetilde{f}_1(\mathbf{A}',\boldsymbol{\varphi}') \tag{8.10}$$

with $\widetilde{h}_1,\widetilde{f}_1$ holomorphic in $W(\widetilde{\rho}_0,\widetilde{\xi}_0;\{\mathbf{0}\})$ and $\widetilde{f}_1$ of size $\sim\varepsilon_0^2$, $\widetilde{h}_1(\mathbf{A}')=h_0(\mathbf{A}')+f_{\mathbf{00}}(\mathbf{A}')$ with $f_{\mathbf{00}}(\mathbf{A}')$ defined as the average $\overline{f}_0(\mathbf{A}',\boldsymbol{\varphi})$ over $\boldsymbol{\varphi}$ of $f_0(\mathbf{A}',\boldsymbol{\varphi})$.

This step is achieved simply by perturbation theory by defining $\mathcal{C}_0$ via the generating function Φ_0 which is the solution of the *linearized* Hamolton–Jacobi equation

8.11
$$\boldsymbol{\omega}_0(\mathbf{A}')\cdot\frac{\partial\,\Phi_0}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})+f_0(\mathbf{A}',\boldsymbol{\varphi})-\overline{f}_0(\mathbf{A}')\,=\,0 \tag{8.11}$$

Of course for the reasons amply illustrated many time in the previous sections the later equation is not soluble because, precisely, of the anisochrony condition $\eta_0<+\infty$ and the fat that, consequently, $\boldsymbol{\omega}_0(\mathbf{A}')\cdot\boldsymbol{\nu}$ will vanish too often (near aby $\mathbf{A}'$ there will be pairs $\mathbf{A},\boldsymbol{\nu}$ with $\boldsymbol{\nu}\neq\mathbf{0}$ such that $\boldsymbol{\omega}(\mathbf{A})\cdot\boldsymbol{\nu}=0$).

Nevertheless one can easily circumvent the difficulty by recalling the way perturbation theory is applied in Astronomy or in other applications seen above. Instead of (8.11) one should solve

8.12
$$\boldsymbol{\omega}_0(\mathbf{A}')\cdot\frac{\partial\,\Phi_0}{\partial\boldsymbol{\varphi}}(\mathbf{A}',\boldsymbol{\varphi})+f_0^{[\le N_0]}(\mathbf{A}',\boldsymbol{\varphi})-\overline{f}_0(\mathbf{A}')\;=\;0 \tag{8.12}$$

where $f_0^{[\le N_0]}(\mathbf{A},\boldsymbol{\varphi})\stackrel{def}{=}\sum_{|\boldsymbol{\nu}|\le N_0}f_{\boldsymbol{\nu}}(\mathbf{A})\,e^{i\boldsymbol{\nu}\cdot\boldsymbol{\varphi}}$ and the *ultraviolet cut–off* N_0 is so chosen that

8.13
$$\sup_{W(\rho e^{-\delta_0},\xi_0-\delta_0;\{\mathbf{0}\})}\Big(\Big|\frac{\partial f^{[>N_0]}}{\partial\mathbf{A}}\Big|+\frac{1}{\rho_0}\Big|\frac{\partial f^{[>N_0]}}{\partial\boldsymbol{\varphi}}\Big|\Big)\;\le\;\varepsilon_0^2C_0 \tag{8.13}$$

where $f_0^{[>N_0]}\stackrel{def}{=}f_0-f_0^{[\le N_0]}$ and $\delta_0>0$ is conveniently chosen ($\delta_0\ll\xi_0$ will be fixed later when its role in the proof becomes clearer); the *r.h.s.* in (8.13) is chosen "arbitrarily" to be of $O(\varepsilon_0^2)$ and C_0 is introduced to fix the dimensions. E_0^{-1} could, of course, replace C_0 for the same purpose, but I hope that the reader will agree that the recurrence time scale C_0 is more significantly compared with the perturbation time scale ε_0^{-1} than the *free time scale* E_0^{-1}. The constant δ_0 has to be introduced because, as we shall soon see, we cannot estimate $f_0^{[>N_0]}$ in the whole domain $W(\rho,\xi_0;\{\mathbf{0}\})$.

The advantage of (8.12) with respect to (8.11) is that $f_0^{[\le N_0]}$ is a trigonometrical polynomial; hence (8.12) can be solved in the region where $\boldsymbol{\omega}_0(\mathbf{A}')\cdot\boldsymbol{\nu}\neq\mathbf{0}$ for all $0<|\boldsymbol{\nu}|\le N_0$: which contains a small vicinity of $\mathbf{A}_0=\mathbf{0}$, by (8.6).

The price that one [ays for replacing (8.11) by]equ(8.12) is that the canonical map generated by Φ_0 is adapted to $H'=h_0+f_0^{[\le N_0]}$ rathe than to $H_0=h_0+f_0$; so it will put the Hamiltonia H' rather than H_0 in the form (8.10). However since H' differs from H_0 by less than $\varepsilon_0^2C_0$ (see (8.13)) it is clear that in the coordinates $(\mathbf{A}',\boldsymbol{\varphi}')$ also H_0 will take the form (8.10). Hence the ultraviolet cut–off, so familiar in Astronomy, does not cause problems *in addition to* those that would already be present if H_0 was replaced by H'.

On the other hand it is clear thta the problems of (8.11) could be just essentially transferred to]equ(8.12) if N_0 is too large. In fact the size of the region in $\mathbf{A}$ where

8.14
$$|\boldsymbol{\omega}_0(\mathbf{A})\cdot\boldsymbol{\nu}|^{-1}\le 2\,C_0\,|\boldsymbol{\nu}|^{-1},\qquad\forall\,0<|\boldsymbol{\nu}|\le N_0 \tag{8.14}$$

can become very small around $\mathbf{0}$ if N_0 is large; the size of this region determines the domain of Φ_0:

8.15
$$\Phi_0(\mathbf{A}',\boldsymbol{\varphi})=\sum_{0<|\boldsymbol{\nu}|\le N_0}\frac{f_{0\boldsymbol{\nu}}(\mathbf{A}')}{-i\,\boldsymbol{\omega}(\mathbf{A}')\cdot\boldsymbol{\nu}} \tag{8.15}$$

and the size of Φ_0 itself; hence it determines the domain of definition of the completely canonical map $\mathcal{C}_0$ generated by Φ_0, which we denote $W(\widetilde{\rho}_0,\widetilde{\xi}_0;\{\mathbf{0}\})$.

To understand what goes on, *i.e.* how small $\widetilde{\rho}_0,\widetilde{\xi}_0$ will turn out to be pone has to find N_0. The basic inequality is the dimensional bound (already discussed in Sect. 3) following from $\rho_0^{-1}|\partial_{\boldsymbol{\varphi}}f_0|\le\varepsilon_0$, see (8.5), and the assumed holomorhy of f_0;

8.16
$$|\boldsymbol{\nu}|\,|f_{o\boldsymbol{\nu}}(\mathbf{A})|\;\le\;\rho_0\varepsilon_0e^{-\xi_0|\boldsymbol{\nu}|},\qquad\forall\,|A_i|<\rho_0,\;\boldsymbol{\nu}\in\mathbb{Z}^\ell \tag{8.16}$$

telling us that for $(\mathbf{A},\mathbf{z})\in W(\rho e^{-\frac12\delta_0},\,\xi_0-\frac12\delta_0;\,\{\mathbf{0}\})$

