# On the stability problem for nearly–integrable Hamiltonian systems

Luigi Chierchia[1]

**Abstract:** *The problem of stability of the action variables in nearly–integrable (real–analytic) Hamiltonian systems is considered. Several results (fully described in* [CG2]*) are discussed; in particular: (i) a generalization of Arnold's method (*[A]*) allowing to prove instability (i.e. drift of action variables by an amount of order* 1*, often called "Arnold's diffusion") for general perturbations of "a–priori unstable integrable systems" (i.e. systems for which the integrable structure carries separatrices); (ii) Examples of perturbations of "a–priori stable sytems" (i.e. systems whose integrable part can be completely described by regular action–angle variables) exhibiting instability. In such examples, inspired by the "D'Alembert problem" in Celestial Mechanics (treated, in full details, in* [CG2]*), the splitting of the asymptotic manifolds is not exponentially small in the perturbation parameter.* **Abstract:** *The problem of stability of the action variables in nearly–integrable (real–analytic) Hamiltonian systems is considered. Several results (fully described in* [CG2]*) are discussed; in particular: (i) a generalization of Arnold's method (*[A]*) allowing to prove instability (i.e. drift of action variables by an amount of order* 1*, often called "Arnold's diffusion") for general perturbations of "a–priori unstable integrable systems" (i.e. systems for which the integrable structure carries separatrices); (ii) Examples of perturbations of "a–priori stable sytems" (i.e. systems whose integrable part can be completely described by regular action–angle variables) exhibiting instability. In such examples, inspired by the "D'Alembert problem" in Celestial Mechanics (treated, in full details, in* [CG2]*), the split-*

[1] Dip. di Matematica, $II^a$ Università di Roma, "Tor Vergata", via della Ricerca Scientifica, 00133 Roma, Italia. *Lecture delivered at the S. Petersburg Conference "Dynamical Systems", November 1991 describing joint work in collaboration with Giovanni Gallavotti,* Dip. di Fisica, Università di Roma, "La Sapienza", P. Moro 5, 00185 Roma, Italia

*ting of the asymptotic manifolds is not exponentially small in the perturbation parameter.*

## §1 A Theorem and a Conjecture by V. I. Arnold

Consider a Hamiltonian system with $N$ degree of freedom with a Hamiltonian of the form $H_\mu \equiv H_0 + \mu f$, where $H_0, f$ are real–analytic functions on the phase space $V \times T^N$ ($V$ being a bounded region in $R^N$, $T \equiv R/2\pi Z$) and $\mu$ is a real parameter. For $\mu = 0$, $H_0$ is assumed to be *integrable i.e.* there exist $N$ integrals, $I_1, ..., I_N$, independent and in involution (see [A1] or [G] for the standard terminology). Typically, for $\mu \neq 0$ the system will be no longer integrable, however the time for a possible variation of $\mathcal{O}(|\mu|^c)$ of one of the above integrals is (again: in typical situations) extremely long as dictated by a result of Nekhoroshev [N].

In 1964 V. I. Arnold conjectured that, *in general, for $\mu \neq 0$ small enough, there exist initial data $z \in V \times T^N$ and a time $T > 0$ such that* $|I(\phi_\mu^T(z)) - I(z)| = \mathcal{O}(1)$; here $I \equiv (I_1, ..., I_N)$ and $t \to \phi_\mu^t$ is the flow generated by the Hamiltonian $H_\mu$. The conjecture was based on the following theorem and on its proof:

**Theorem** ([A2]). *Fix $\varepsilon > 0$ and let $H_\mu \equiv H_0 + \mu f$ with:*

$$H_0 \equiv \frac{a^2}{2} + b + \frac{p^2}{2} + \varepsilon(\cos q - 1) \ , \qquad f \equiv \varepsilon(\cos q - 1)(\sin\alpha + \cos\beta) \tag{1.1}$$

*where $(a,\alpha), (b,\beta), (p,q) \in (R \times T)^3$ are standard canonical (symplectic) coordinates i.e. the equations of motion are:*

$$\begin{array}{lll} \dot\alpha = a & \dot\beta = 1 & \dot q = p \\ \dot a = \varepsilon\mu(1 - \cos q)\cos\alpha & \dot b = \varepsilon\mu(\cos q - 1)\sin\beta & \dot p = \varepsilon \sin q[1 + \mu(\sin\alpha + \cos\beta)] \end{array}$$

*For each $0 < a' < a''$ there exists a motion $(a(t), b(t), p(t), \alpha(t), \beta(t), q(t))$ and a time $T > 0$ such that $a(0) < a'$ while $a(T) > a''$, provided $\mu > 0$ is small enough.*

Before discussing the method of proof let us comment on the presence of the *two parameters* $\varepsilon, \mu$.