8.17
$$\begin{aligned}|f_0^{[>N_0]}(\mathbf{A},\mathbf{z})|\;&\le\;\rho_0\varepsilon_0\sum_{|\boldsymbol{\nu}|>N_0}\frac{e^{-\xi_0|\boldsymbol{\nu}|}}{|\boldsymbol{\nu}|}\,e^{(\xi_0-\frac12\delta_0)|\boldsymbol{\nu}|}\;\le\;\rho_0\varepsilon_0e^{\frac14\delta_0N_0}\sum_{|\boldsymbol{\nu}|>0}\frac{e^{-\frac14\delta_0|\boldsymbol{\nu}|}}{|\boldsymbol{\nu}|}\\&\le B_1'\,\rho_0\varepsilon_0\delta_0^{-(\ell+1)}e^{-\frac14\delta_0N_0}\end{aligned} \tag{8.17}$$

so that the *l.h.s.* of (8.13) is estimated by a dimensional bound[5] as

8.18*a*
$$\|f_0^{[>N_0]}\|_{\rho_0 e^{-\delta_0},\xi_0-\delta_0} \stackrel{def}{=} \sup_{W(\rho e^{-\delta_0},\xi_0-\delta_0;\{\mathbf{0}\})} \Big(\Big|\frac{\partial f^{[>N_0]}}{\partial \mathbf{A}}\Big| + \frac{1}{\rho_0 e^{-\delta_0}}\Big|\frac{\partial f^{[>N_0]}}{\partial \boldsymbol{\varphi}}\Big|\Big) \le \\ \le B_1''\,\varepsilon_0\delta_0^{-\ell}e^{-\frac14 N_0\delta_0} \tag{8.18}$$

and a similar calculation leads to

8.18*b*
$$\|f_0^{[>N_0]}\|_{\rho_0 e^{-\delta_0},\xi_0-\delta_0} \le B_1''\varepsilon_0\delta_0^{-\ell} \tag{8.19}$$

Setting the *r.h.s.* of (8.18) equal to $C_0\varepsilon_0^2$ one gets

8.19
$$N_0 = \frac{4}{\delta_0}\log(C_0\varepsilon_0\delta_0^{-\ell}/B_1'')^{-1} \tag{8.20}$$

which we assume to be lerger than 1, without loss of generality. In othe words "the exponential decay of the harmonics of f_0 implies that the cut–off can be chosen to depend on ε_0 *only logarithmically*".

So we see that N_0 is not too large, growing only logarithmically with ε_0 provided δ_0 is not too small. For this reason we take

8.20
$$\delta_0 \stackrel{def}{=} \frac{\xi_0}{\log(C_0\varepsilon_0)^{-1}} \tag{8.21}$$

implying

8.21
$$N_0 \;=\; \frac{4}{\xi_0}(\log(C_0\varepsilon_0)^{-1})\,\log\Big(\frac{C_0\varepsilon_0\xi_0^\ell}{B_1\,(\log(C_0\varepsilon_0)^{-1})^\ell}\Big)^{-1} \tag{8.22}$$

To avoid carrying along such an involved expression we replace it by defining N_0 to equal a slightly larger (but simpler) quantity

8.22
$$N_0 \stackrel{def}{=} \frac{B_1}{\xi_0^2}\big(\log(C_0\varepsilon_0)^{-1}\big)^2 \tag{8.23}$$

where B_1 is a suitable constant: *i.e.* such that (8.23) implies (8.22), hence that thr *r.h.s.* of (8.18) is $\le C_0\varepsilon^2$.

With the choices (8.21),(8.23) the above bounds yield

8.23
$$\begin{aligned} &||f_0^{[>N_0]}||_{\rho_0 e^{-\delta_0},\xi_0-\delta_0} \le C_0\varepsilon_0^2 \\ &||f_0^{[\le N_0]}||_{\rho_0 e^{-\delta_0},\xi_0-\delta_0} \le B_1'\varepsilon_0\delta_0^{-\ell} \le B_2\varepsilon_0\xi_0^{-\ell}\big(\log(C_0\varepsilon_0)^{-1}\big)^\ell \end{aligned} \tag{8.24}$$

Having determined N_0,δ_0 one can easily find the size of the vicinity of $\mathbf{0}$ where Φ_0 can be defined via (8.15); this can be chosen to be $W(\rho_0',\xi_0';\{\mathbf{0}\})$ where $\xi_0' = \xi_0 - \delta_0$ and C_0' is so small that (8.14) holds.

If $\rho_0' < \frac12\rho_0$ and $0 < |\boldsymbol{\nu}| \le N_0$, (here $\boldsymbol{\omega}_0(\mathbf{0}) \equiv \boldsymbol{\omega}_0$) and using a dimansional estimate we obtain

[5] Dimensional bound = bounding a derivative of a holomorphic function at a point by the ratio of the maximum of the function and the distance of the point to the boundary of the holomorphy domain.

8.24
$$\begin{aligned}|\boldsymbol{\omega}_0(\mathbf{A})\cdot\boldsymbol{\nu}|^{-1}&\le|\boldsymbol{\omega}_0(\mathbf{0})\cdot\boldsymbol{\nu}|^{-1}\Big|1-\frac{|\boldsymbol{\omega}_0(\mathbf{A})-\boldsymbol{\omega}_0(\mathbf{0})|\,|\boldsymbol{\nu}|}{|\boldsymbol{\omega}_0(\mathbf{0})\cdot\boldsymbol{\nu}|}\Big|^{-1}\le\\&\le C_0\,|\boldsymbol{\nu}|^{\ell}\Big(1-C_0|\boldsymbol{\nu}|^{\ell+1}\sup_{|\mathbf{a}|\le\frac12\rho_0}\big|\partial_{\,\mathbf{A}}\boldsymbol{\omega}(\mathbf{A})|\,\rho_0'\Big)^{-1}\le\\&\le C_0\,|\boldsymbol{\nu}|^{\ell}\Big(1-C_0\,N_0^{\ell+1}\frac{E_0}{\rho_0-\frac12\rho_0}\,\rho_0'\Big)^{-1}\le\\&\le 2\,C_0\,|\boldsymbol{\nu}|^{\ell},\qquad\forall\,0<|\boldsymbol{\nu}|\le N_0\end{aligned}\tag{8.25}$$

if

8.24′
$$\rho_0'\overset{def}{=}\frac{\rho_0}{4C_0E_0N_0^{\ell+1}}\tag{8.26}$$

and, indeed, $\rho_0'<\frac12\rho_0$ because $N_0\ge1$, $C_0E_0\ge1$.

Therefore the function Φ_0 in (8.15) can be defined in the region $W(\rho_0',\xi_0-\delta_0;\{\mathbf{0}\})$ and there it is bounded by (from (8.14),(8.15),(8.16))

8.25
$$\sup_{W(\rho_0',\xi_0-\delta_0;\{\mathbf{0}\})}\big|\Phi_0(\mathbf{A}',\boldsymbol{\varphi})\big|\le\varepsilon_0\rho_0\sum_{|\boldsymbol{\nu}|\le N_0}2\,C_0\,|\boldsymbol{\nu}|^{\ell}\frac{e^{-\delta_0|\boldsymbol{\nu}|}}{|\boldsymbol{\nu}|}\le B_2\rho_0C_0\varepsilon_0\delta_0^{-2\ell+1}\tag{8.27}$$

Hence by dimensional estimates, for $(\mathbf{A}',\boldsymbol{\varphi})\in W(\rho_0'e^{-\delta_0},\xi_0-2\delta_0;\{\mathbf{0}\})$ we get

8.26
$$\begin{aligned}&\Big|\partial_{\,\mathbf{A}}\Phi_0\Big|+\frac{1}{\rho_0'e^{-\delta_0}}\Big|\partial_{\,\boldsymbol{\varphi}}\Phi_0\Big|\le\\&\le B_2\rho_0C_0\varepsilon_0\delta_0^{-2\ell+1}\Big(\frac{1}{\rho_0'-\rho_0'e^{-\delta_0}}+\big(\frac{4E_0C_0N_0}{\rho_0'e^{-\delta_0}}\big)^{\ell+1}\frac{e^{\xi_0}}{e^{-\xi_0}(1-e^{-\delta_0})}\Big)\le\\&\le B_3(C_0\varepsilon_0)\,\delta_0^{-2\ell}\,(C_0E_0)\,N_0^{\ell+1}\end{aligned}\tag{8.28}$$

The dimensional estimate of $\partial_{\,\boldsymbol{\varphi}}$ is slightly bmore involved because ∂_{φ} has the meaning of $iz\partial_z$ and it is obtained by boundig $|z|$ by $e^{\xi_0-\delta_0}$ and the distance to the boundary by $|e^{-(\xi_0-\delta_0)}-e^{-(\xi_0-2\delta_0)}|$; this explains the factpr $e^{\xi_0}/e^{-\xi_0}(1-e^{-\delta_0})$ which is a bound on $e^{\xi_0-\delta_0}/e^{-(\xi_0-\delta_0)}(1-e^{-\delta_0})$ in (8.28).

In the same way one obtains on $W(\rho_0'e^{-\delta_0},\xi_0-2\delta_0;\{\mathbf{0}\})$ the bound

8.27
$$\Big|\partial^{\,2}_{\mathbf{A}\boldsymbol{\varphi}}\Phi_0\Big|\le B_4\varepsilon_0C_0\,\delta_0^{-2\ell-1}\,C_0E_0\,N_0^{\ell+1}\tag{8.29}$$

In this way the relations

8.28
$$\mathbf{A}=\mathbf{A}'+\partial_{\,\boldsymbol{\varphi}\Phi_0}(\mathbf{A},\boldsymbol{\varphi}),\qquad\boldsymbol{\varphi}'=\boldsymbol{\varphi}+\partial_{\,\mathbf{A}'}\Phi_0(\mathbf{A}',\boldsymbol{\varphi})\tag{8.30}$$

can generate the completely canonical map $\mathcal{C}_0$ via the implicit functions theorem used already in Sect. 3.

The map $\mathcal{C}_0$ will be defined iin a region slightly smaller than the one where Φ_0 is defined, *i.e.* smaller than $W(\rho_0'e^{-\delta_0},\xi_0-2\delta_0;\{\mathbf{0}\})$. To fix the ideas it will be defined in $W(\frac12\rho_0',\xi_0-3\delta_0;\{\mathbf{0}\})$ and it will take values in $W(\rho_0'e^{-\delta_0},\xi_0-2\delta_0;\{\mathbf{0}\})$. As discussed in Sect. 3 the concdition for this is that the Jacobian $\det\partial^{\,2}_{\mathbf{A}\boldsymbol{\varphi}}\Phi_0$ is sufficiently small compared to 1; *i.e.* , recalling (8.29) for some large enough $\widetilde{B}$

8.29 $$\widetilde{B}\, B_4\,]e_0 C_0\, \delta_0^{-2\ell-1}\, E_0 C_0\, N_0^{\ell+1} \tag{8.31}$$

It is then convenient to write

8.30 $$\begin{aligned}\mathbf{A} =&\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}',\boldsymbol{\varphi}')\\ \boldsymbol{\varphi} =&\boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}')\end{aligned}, \qquad (\mathbf{A},\boldsymbol{\varphi}) = \mathcal{C}_0(\mathbf{A},\boldsymbol{\varphi}) \tag{8.32}$$

and to remark that

8.31 $$\boldsymbol{\Xi}(\mathbf{A}',\boldsymbol{\varphi}') = \partial_{\boldsymbol{\varphi}}\Phi_0(\mathbf{A}',\boldsymbol{\varphi}), \qquad \boldsymbol{\Delta}(\mathbf{A}',\boldsymbol{\varphi}') = -\partial_{\mathbf{A}'}\Phi_0(\mathbf{A}',\boldsymbol{\varphi}) \tag{8.33}$$

are bounded by (8.28) which, combined with (8.31) (note that (8.31) contains $\delta_0^{-2\ell-1}$ while (8.28) only contains the factor $\delta_0^{-2\ell}$) implies $\boldsymbol{\Xi}| \le]r_0' e^{-\delta_0}\delta_0$, $|\boldsymbol{\Delta}| \le \delta_0$, which guarantee tha

8.32 $$\mathcal{C}_0, \big(W(\frac{1}{2}\rho_0', \xi_0 - 3\delta_0; \{\mathbf{0}\})\big) \subset W(\rho_0' e^{-\delta_0}, \xi_0 - 2\delta_0; \{\mathbf{0}\}) \tag{8.34}$$

Similar bounds hold for $\mathcal{C}_0^{-1}$ defined by (8.30) through the appropriate inversions.

Heving defined $\mathcal{C}_0$ we can compute the Hamiltonian H_0 in the coordinates $\mathbf{A}', \boldsymbol{\varphi}'$. Using the above notations such a function is denoted as $\widetilde{H}_1$ and

8.33 $$\widetilde{H}_1(\mathbf{A}',\boldsymbol{\varphi}) = H_0((\mathcal{C}_0(\mathbf{A}', Bf')) = \widetilde{h}_1(\mathbf{A}') + \widetilde{f}_1(\mathbf{A}',\boldsymbol{\varphi}') \tag{8.35}$$

where $\widetilde{h}_1$ is by definition (a natural definition, in fact, on the basis of perturbation theory)

8.34 $$\widetilde{h}_1(\mathbf{A}') = h_0(\mathbf{A}') + \overline{f}_0(\mathbf{A}') \tag{8.36}$$

and $\widetilde{f}_1 = \widetilde{H}_1 - \widetilde{h}_1$. Of course $\widetilde{h}_1, \widetilde{f}_1$ are holomorphic in $W(\frac{1}{2}\rho_0', \xi_0 - 3\delta_0; \{\mathbf{0}\})$. Ther sizes can be easily estimated on their domain

8.35 $$\begin{aligned}&\Big|\partial_{\mathbf{A}'}\widetilde{h}_1(\mathbf{A}')\Big| \le E_0 + \varepsilon_0\\ &\Big|\Big(\partial^2_{\mathbf{A}'\mathbf{A}'}\widetilde{h}_1(\mathbf{A}')\Big)^{-1}\Big| = \Big|\Big(\partial^2_{\mathbf{A}'\mathbf{A}'}h_0(\mathbf{A}') + \partial^2_{\mathbf{A}'\mathbf{A}'}\overline{f}_0(\mathbf{A}')\Big)^{-1}\Big| =\\ &= \Big|\partial^2_{\mathbf{A}'\mathbf{A}'}h_0(\mathbf{A}')\Big[1 + \big(\partial^2_{\mathbf{A}'\mathbf{A}'}h_0(\mathbf{A}')\big)^{-1}\partial^2_{\mathbf{A}'\mathbf{A}'}\overline{f}_0(\mathbf{A}')\Big]^{-1}\Big| \le\\ &\le \eta_0\,(1 + 4\eta_0\varepsilon_0\rho_0^{-1})\end{aligned} \tag{8.37}$$

where $f_0(\mathbf{A}')$ has been estimated dimensionally (making use of $|\mathbf{A}'| \le \frac{1}{2}\rho_0' < \frac{1}{2}\rho_0$ and $|\partial_{\mathbf{A}'} f_0(\mathbf{A},\boldsymbol{\varphi})| < \varepsilon_0$ for $|\mathbf{A}| < \rho_0$), and of course (8.37) holds only if

8.36 $$\eta_0\varepsilon_0\rho_0^{-1} < \frac{1}{4} \tag{8.38}$$

The estimate of $\widetilde{f}_1$ is slightly longer; we expect it to be *roughly* of $O(\varepsilon_0^2)$ by construction; therefore it is not surprising that

8.37 $$\sup_{W(\frac{1}{2}\rho_0',\xi_0-4\delta_0;\{\mathbf{0}\})} \Big|\partial_{\mathbf{A}'}\widetilde{f}_1\Big| + \frac{1}{\rho_0'/4}\Big|\partial_{\boldsymbol{\varphi}'}\widetilde{f}_1\Big| \;\le\; B_5\,\varepsilon_0\,(C_0\varepsilon_0)\,(C_0 E_0)^2\, N_0^{\ell+1}\,\delta_0^{-4\ell} \tag{8.39}$$

Let us postpone the proof of (8.39), see Appendix A1, to continue the argument. NOte that (8.39) has to be expected just by our choice of $|F_0, \mathcal{C}_0$. The only quantities that should be computed are B_5 andthe various exponents of $(C_0E_0), \delta_0, N_0$ (but not that of ε_0).

We now look whther tere is an $\mathbf{A}_1$ close to $]V0$ where the "new unperturbed Hamiltonian" $\widetilde{h}_1$ has frequencies $\boldsymbol{\omega}_0$; *i.e.* we write the equation

8.38 $$\boldsymbol{\omega}_0=\boldsymbol{\omega}_0(\mathbf{0})=\widetilde{\boldsymbol{\omega}}_1(\mathbf{A}_1)+\partial_{\mathbf{A}'}\overline{f}_0(\mathbf{A}_1) \tag{8.40}$$

where $\widetilde{\boldsymbol{\omega}}_1\stackrel{def}{=}\partial_{\mathbf{A}'}\widetilde{h}_1$.