In the theory of nearly–integrable Hamiltonian systems one considers Hamiltonians of the form:

$$H \equiv \; h_0(J) + \varepsilon F_0(J, \psi)$$

$(J, \psi) \in V \times T^N$ being the so–called action–angle variables and $\varepsilon$ a small parameter. Under suitable non–degeneracy assumptions, in a neighborhood of a *simple resonance* (a "simple resonance" is a hypersurface in $V$ of the form $\{J \in V : \partial_J h_0 \cdot \nu_0 = 0$ for some $\nu_0 \in Z^N \backslash \{0\}$ and $\partial_J h_0 \cdot \nu \neq 0$ for all $\nu$ not parallel to $\nu_0$ $\})$ one can find, for any $Q > 0$, a canonical transformation $(J, \psi) \to (a, \alpha), (p, q) \in R^{N-1} \times T^{N-1} \times R \times T$ such that in the new variables $H$ takes the form:

$$\tilde{h}(a; \varepsilon) + P(p, q; a, \varepsilon) + \varepsilon^{Q+1} \tilde{F}(a, \alpha, p, q; \varepsilon)$$

In general, the integrable Hamiltonian $P$ (*parametrized* by $a \in V \subset R^{N-1}$ and $\varepsilon$) generates a structure similar to that of the standard pendulum (*i.e.* isolated unstable equilibria, separatrices, etc.). The relation of Arnold's Hamiltonian (1.1) to the general problem of nearly–integrable motions near simple resonances is transparent.

The proof of the above Theorem is based on the idea of *transition chains of whiskered tori*.

For $\mu = 0$ the (integrable) system with Hamiltonian $H_0$ carries a "partially–hyperbolic structure": there are *lower dimensional invariant tori*

$$\mathcal{T}(\omega) \equiv \; \{p = q = 0\} \times \{a = \omega, b = b_0\} \times \{(\alpha, \beta) \in T^2\}$$

on which the motion is *quasi–periodic*: $(\alpha, \beta) \to (\alpha + \omega t, \beta + t)$. These tori are linearly unstable and admit asymptotic manifolds called by Arnold "whiskers":

$$\begin{aligned} W^{\pm}(\mathcal{T}(\omega)) \equiv \; & \{a = \omega, b = b_0\} \times \{(\alpha, \beta) \in T^2\} \times \{(p, q) : \frac{p^2}{2} + \varepsilon(\cos q - 1) = 0\} \\ = \; & \{\text{phase points } z : \text{ dist. } (\phi_0^t(z), \mathcal{T}(\omega)) \to 0 \text{ as } t \to \pm\infty\} \end{aligned}$$

The whiskers $W^{\pm}(\mathcal{T}(\omega))$ are 3–dimensional manifolds lying in the same 5–dimensional ambient space (energy level). In general two 3–dimensional surfaces in a 5–dimensional space intersect in a line; however in the integrable (and hence degenerate) situation ($\mu = 0$) it is $W^+ \equiv \; W^-$. The perturbation $f$ in Arnold's Theorem is taken so that *all whiskered tori*

$\mathcal{T}(\omega)$ *are preserved for* $\mu \neq 0$: this, of course, is a highly non–generic property. The perturbation removes the degeneracy of the integrability and persistent whiskers will, in general, intersect transversally along a curve $\zeta$, which is a trajectory such that dist.$(\zeta(t), \mathcal{T}_\mu(\omega)) \to 0$ as $t \to \pm\infty$, where $\mathcal{T}_\mu$ denotes the continuation in $\mu$ of the persistent torus $\mathcal{T}(\omega)$. Trajectories of this type were called by Poincarè *homoclinic*.