This is again an implicit function problem which can be informally discussed as follows. We look for a solution of (8.40) with $|\mathbf{A}_1|<\frac{1}{16}\rho_0'$ (say), *i.e.* with $|\mathbf{A}_1|$ very close to 0, because we want ro consider $\widetilde{H}_1$ as defined on

8.39 $$W(\frac{1}{4}\rho_0',\xi_0-4\delta_0;\{\mathbf{A}_1\}), \tag{8.41}$$

i.e. as defined on a rather large polydisk with center at a point $\mathbf{A}_1$ where $\widetilde{\boldsymbol{\omega}}_1(\mathbf{A}_1)=\boldsymbol{\omega}_0$. Rewrite (8.40), by interpolation, as

8.40 $$\begin{aligned}\mathbf{0}=&\boldsymbol{\omega}_0(\mathbf{A}_1)-\boldsymbol{\omega}_0(\mathbf{0})+\partial_{\mathbf{A}}\overline{f}_0(\mathbf{A}_1)=\\ =&\partial_{\mathbf{A}}\boldsymbol{\omega}_0(\mathbf{0})\cdot\mathbf{A}_1+\Big[\boldsymbol{\omega}_0(\mathbf{A}_1)-\boldsymbol{\omega}_0(\mathbf{0})-\partial_{\mathbf{A}}\boldsymbol{\omega}_0(\mathbf{0})\cdot\mathbf{A}_1+\partial_{\mathbf{A}}\overline{f}_0(\mathbf{A}_1)\Big],\qquad i.e.\\ \mathbf{A}_1=&\Big\{\Big(\partial_{\mathbf{A}}\boldsymbol{\omega}_0(\mathbf{0})\Big)^{-1}\cdot\Big[\int_0^1(1-\lambda)\,d\lambda\,\partial^2_{\mathbf{A}^2}(\lambda\mathbf{A}_1)\cdot\mathbf{A}_1^2+\partial_{\mathbf{A}}\overline{f}_0(\mathbf{A}_1)\Big]\Big\}=\mathbf{0}\end{aligned} \tag{8.42}$$

Denoting by $\mathbf{n}(\mathbf{A})$ the function in curly brackets and regarding it as a function defined for $|\mathbf{A}|<\rho<\frac{1}{2}\rho_0$ one can bound it dimensionally as

8.41 $$|\mathbf{n}|\le\eta_0\cdot\big(\frac{E_0}{(\rho_0/2)^2}\rho^2+\varepsilon_0\big),\qquad \big|\partial_{\mathbf{A}}\mathbf{n}\big|\le\eta_0\cdot\big(\frac{4E_0}{\rho_0^2}\rho^2+\varepsilon_0\big)\frac{1}{\rho_0}. \tag{8.43}$$

In order to have a solution with $\mathbf{A}_1|<\frac{1}{16}\rho_0'$ one imposes that $|\mathbf{n}|<\frac{1}{16}\rho_0'$ and $|\partial_{\mathbf{A}}\mathbf{n}|\ll 1$ for $|\mathbf{A}|<\rho<\frac{\rho_0'}{16}$ ("size and Jacobian conditions"). We choose therefore $\rho=\frac{1}{2}\rho_0\sqrt{\frac{\varepsilon_0}{E_0}}$ so that (8.4) implies

8.42 $$|\mathbf{n}|\le 2\eta_0\varepsilon_0,\qquad |\partial_{\mathbf{A}}\mathbf{n}|\le 2\eta_0\varepsilon_0\rho_0^{-1} \tag{8.44}$$

Hence if for some large B_6' one has

8.43 $$\begin{aligned}&B_6'\eta_0\rho_0^{-1}\varepsilon_0<1,\quad 2\eta_0\varepsilon_0<\frac{1}{2}\rho_0\sqrt{\frac{\varepsilon_0}{E_0}}<\frac{\rho_0}{16}\frac{1}{4C_0E_0N_0^{\ell+1}},\qquad i.e.\text{ if}\\ &B_6'\eta_0\rho_0^{-1}\varepsilon_0<1,\quad 4\eta_0\varepsilon_0\rho_0^{-1}(\sqrt{\frac{\varepsilon_0}{E_0}})^{-1}<1,\quad 32C_0E_0N_0^{\ell+1}\sqrt{\frac{\varepsilon_0}{E_0}},\text{or, more simply, if}\\ &B_6\,(\eta_0\rho_0^{-1}E_0)^2(C_0E_0)N_0^{2(\ell+1)}\varepsilon_0C_0<1\end{aligned} \tag{8.45}$$

with a suitable B_6 then there will be an $\mathbf{A}_1$ satisfying $\widetilde{\boldsymbol{\omega}}_1(\mathbf{A}_1)=\boldsymbol{\omega}_0$ and

8.44 $$|\mathbf{A}_1| \le \frac{1}{2}\rho_0\sqrt{\frac{\varepsilon_0}{E_0}} < \frac{1}{16}\rho'_0 \tag{8.46}$$

For a more detailed discussion of the implicit function theorem used here see Ref. [1] at p. 490 (for instance or "do it yourself" which is easier).

At this point the renormalization transformation is completed as follows: define new coordinates $(\mathbf{A}'', \boldsymbol{\varphi}'')$ by $(\mathbf{A}, \boldsymbol{\varphi}) = \mathcal{C}_0(\mathbf{A}', \boldsymbol{\varphi}')$ and $(\mathbf{A}', \boldsymbol{\varphi}') = \mathcal{L}_0(\mathbf{A}'', \boldsymbol{\varphi}'')$ where

8.45 $$\mathbf{A}'' = \frac{4\rho_0}{\rho'_0}(\mathbf{A}' - \mathbf{A}_1), \qquad \boldsymbol{\varphi}'' = \boldsymbol{\varphi}' \tag{8.47}$$

The map $\overline{\mathcal{C}}_0 = \mathcal{C}_0\mathcal{L}_0$ maps $W(\rho_0, \xi_0 - 4\delta_0; \{\mathbf{0}\})$ into $W(\rho_0, \xi_0; \{\mathbf{0}\})$ and is the composition of a map which is a simple translation and rescaling (*i.e.* $\mathcal{L}_0$) with a map which is "very close" to the identity together with its derivatives (*i.e.* $\mathcal{C}_0$).

In fact a derivative of order p with respect to $\boldsymbol{\varphi}'$ and of order q with respect to $\mathbf{A}'$ of $\mathcal{C}_0 -$identity can be estimated dimensionally: recall that $\mathcal{C}_0$ is defined and holomorphic in the large domain $W(\rho'_0 e^{-\delta_0}, \xi_0 - 3\delta_0; \{\mathbf{0}\})$ and bounded there by (8.33),(8.28); however $\mathcal{C}_0$ is only considered inside the smaller domain $W(\frac{1}{4}\rho'_0, \xi_0 - 4\delta_0; \{\mathbf{0}\})$ so that its derivatives can be dimensionally bounded. And one finds that the p-th $\boldsymbol{\varphi}'$-derivative and the q-th $\mathbf{A}'$-derivative of $\mathcal{C}_0 -$ identity are bounded by

8.46 $$\begin{aligned} &(const)^{p+q}\, \rho_0\, C_0\varepsilon_0\, \delta_0^{-2\ell}\, \delta_0^{-p} {\rho'_0}^{-q}, \text{for } \boldsymbol{\Xi} \\ &(const)^{p+q}\, C_0\varepsilon_0\, \delta_0^{-2\ell}\, C_0 E_0\, \delta_0^{-p} {\rho'_0}^{-q}, \text{for } \boldsymbol{\Delta} \end{aligned} \tag{8.48}$$

in $W(\frac{1}{4}\rho'_0, \xi_0 - 4\delta_0; \{\mathbf{A}_1\})$.

Define $\lambda_0 \stackrel{def}{=} 4\frac{\rho_0}{\rho'_0}$ (which is, by our choices, $\gg 1$) and

8.47 $$H_1(\mathbf{A}'', \boldsymbol{\varphi}'') = \lambda_0 \Big(H_0(\mathcal{C}_0\mathcal{L}_0(\mathbf{A}'', \boldsymbol{\varphi}'')) - \widetilde{h}_1(\mathbf{A}_1)\Big) \stackrel{def}{=} h_1(\mathbf{A}'') + f_1(\mathbf{A}'', \boldsymbol{\varphi}'') \tag{8.49}$$

where

8.48 $$h_1(\mathbf{A}'') \stackrel{def}{=} \lambda_0 \Big(\widetilde{h}_1(\mathbf{A}_1 + \lambda_0^{-1}\mathbf{A}'') - \widetilde{h}_1(\mathbf{A}_1)\Big), \qquad f_1(\mathbf{A}'', \boldsymbol{\varphi}'') = \lambda_0\, \widetilde{f}_1(\mathbf{A}_1 + \lambda_0^{-1}\mathbf{A}'', \boldsymbol{\varphi}''), \tag{8.50}$$

and remark that the $\overline{\mathcal{C}}_0$-image of a motion in $W(\rho_0, \xi_0 - 4\delta_0; \{\mathbf{0}\})$ described by the Hamltonian $\widetilde{H}_1$ is a motion in $W(\rho_0, \xi_0; \{\mathbf{0}\})$ described by the Hamiltonian H_0.

In fact if $\mathbf{A}' = \mathbf{A}_1 + \lambda_0^{-1}\mathbf{A}'', \boldsymbol{\varphi}' = \boldsymbol{\varphi}''$ the motion with Hamiltonian $\widetilde{H}_1(\mathbf{A}', \boldsymbol{\varphi}')$ are describednin the variables $\mathbf{A}'', \boldsymbol{\varphi}''$ by $\lambda_0\widetilde{H}_1(\mathbf{A}_1 + \lambda_0^{-1}\mathbf{A}'', \boldsymbol{\varphi}'')$ as it can be immediately realized by writing hte equations of motion. The addition of the *constant* term $-\lambda_0\widetilde{h}_1(\mathbf{A}_1)$ in (8.49),(8.50) does not affect the equations of motion but it is convenient as it will become clear shortly.

We can define explicitly the renormaliztion transformation $\mathcal{K}$ as a map acting on the space of the pairs (h_0, f_0) of holomorphic functions on $W(\rho_0, \xi_0; \{\mathbf{0}\})$ with h_0 depending only on $\mathbf{A}$ as:

8.49 $$\mathcal{K}(h_0, f_0) = (h_1, f_1) \tag{8.51}$$

with values h_1, f_1 defined on $W(\rho_0, \xi_0 - 4\delta_0; \{\mathbf{0}\})$ by (8.49), (8.50) and (see also (8.21),(8.23))

8.50 $$\delta_0\frac{\xi_0}{\log(C_0\varepsilon_0)^{-1}},\quad \lambda_0=\frac{4\rho_0}{\rho_0'}=16\,(E_0C_0)\,N_0^{\ell+1},\quad N_0=\frac{B_1}{\xi_0^2}\Big(\log(C_0\varepsilon_0)^{-1}\Big)^2 \tag{8.52}$$

The domain of definition of $\mathcal{K}$ is restricted by the conditins that have been imposed while constructing it, namely $\xi_0>4\delta_0$ and (8.31) (to define $\mathcal{C}_0$), (8.38) and the stronger (8.45) (to define $\mathbf{A}_1$); *all such conditions* can be implied by a condition of the form

8.51 $$B_7'\,(\varepsilon_0C_0)\,(E_0C_0)^{\alpha_1'}\,(E_0\eta_0\rho_0^{-1})^{\alpha_2'}\,\xi_0^{-\alpha_3'}\,\big(\log(C_0\varepsilon_0)^{-1}\big)^{\alpha_4'} \tag{8.53}$$

with, for instance, $]a_1'=1,\alpha_2'=2.\alpha_3'=\alpha_4'=4\ell+1$ (just substitute in (8.31),(8.45) the expressions for δ_0,N_0 in terms of ξ_0,ε_0).

The new pair (h_1,f_1) satisfies the following bounds on $W(\rho_0,\xi_0-4\delta_0;\{\mathbf{0}\})$

8.52 $$\big|\partial_{\mathbf{A}}\big|\le E_0+\varepsilon_0,\qquad \Big|\Big(\partial^2_{\mathbf{A}^2}h_1\Big)^{-1}\Big|\le\lambda_0\,\eta_0\,(1+4\eta_0\varepsilon_0\rho_0^{-1}, \tag{8.54}$$

obtained by transcription and rescaling from (8.37), and

8.53 $$\big|\partial_{\mathbf{A}}\big|+\frac{1}{\rho_0}\big|\partial_{\boldsymbol{\varphi}}f_1\big|\le B_5\,\varepsilon_0\,(C_0\varepsilon_0)\,(C_0E_0)^2\,N_0^{\ell+1}\,\delta_0^{-4\ell} \tag{8.55}$$

obtained from (8.39)

It is convenient to simplify the above relations at the expense of assuming a somewhat more involved condition, like (8.53). In fact if we assume

8.54 $$\begin{aligned}&16\,(E_0C_0)\,N_0^{\ell+1}\,(1+4\eta_0\varepsilon_0\rho_0^{-1})\;\le\;\Big(\log(C_0\varepsilon_0)^{-1}\Big)^{2(\ell+2)}\overset{def}{=}\overline{\lambda}_0\\ &B_5\sqrt{C_0\varepsilon_0}\,(C_0E_0)^2\,N_0^{\ell+1}\,\delta_0^{-4\ell}\;<\;1.\end{aligned} \tag{8.56}$$

And one can define the new parameters $E_1,\eta_1,\varepsilon_1,\rho_1,\xi_1$ for H_1, *i.e.* for (h_1,f_1), see (8.5), as

8.55 $$E_1=E_0+\varepsilon_0,\qquad \eta_1=\overline{\lambda}_0\eta_0,\qquad C_1\varepsilon_1=(C_0\varepsilon_0)^{3/2},\qquad \xi_1=\xi_0\,(1-\frac{4}{\log(C_0\varepsilon_0)^{-1}}),\qquad \rho_1=\rho_0 \tag{8.57}$$

and (8.56) can be imposed by requiring a single (stronger) condition like

8.56 $$B_7\,(\varepsilon_0C_0)\,(E_0C_0)^2\,(E_0\eta_0\rho_0^{-1})^2\Big(\xi_0^{-1}\log(C_0\varepsilon_0)^{-1}\Big)^{10\ell+3}\;<\;1 \tag{8.58}$$

with B_7 suitably large (depending only on ℓ.

p.8.2 **8.2.** *Iteration and fixed point for the renormalization.*

We are now in a position to iterate: starting from H_1 one can apply $\mathcal{K}$ to it and define H_2 *etc*In general H_n can be defined in terms of H_{n-1} provided condition (8.58) written with the parameters $\varepsilon_{n-1},E_{n-1},\eta_{n-1},\xi_{n-1}$ instead of $\varepsilon_0,E_0,\eta_0,\xi_0$ holds.

Assuming that such a condition holds we would deduce from (8.57)

8.57
$$\begin{aligned} C_0\varepsilon_n =&(C_0\varepsilon_0)^{(3/2)^n}\\ \xi_n \ge&\xi_0\prod_{k=0}^\infty\Big(1-\frac{4}{(3/2)^k\log(C_0\varepsilon_0)^{-1}}\Big)\ge\frac12\xi_0\\ E_n\le& E_0+C_0^{-1}\sum_{k=0}^\infty(C_0\varepsilon_0)^{(3/2)^k}\le 2\,E_0\\ \eta_n =&\eta_0\,\Big(\log(C_0\varepsilon_0)^{-1}\Big)^{2(\ell+2)n}\Big(\frac32\Big)^{\frac12 n\,(n-1)\,2(\ell+2)}, \end{aligned}\tag{8.59}$$

where the two intermediate inequalities hld if $C_0\varepsilon_0$ is small enough: and we may and shall impose that the latter condition is already implied by a suitably large choice of B_7 in (8.58). So (8.7) will hold for all n if

8.58
$$B_7\,(\varepsilon_nC_0)\,(E_nC_0)^2\,(E_n\eta_n\rho_0^{-1})^2\Big(\xi_n^{-1}\log(C_0\varepsilon_n)^{-1}\Big)^{10\ell+3}\,<\,1\tag{8.60}$$

with B_7 suitably large (depending only on ℓ.