Transversality is checked by a first order (in $\mu$) computation by means of "Poincarè [P]–Melnikov [Me]–Arnold [A2] integrals" (we shall come back on this point).

Persistent whiskers are arbitrarly close one to the other, therefore transversality yields *heteroclinic trajectories i.e.* trajectories $\zeta(t) \equiv \zeta(t;\omega,\omega')$ such that $\zeta \equiv \{\zeta(t) : t \in R\}$ $= W^+(\mathcal{T}_\mu(\omega)) \cap W^-(\mathcal{T}_\mu(\omega'))$ with $\omega \neq \omega'$.

Since in Arnold's example *all tori* are preserved, one can construct *long* chains of whiskered tori $\mathcal{T}_i \equiv \mathcal{T}_\mu(\omega_i)$ for which $|\omega_{i+1} - \omega_i| = o(|\mu|)$ (say, $|\omega_{i+1} - \omega_i| = c|\mu|^2$) and $\emptyset \neq W^+(\mathcal{T}_i) \cap W^-(\mathcal{T}_{i+1}) = \zeta_i$; $\omega_1 < a'$, $\omega_M > a''$. Here $M = \mathcal{O}((a'' - a')/c\mu^2)$ and the reason for taking $|\omega_{i+1} - \omega_i| = o(\mu)$ is related to the fact that transversality is measured by a suitable determinant, which is, in general, of $\mathcal{O}(|\mu|)$ (see below).

The idea is now to "shadow" the "pseudo–orbit" $\mathcal{P} \equiv \cup_{i=1}^M \zeta_i$ with a true orbit starting near $\mathcal{T}_1$, staying close for $0 \le t \le T$ to $\mathcal{P}$ and passing at time $T$ near $\mathcal{T}_M$. In [A2] it is claimed that this is possible as long as all $\omega_i$'s are irrational. Below we shall sketch the argument under stronger (*i.e.* "Diophantine") assumptions on the frequencies $\omega_i$: the argument will rely upon a "strong" KAM linearization around the whiskered tori.

## §2 Three classes of model problems

We shall now discuss a few generalizations of the results and methods outlined above. Rather than introducing the general setting to which our theory applies, we shall restrict here on the following three *model problems.*

Consider a *real–analytic* Hamiltonian of the form:

$$H_\mu \equiv H_0(a,p,q;\varepsilon) + \mu\varepsilon^Q f(a,\alpha,p,q)$$

with $(p,q) \in R \times T$, $(a,\alpha) \in R^{N-1} \times T^{N-1}$ standard symplectic coordinates, $Q \geq 0$, and the integrable part $H_0$ given by:

$$H_0 \equiv \; h(a) + P(p,q;a,\varepsilon) \;, \qquad P \equiv \; \frac{p^2}{2J_0} + \varepsilon g_0^2 J_0(\cos q - 1) \tag{2.1}$$

where $g_0, J_0$ may depend on $a$. We shall then consider the following three models:

$$h \equiv \; \sum_{i=1}^{N-1} \frac{a_i^2}{2J_i} \;, \quad \varepsilon > 0 \quad \text{fixed} \;, \quad |\mu| \ll 1 \;, \quad (Q \text{ arbitrary}) \tag{\textbf{M1}}$$

$$h \equiv \; \sum_{i=1}^{N-1} \frac{a_i^2}{2J_i} \;, \quad Q > Q_0, \quad 0 < \varepsilon < \varepsilon_0 \ll 1, \quad |\mu| \le 2, \quad (\text{suitable } Q_0, \varepsilon_0) \tag{\textbf{M2}}$$

$$h \equiv \; \omega_1 a_1 + \varepsilon \frac{a_2^2}{2J} \;, \quad Q, \varepsilon, \mu \text{ as in } (\textbf{M2}) \qquad (N = 3) \tag{\textbf{M3}}$$

**Remarks:** (i) To cover Arnold's example one should substitute in **(M1)** $h$ with $\sum_{i=1}^{N-2}(a_i^2/2J_i) + \omega a_{N-1}$, to which, with the due (and well known) modifications all the theory below applies (see [CG2]).

(ii) In the models **(M2)** and **(M3)** $\mu$ is just an *auxiliary parameter* as its large domain of definition allows to set it equal to 1.

(iii) The Hamiltonian $H_0$ in the third model **(M3)** (for which the above comment (ii) holds as well) is an example of *degenerate system*: for $\varepsilon = 0$, $H_0$ is independent of the action variable $a_2$. Such systems are important as they are common in Celestial Mechanics. Consider the *D'Alembert model of the Earth precession*: a planet assimilated to a rigid rotational ellipsoid with small "flattness" $\varepsilon$ (*i.e.* with "equatorial radius" $R$ and "polar radius" $R/(1+2\varepsilon)^{1/2}$), revolving on a given Keplerian orbit of eccentricity $e \equiv \; \varepsilon^Q$, around a fixed star and subject only to Newtonian gravitational forces. One can show that the model just described is a suitable generalization of **(M3)** above: see [CG2] where "Arnold's diffusion" for such a model is proved.

(iv) The model **(M1)** will be called *a–priori unstable* as the integrable part ($\mu = 0$) contains separatrices on which the motion is partially hyperbolic as already explained above, The models **(M2)**, **(M3)** are, instead, examples of *a–priori stable systems* as the perturbative (small) parameter is $\varepsilon$ [see (ii)] and separatrices may be introduced only at resonances by the perturbation (see also the comment after the Theorem in §1).

(v) In fact the theory sketched below can be extended so as to cover cases where $h(a)$ is a rather arbitrary function verifying standard non–degeneracy conditions and where $J_i, g_0$ could also depend on $p, q$ (besides depending on the actions $a$ as assumed here): this extensions allow to cover nearly–integrable situations in a neighborhood of a simple resonance (included the D'Alembert model).

## §3 Instability mechanism

Roughly speaking the mechanism for instability (or "drift" or "diffusion") of the perturbed integrals is based on the following idea.

Fixed a reference energy $E$, consider a path $\mathcal{L} \subset \{a : H_0(a, 0, 0; \varepsilon) = E\}$, piecewise analytic. Attached to each $a_0 \in \mathcal{L}$ we shall think a lower dimensional invariant torus

$$\mathcal{T}(\omega) \equiv \{\alpha \in T^{N-1},\ a = a_0,\ p = q = 0\}\ , \quad \omega \equiv \partial_a h(a_0)$$

(which is actually independent of $\varepsilon$) together with its whiskers

$$W^{\pm}(\mathcal{T}(\omega)) \equiv \{\alpha \in T^{N-1}; a = a_0; (p, q) : P(p, q; a_0, \varepsilon) = 0\}$$

According to (partially hyperbolic) KAM theory ([M],[Gr],[Z]), for general $\mathcal{L}$, "most" of the whiskered tori persist for $\mu \neq 0$ but small enough. However, *gaps* of size as big as $\mathcal{O}(\sqrt{|\mu|\varepsilon^Q})$ where none of the above tori persist, have to be expected.

It is therefore clear which are the problems to be overcome:

**(P1)** Persistence of whiskered tori emerging from the unperturbed ones attached to the path $\mathcal{L}$; smoothness properties of the whiskers (with respect to parameters, smooth interpolation properties, etc.); behaviour of the trajectories around the whiskered tori; analytic continuation for large values of $\mu$ for the models **(M2)**, **(M3)**;

**(P2)** Transversal splitting and creation of homo/heteroclinic orbits;

**(P3)** Quantitative relation between the size of gaps and the "size of transversality" (location of high–density zones of persistent tori and admissible choises of "diffusion" paths $\mathcal{L}$).