Substituting the bounds (8.59) in (8.60) shows that (8.60) holds *eventually* in n because the renormalized perturbation size $\varepsilon_n \xrightarrow[n\to\infty]{} 0$ *very fast* compared to the speed with which $\eta_n \xrightarrow[n\to\infty]{}\infty$. Therefore if one imposes (8.60) for sufficiently many values of n one deduces from (8.59) that (8.60) holds *for all* n. This means that a condition like

8.59
$$B\,(\varepsilon_0C_0)\,(E_0C_0)^{\alpha_1}\,(E_0\eta_0\rho_0^{-1})^{\alpha_2}\,\xi_0^{-\alpha_3}\,<\,1,\tag{8.61}$$

with $B,\alpha_1,\alpha_2,\alpha_3$ conveniently chosen and depending ony on ℓ sufficies to imply (8.59) and (8.60) for all n.

We conclude that under the condition (8.61) all iterates of $\mathcal{K}$ can be applied to (h_0,f_0). If $N_n=B_1\xi_n^{-2}\big(\log(C_0\varepsilon_n)^{-1}\big)^2$ and $\lambda_n=16E_nC_0N_n^{\ell+1}$ (see (8.52)) then the (8.50) yield for some $c>0$ (only dependent on ℓ)

8.60
$$\begin{aligned} h_n(\mathbf{A}'')=&\lambda_{n-1}\,\big(\widetilde h_{n-1}(\mathbf{A}_n+\lambda_{n-1}^{-1}\mathbf{A}'')-\widetilde h_n(\mathbf{A}_n)\big)=\\ =&\boldsymbol{\omega}_0\cdot\mathbf{A}''+\lambda_{n-1}\frac{E_{n-1}+|e_{n-1}}{\rho_{n-1}/2}\,(\lambda_{n-1}\rho_0)^2\xrightarrow[n\to\infty]{}\boldsymbol{\omega}\cdot\mathbf{A}''\\ &Dpr_{\mathbf{A}''}h_n(\mathbf{A}'')\xrightarrow[n\to\infty]{}\boldsymbol{\omega}_0,\qquad \partial^p_{\mathbf{A}^p}h_n(\mathbf{A}'')\xrightarrow[n\to\infty]{}0\quad\forall p\ge2\\ \partial^{p+q}_{\mathbf{A}''^p\,\boldsymbol{\varphi}''^q}f_n(\mathbf{A}'',\boldsymbol{\varphi}'')=&O(\varepsilon_nC_0\delta_n^{-c-q}\rho_0^{c-p})\xrightarrow[n\to\infty]{}0 \end{aligned}\tag{8.62}$$

by dimensional estimates.

The bounds (8.48),(8.49) and (8.59) imply that

8.61
$$\widetilde{\mathcal{C}}=\lim_{j\to\infty}\widetilde{\mathcal{C}}_0\,\widetilde{\mathcal{C}}_1\,\cdots\widetilde{\mathcal{C}}_j\tag{8.63}$$

exists and is a holomorphic map from $W(\rho_0,\frac12\xi_0;\{\mathbf{0}\})$ into $W(\rho_0,\xi_0;\{\mathbf{0}\})$ with the property that

8.62
$$\big|\partial_{\,\mathbf{A}''}\widetilde{\mathcal{C}}(\mathbf{A}'',\boldsymbol{\varphi}'')\big|\le const\,\prod_{j=0}^\infty\lambda_0^{-1}=0,\tag{8.64}$$

because $\mathcal{L}_k$ is a contraction by $\sim \lambda_k^{-1}$. Hence $\widetilde{\mathcal{C}}(\mathbf{A}'', \boldsymbol{\varphi}'')$ is $\mathbf{A}''$-independent and it can be written

8.63
$$\mathbf{A} = \boldsymbol{\alpha}(\boldsymbol{\varphi}''), \qquad \boldsymbol{\varphi} = \boldsymbol{\varphi}'' + \boldsymbol{\beta}(\boldsymbol{\varphi}''), \tag{8.65}$$

defining, therefore, an analytic torus $\mathcal{T}_{\boldsymbol{\omega}_0}$ in $W(\rho_0, \xi_0; \{\mathbf{0}\})$.

The uniformity of the convergence of (8.63) in $W(\frac{1}{2}\rho_0, \frac{1}{4}\xi_0; \{\mathbf{0}\})$ (implied by the convergence and analyticity in the larger domain $W(\rho_0, \frac{1}{2}\xi_0; \{\mathbf{0}\})$) allows usto infer that, denoting S_t^H the time evolution generated by the generic Hamiltonian H,

8.64
$$\begin{aligned} S_t^{H_0}(\mathbf{A}, \boldsymbol{\varphi}) &= \lim_{j\to\infty} S_t^{H_0}\Big(\widetilde{\mathcal{C}}_0\,\widetilde{\mathcal{C}}_1\,\cdots\widetilde{\mathcal{C}}_j(\mathbf{0}, \boldsymbol{\varphi}'')\Big) = \\ &= \lim_{j\to\infty} \widetilde{\mathcal{C}}_0\,\widetilde{\mathcal{C}}_1\,\cdots\widetilde{\mathcal{C}}_j(S_t^{H_{j+1}}(\mathbf{0}, \boldsymbol{\varphi}'')) = \\ &= \lim_{j\to\infty} \widetilde{\mathcal{C}}_0\,\widetilde{\mathcal{C}}_1\,\cdots\widetilde{\mathcal{C}}_j(\mathbf{0} + O(\varepsilon_j C_0)\delta_j^{-1}\,t, \boldsymbol{\varphi}'' + \boldsymbol{\omega}_0 t + O(\varepsilon_j C_0)\delta_j^{-1} t) = \\ &= \lim_{j\to\infty} \widetilde{\mathcal{C}}_0\,\widetilde{\mathcal{C}}_1\,\cdots\widetilde{\mathcal{C}}_j(\mathbf{0}, \boldsymbol{\varphi}'' + \boldsymbol{\omega}_0\,t) \equiv \widetilde{\mathcal{C}}(\mathbf{0},'' + \mathbf{Bo_0\,t}). \end{aligned} \tag{8.66}$$

which proves that $\mathcal{T}_{\boldsymbol{\omega}_0}$ is invariant and the motion on it is quasi periodic with frequencies $\boldsymbol{\omega}_0$; actually, in the parametrization (8.65) it is simply $\boldsymbol{\varphi}'' \to \boldsymbol{\varphi}'' + \boldsymbol{\omega}_0 t$.

This completes the proof of the theorem, provided one accepts the bound (8.39): the latter bound is checked in Appendix A1.

Remark: The above analysis can be regarded as a proof that the iteration of he renormalization map $\mathcal{K}$ "*drives*" any Hamiltonian (h_0, f_0) close to integrable (in the sense that it satisfies (8.9)) to the *harmonic oscillator* $(\boldsymbol{\omega}_0 \cdot \mathbf{A}, 0)$ near a Diophantine rotation vector $\boldsymbol{\omega}_0$: more precisely it *magnifies* phase space around an invariant torus with given rotation vector $\boldsymbol{\omega}_0$ and shows that in the magnified view the Hamiltonian can be identified with that of an harmonic oscillator.

*sec.*9

9. Concluding remarks

*p.*9.1 **9.1.** *An extension of the results in Sect. 8*

The following (see [8]) extension shows that the invariant tori although filling a set with an open dense complement (at least if one only considers the ones that we can prove to exist by interpreting the result in Sect. 8) nevertheless they can be *smoothly interpolated*.