## §4 Persistence of whiskered tori and their analyticity properties

Regarding problem **(P1)** the following result holds. Let $\mathcal{L} \equiv \{a_\sigma : \sigma \in [0,1]\}$ be a piecewise analytic (in $\sigma$) path on a fixed energy level $\mathcal{E}_E \equiv \{a : H_0(a,0,0;\varepsilon) = 0\}$ and, if $\omega_\sigma \equiv \partial_a h(a_\sigma)$, let

$$\tilde{\mathcal{L}} \equiv \{a \equiv a_{\sigma\gamma} : \partial_a h(a) = \omega_\sigma(1+\gamma)\ ,\ \sigma \in [0,1], \gamma \in [-\bar\gamma, \bar\gamma]\}$$

Clearly if $\bar\gamma$ is small enough, $\tilde{\mathcal{L}}$ is $|\gamma|$–close to $\mathcal{L}$. Finally, let

$$\Sigma(C,\tau) \equiv \{\sigma \in [0,1] : |\omega_\sigma \cdot \nu|^{-1} \le C|\nu|^\tau,\ \forall\ \nu \in Z^{N-1}, \nu \ne 0\}$$

and denote by $B^n_r \equiv \{x \in R^n : |x_i| \le r\ , i = 1, ..., n\}$.

**Theorem 1** ([CG2]). **(i):** *Consider the model* **(M1)** *and fix* $k > 0$. *There exist* $\mu_0, \kappa_0, \bar\gamma > 0$ *and* $C^k$*–functions*

$$\zeta :\ (\mu,\psi,y,x,\sigma,\gamma) \in B^1_{\mu_0} \times T^{N-1} \times B^2_{\kappa_0} \times [0,1] \times B^1_{\bar\gamma} \to \zeta_\mu(\psi,y,x;\sigma,\gamma) \in R^N \times T^N$$
$$g :\ (\mu,z,\sigma,\gamma) \in B^1_{\mu_0} \times B^1_{\kappa_0^2} \times [0,1] \times B^1_{\bar\gamma} \to g_\mu(z;\sigma,\gamma) \in R$$

*which, for* $\sigma \in \Sigma(C,\tau)$ *and* $\gamma \in B^1_{\bar\gamma}$ *fixed, are analytic in the remaining arguments, and are such that:*

$$\phi^t_\mu\ \zeta_\mu(\psi,y,x;\sigma,\gamma) = \zeta_\mu(\psi + \omega_\sigma(1+\gamma)t, ye^{-g_\mu t}, xe^{g_\mu t};\sigma,\gamma)\ ,\quad (\sigma \in \Sigma(C,\tau), |\gamma| \le \bar\gamma)$$

*where* $\phi^t_\mu$ *denotes the flow generated by* $H_\mu$. *In particular*

$$\mathcal{T}_{\sigma\gamma} \equiv \{\zeta_\mu(\psi,0,0;\sigma,\gamma) : \psi \in T^{N-1}\}$$

*is a* $H_\mu$*–invariant torus with local whiskers given by:*

$$\begin{aligned} W^+_{loc}(\mathcal{T}_{\sigma\gamma}) &\equiv \{\zeta_\mu(\psi,y,0;\sigma,\gamma) : \psi \in T^{N-1}, y \in [-\kappa_0,\kappa_0]\} \\ W^-_{loc}(\mathcal{T}_{\sigma\gamma}) &\equiv \{\zeta_\mu(\psi,0,x;\sigma,\gamma) : \psi \in T^{N-1}, x \in [-\kappa_0,\kappa_0]\} \end{aligned} \tag{4.1}$$

*Finally, the value* $\gamma$ *can be chosen to be a smooth function of* $\mu, \sigma$ *and so that the energy* $H_\mu(\zeta_\mu(\psi,0,0;\sigma,\gamma))$ *has a preassigned value* $E$ *within the range of the possible unperturbed energy values.*

**(ii):** *Consider the models* **(M2)**, **(M3)**: *There exist* $Q_0 > 1$, $\varepsilon_0 > 0$ *such that for* $Q > Q_0$, $0 < \varepsilon < \varepsilon_0$ *all the above results hold with* $\mu_0 = 2$. *More precisely: Let* $\Omega \equiv \sup_{\sigma,\gamma} |\omega_{\sigma\gamma}|$; *then there exist a (universal) number* $Q_0 > 1$ *and a positive constant* $K$ *(depending on* $\tau, \kappa_0$ *and other natural parameters associated to the functions* $H_0, f$*) such that if:*

$$K\mu_0\varepsilon_0^{Q-Q_0}\ (C\Omega)^{14} < 1 \tag{4.2}$$

*then the results in* **(i)** *hold for all* $|\mu| \le \mu_0$, $0 < \varepsilon \le \varepsilon_0$, $Q > Q_0$.