There exit positive constants $B, B', \alpha_1, \alpha_2, \ldots, \alpha_6$ depending only on ℓ which control the following statements. Let h_0, f_0 be holomporphic in $W(\rho_0, \xi_0; V)$ where V is a sphere. Then there exists a completely canonical map $\mathcal{C}$ of class C^∞ defined in $\mathbb{R}^\ell \times \mathbb{T}^\ell$ and with values in $\mathbb{R}^\ell \times \mathbb{T}^\ell$ and a C^∞ function $\overline{h}$ on $\mathbb{R}^\ell$ such that

(1) For $(\mathbf{A}', \boldsymbol{\varphi}') \in V_{f_0} \times \mathbb{T}^\ell$, with $V_{f_0} \subset V$ and V_{f_0} suitably chosen, it is

9.1
$$H_0(\mathcal{C}(\mathbf{A}', \boldsymbol{\varphi}')) \overset{C^\infty}{=} \overline{h}(\mathbf{A}') \tag{9.1}$$

where $\overset{C^\infty}{=}$ means that the r.h.s. and the l.h.s. are identical together with all their derivatives on the set $V_{f_0} \times \mathbb{T}^\ell$. Note that if V_{f_0} does not contain open subsets this is a non trivial property which in general does not follow from the simple equality $H_0(\mathcal{C}(\mathbf{A}', \boldsymbol{\varphi}')) = \overline{h}(\mathbf{A}')$ on $V_{f_0} \times \mathbb{T}^\ell$.

(2) $V_{f_0} \supseteq \big\{\, \mathbf{A}' \,\big|\, \big|\partial_{\mathbf{A}'}\overline{h}(\mathbf{A}') \cdot \boldsymbol{\nu}\big|^{-1} \le C_0\, |\boldsymbol{\nu}|^\ell \big\}$ for $C_0 > 0$ given in (9.2) below.

(3) $\mathrm{vol}(V_{f_0} \times \mathbb{T}^\ell) \geq (1-\lambda)\,\mathrm{vol}(V \times \mathbb{T}^\ell)$
provided ε_0 *is small enough. The constraints on* C_0, ε_0 *are*

9.2
$$\begin{aligned}&\frac{\varepsilon_0}{E_0} \leq B\,(E_0\eta_0\rho_0^{-1})^{-\alpha_1}\,\xi_0^{\alpha_2}\,\lambda^{\alpha_3},\\ &C_0 = B'\,E_0^{-1}\,(\frac{\varepsilon_0}{E_0})^{-\alpha_4}\,\xi_0^{\alpha_5}\,(E_0\eta_0\rho_0^{-1})^{-\alpha_6}\end{aligned} \tag{9.2}$$

The above theorem says that quasi periodic motions fill most of phase space if the perturbation is very small; furthermore *the invariant tori fill the phase space in a smooth way.* In fact they can be embeddein a smooth foliation (of class C^∞) of tori with parametric equations $(\mathbf{A},\boldsymbol{\varphi}) = \mathcal{C}(\mathbf{A}',\boldsymbol{\varphi}')$, $\boldsymbol{\varphi}' \in \mathbb{T}^\ell$. Such tori are, however, invariant only if $\mathbf{A}' \in V_{f_0}$.

Furthermore item (2) shows that V_{f_0} as given by the proof in Sect. 8 must be thought of as the complement of an open dense set whose complement has small measure (by item (3)), [19].

One can prove that the above theorem implies "absence of Arnold diffusion" in the anisochronous systems considered here if $\ell = 2$.

p.9.2 **9.2.** *Extension to perturbations of isochronous systems*

A naturally modifed form of the above theorem holds also for systems of the form

9.3
$$\boldsymbol{\omega}_0 \cdot \mathbf{A} + \varepsilon f_0(\mathbf{A},\boldsymbol{\varphi}) \tag{9.3}$$

i.e. for perturbations of harmonic oscillators, provided $|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}|^{-1}| < C|\boldsymbol{\nu}|^\ell, \forall\boldsymbol{\nu} \neq \mathbf{0}$ for some $C > 0$ and provided ε is small compared to $\lambda, \xi_0, E_0\eta_0'\rho_o^{-1}$, where $\eta_0' - \sup|\partial^2_{\mathbf{A}^2}\overline{f}_0$, see Ref. [19],[20].

p.9.3 **9.3.** *Integrability conditions for perturbations of harmonic oscillators*

Another theorem that can be proved is

Consider the perturbation expansion for (9.3) with $\boldsymbol{\omega}_0$ *satisfying the non resonance condition* $|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}|^{-1} \leq C|\boldsymbol{\nu}|^\ell$, $\forall\boldsymbol{\nu} \neq \mathbf{0}$. *Suppose that to all orders of perturbation theory one finds that* $h_n(\mathbf{A}')$ *depends on* $\mathbf{A}'$ *only through* $\boldsymbol{\omega} \cdot \mathbf{A}'$*, i.e.* $h_n(\mathbf{A}') = s_n(\boldsymbol{\omega}_0 \cdot \mathbf{A}')$ *for some* s_n*, i.e. to the* n*-th order Hamiltonian of the Birkhoff series depends on* $\mathbf{A}'$ *only through* $\boldsymbol{\omega} \cdot \mathbf{A}'$*. Then the system is canonically integrable and the Birkhoff series for the Hamiltonian and for the integrating canonical map converge.*

See [21] and, for a proof similar to the ones developed here, see [22].

p.9.4 **9.4.** *Breakdown of invariant tori.*

Finally consider a one parameter family of perturbations

9.4
$$h_0(\mathbf{A}) + \varepsilon f(\mathbf{A},\boldsymbol{\varphi}) \tag{9.4}$$

with $h_0, \varepsilon f = f_0$ satisfying the assumptions of the KAM theorem of Sect. 8. Fix $\boldsymbol{\omega}_0 = \boldsymbol{\omega}_0(\mathbf{0})$ as in Sect. 8. THEN

9.5
$$\mathcal{K}^n(h_0,\varepsilon f) \xrightarrow[n\to\infty]{} (\boldsymbol{\omega}_0 \cdot \mathbf{A}, 0) \tag{9.5}$$

for ε small enough.

It is tempting to hope that there is a *non trivial fixed point* (h^*, f^*) for $\mathcal{K}$ such that for $\varepsilon = \varepsilon_c$ it is $\mathcal{K}^n(h_0,\varepsilon f) \xrightarrow[n\to\infty]{} (h^*, f^*)$ while for $0 < \varepsilon < \varepsilon_c$ it is $\mathcal{K}^n(h_0,\varepsilon f) \xrightarrow[n\to\infty]{} (\boldsymbol{\omega}_0 \cdot \mathbf{A}, 0)$.

A discussion of similar problems can be found in Ref. [18]. A somewhat different approach has been developed in detail to study the transition to chaos near a torus of fixed frequencies $\boldsymbol{\omega}_0$ in the cases $\ell = 2$, see Ref. [23],[24].

Acknowledgement: I am grateful to Gian Carlo Benettin and Luigi Chierchia for critically reading some parts of these lectures.

app.A1

Appendix A1. Check of the bound (8.39)

This is a dimensional bound. Let $(\mathbf{A}', \boldsymbol{\varphi}') \in W(\rho_0 e^{-\delta_0}, \xi_0 - 3\delta_0; \{\mathbf{0}\})$ and

A1.1
$$\begin{aligned}\widetilde{f}_1(\mathbf{A}', \boldsymbol{\varphi}') = \widetilde{H}_1(\mathbf{A}', \boldsymbol{\varphi}') - \widetilde{h}_1(\mathbf{A}') = \\ h_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}', \boldsymbol{\varphi}')) + f_0(\mathbf{A}' + \boldsymbol{\Xi}(\mathbf{A}', \boldsymbol{\varphi}'), \boldsymbol{\varphi}' + \boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}')) - h_0(\mathbf{A}') - \overline{f}_0(\mathbf{A}').\end{aligned} \tag{A1.1}$$

The main remark is that the functions $\boldsymbol{\Xi}, \boldsymbol{\Delta}$ satisfy (8.33); hence (8.28) provides bounds for their size. The (8.33) means that $\boldsymbol{\Xi}, \boldsymbol{\Delta}$ satisfy

A1.2
$$\boldsymbol{\omega}_0(\mathbf{A}') \cdot \boldsymbol{\Xi}(\mathbf{A}', \boldsymbol{\varphi}) + f_0^{[\le N_0]}(\mathbf{A}', \boldsymbol{\varphi} + \boldsymbol{\Delta}(\mathbf{A}', \boldsymbol{\varphi}')) - \overline{f}_0(]VA') = 0 \tag{A1.2}$$

Hence (A1.1) can be rewritten, subtracting (A1.2) from it, on $W(\rho_0' e^{-\delta_0}, \xi_0 - 3\delta_0; \{\mathbf{0}\})$ as

A1.3
$$\begin{aligned}\widetilde{f}_1(\mathbf{A}', \boldsymbol{\varphi}') = &\big[\, h_0(\mathbf{A}' + \boldsymbol{\Xi}) - h_0(\mathbf{A}') - \boldsymbol{\omega}_0(\mathbf{A}') \cdot \boldsymbol{\Xi} \,\big] + \\ &+ \big[\, f_0^{[\le N_0]}(\mathbf{A}' + \boldsymbol{\Xi}, \boldsymbol{\varphi}' + \boldsymbol{\Delta}) - f_0^{[\le N_0]}(\mathbf{A}', \boldsymbol{\varphi}' + \boldsymbol{\Delta}) \,\big] \, - \big[\, f^{[>N_0]}(\mathbf{A}' + \boldsymbol{\Xi}, Bf' + \boldsymbol{\Delta}) \,\big] \, = \\ &\stackrel{def}{=} \big[\, f^I \,\big] + \big[\, f^{II} \,\big] + \big[\, f^{III} \,\big]\end{aligned} \tag{A1.3}$$

And, denoting $\big|\big|\boldsymbol{\Xi}\big|\big|$ the maximum of $\big|\boldsymbol{\Xi}\big|$ in the domain $W(\rho_0 e^{-\delta_0}, \xi_0 - 3\delta_0; \{\mathbf{0}\})$ we can bound F^I, f^{II}, f^{III} by

A1.4
$$\begin{aligned}\big|f^I\big| = &\Big| \int_0^1 (1-\lambda)\, d\lambda\, \partial^{\,2}_{\mathbf{A}'^2} h_0(\mathbf{A}' + \lambda\boldsymbol{\Xi})\, \boldsymbol{\Xi}\, BX \Big| \le \frac{E_0}{\rho_0 - \rho_0/2}\, \big|\big|\boldsymbol{\Xi}\big|\big|^2, \\ \big|f^I I\big| = &\Big| \int_0^1 d\lambda\, \partial_{\,\mathbf{A}'} f_0^{[\le N_0]}(\mathbf{A}' + \boldsymbol{\Xi}, \boldsymbol{\varphi}' + \boldsymbol{\Delta})\, \boldsymbol{\Xi} \Big| \le B_1' \varepsilon_0 \delta_0^{-1}\, \big|\big|\boldsymbol{\Xi}\big|\big|,\end{aligned} \tag{A1.4}$$

$$\big|f^I II\big| \le \sum_{|\boldsymbol{\nu}| > N_0} \varepsilon_0 \rho_0 \frac{e^{-\delta_0 |\boldsymbol{\nu}|}}{|\boldsymbol{\nu}|} \le \rho_0 C_0 \varepsilon_0^2$$

having used (8.24), the definition of N_0 and some easy dimensional estimates. Substituting the bounds on $|\boldsymbol{\Xi}| \equiv \partial_{\,\boldsymbol{\varphi}} \Phi_0|$ following (8.28), a bound on $\widetilde{f}_1$ in $W(\rho_0' e^{-\delta_0}, \xi_0 - 3\delta_0; \{\mathbf{0}\})$ of the form

A1.5
$$B_3'\, \varepsilon_0\, (C_0 \varepsilon_0)\, (C_0 E_0)\, \delta_0^{-4\ell}\, \rho_0 \tag{A1.5}$$

is obtained and it immediately yields (8.39) by a dimensional estimate.

References

Ga [1] G. Gallavotti, *The elements of mechanics*, Springer-Verlag, New York, 1983.

CEG [2] P. Collet, H. Epstein, G. Gallavotti, *Perturbations of geodesic flows on surfaces of constant negative curvature and their mixing properties*, Communications in Mathematical Physics, **95**, 61–112, 1984.

Ar [3] V. Arnold, *Les Méthodes mathématiques de la Mécanique classique*, MIR, Moscow, 1978.

Jo [4] R. Jost, *Winkel und Wirkungsvariablen für allgemeine mechanische systeme*, Helvetica Physica Acta, **41**, 965–968, 1968.

Fa [5] F. Fassó, *Quasi-periodicity of motions and complete integrability of Hamiltonian systems*, Ergodic Theory and Dynamical Systems **18**, 1349-1362, 1998. And FF, A. Giacobbe, *Geometric structure of "broadly integrable" Hamiltonian systems*, Journal of Geometry and Physics **44**, 156-170, 2002.

Po [6] H. Poincaré, *Méthodes nouvelles de la mécanique celèste*, Vol. I, Gauthier–Villars, reprinted by Gabay, Paris, 1987.

Ne [7] V. Nekhorossev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Russian Mathematical Surveys, **32 (6)**, 1–65, 1977.

BGG [8] G. Benettin, L. Galgani, A. Giorgilli, *A proof of Nekhorossev theorem for the stability times in nearly integrable Hamiltonian systems*, Celestial Mechanics, **??**, ???-???, 198?.

BG [9] G. Benettin, G. Gallavotti, *Stability of motions near resonances in quasi-integrable hamiltonian systems*, Journal of Statistical Physics, **44**, 293-338, 1986.

Ar2 [10] V. Arnold, *Instabilities of dynamical systems with several degrees of freedom*, Soviet MAthematicesky Doklady, **5** (3), 581–585, 1964.

BGGS [11] G. Benettin, L. Galgani, A. Giorgilli, J.M. Strelcyn, *A proof of Kolmogorov theorem on invariant tori using canonical transformations defined by the Lie method*, Il Nuovo Cimento, **79-B**, 201–223, 1984.

Ru [12] H. Rüssmann, *Note on sums containing smalldivisors*, Communications in Pure and Applied Mathematics, **29**, 755–758, 1976.

Ru1 [13] H. Rüssmann, *One dimensional Schödinger equation*, Annals of the New York Academy of Sciences, **357**, 90–107, 1980.

Br [14] A. Brjuno, *Analytic form of differential equations*, Transactions of the Moscow Mathematical Society, **25** 131–288, 1971. And **26**, 199–339, 1972.

He [15] M. Herman, *Sur les courbes invariantes par les diffeomorphismes de l'anneau*, Asterisque, **103-104**, 1–221, 1983.

Ll[16] R. de la Llave, *A simple proof of a particular case of Siegel's center theorem*, Journal of Mathematical Physics, **24**, 2118–2121, 1983.

Ga1[17] G. Gallavotti, *The integrability problem and the Hamilton–Jacobi equation*, Physics Reports, **103**, 177–184, 1984.

Ga2[18] G. Gallavotti, *Classical Mechanics and Renormalization Group*, in [*Regular and chaotic motions in Dynamic Systems], ed. G.Velo, A.Wightman, Proceedings of the fifth international School of Mathematical Physics at Erice, 1983, Plenum Press, 1985, p. 185–232.*

CG[19] L. Chierchia, G. Gallavotti, *Smooth prime integrals in quasi integrable systems*, Il Nuovo Cimento, **67B**, 277–297, 1982.

Ga3[20] G. Gallavotti, *Perturbation theory forclassical Hamiltonian systems*, in *Scaling and self similarity in Physics*, ed. J. Fröhlich, Birkhauser, Boston, 1985, p.359–426.

Ru3[21] H. Rüssmann, *Über die normalform analytischer Hamiltonsher Differentialgleichungen in der Nähe einer Gleichgewichtlösung*, Mathematische Annalen, **169**, 55–72, 1967.

Ga4[22] G. Gallavotti, *A criterion of integrability for perturbed harmonic oscillators. Wick ordering in Classical Mechanics and invariance of the frequency spectrum*, Communications in Mathematical Physics, **87**, 365–383, 1982. The main part of this paper consists of a derivatin of the results of Ref. [21] in a slightly different (though equivalent) setting.

FKS[23] M. Feigenbaum, L. Kadanoff, S. Shenker, *Quasi periodicity in dissipative systems. A renormalization group analysis*, Physica D, **5**, 370–386, 1982.

ORSS[24] S. Ostlund, D. Rand, J. Sethna, E. Siggia, *Universal properties of the transition from quasi periodicity to chaos in dissipative systems*, Physica D, **8**, 303-342, 1982.