¿From (4.2) it follows that the radius of analyticity in $\mu$ is in fact proportional to $\varepsilon_0^{-m}$ for a suitable $m$.

Clearly for $\mu_0$ (respectively, $\varepsilon_0$) small enough, the above functions defining the local whiskers can be analytically extended, using the flow, to "long stretches": *e.g.*, if $\pi_q$ denotes the projection on the $q$–variable, the $y, x$–domain of $\zeta$ can be extended so as to contain a segment $[-\kappa, \kappa]$ so that $|\pi_q\zeta_\mu(\psi, \pm\kappa, 0; \sigma, \gamma)|$ and $|\pi_q\zeta_\mu(\psi, 0, \pm\kappa; \sigma, \gamma)|$ are larger than $3\pi/2$; of course, in such an extended region, the local whiskers are $\mathcal{O}(\varepsilon^{Q-Q_0}|\mu|)$–close to the unperturbed ones.

The existence of whiskered tori was first established by [M] and [Gr], and, more recently, it has been re–examined by [LW]. For the joint smoothness of the parametric representation (which yields smooth interpolation of the whiskered tori) compare with the analogous results for maximal ($N$–dimensional) tori by [L1], [S], [Pö], [CG1].

One can show (see §8 of [CG2]) that from the above linearization it follows that the tori $\mathcal{T}_{\sigma\gamma}$ are *transition tori* in the sense of [A2].

Finally, notice that the $\mu$ analyticity allows to give a precise meaning to $\mu$–expansions also for the models **(M2)**, **(M3)** and yields an algorithm to compute ("perturbatively") various objects of interest.

## §5 Transversal homoclinic points

Here we discuss **(P2)**. Consider first the model **(M1)**.

The analyticity properties established in Theorem 1 imply that $W^\pm \cap \{q=\pi\}$ is a *graph* over the angles $\alpha \in T^{N-1}$. By the Implicit Function Theorem there exist real–analytic functions $x_\mu, y_\mu, \psi^\pm_\mu$ such that (dropping the $\sigma,\gamma$ parameters in the notation and denoting $\pi_q$,...,$\pi_a$ the projections on the coordinates $q$,...,$a$):

$$\pi_q\ \zeta_\mu(\psi^+_\mu(\alpha), y_\mu(\alpha), 0) = \pi = \pi_q\ \zeta_\mu(\psi^-_\mu(\alpha), 0, x_\mu(\alpha))$$
$$\pi_\alpha\ \zeta_\mu(\psi^+_\mu(\alpha), y_\mu(\alpha), 0) = \alpha = \pi_\alpha\ \zeta_\mu(\psi^-_\mu(\alpha), 0, x_\mu(\alpha))$$

Define:

$$a^+_\mu(\alpha) \equiv\ \pi_a\ \zeta_\mu(\psi^+_\mu, y_\mu, 0)\ , \qquad p^+_\mu(\alpha) \equiv\ \pi_p\ \zeta_\mu(\psi^+_\mu, y_\mu, 0)$$
$$a^-_\mu(\alpha) \equiv\ \pi_a\ \zeta_\mu(\psi^-_\mu, 0, x_\mu)\ , \qquad p^-_\mu(\alpha) \equiv\ \pi_p\ \zeta_\mu(\psi^-_\mu, 0, x_\mu)$$

Then:

$$W^\pm \cap \{q=\pi\} = \{(a^\pm_\mu(\alpha), \alpha, p^\pm_\mu(\alpha), \pi) : \alpha \in T^{N-1}\}$$

Since the $p$–variable can be eliminated by conservation of energy, we see that finding non–degenerate homoclinic intersections is equivalent to find $\alpha_0$ such that:

$$\Delta(\alpha_0;\mu) \equiv\ a^+_\mu(\alpha_0) - a^-_\mu(\alpha_0) = 0\ , \quad D(\alpha_0;\mu) \equiv\ \det[\partial_\alpha \Delta(\alpha_0)] \neq 0 \tag{5.1}$$

The quantity $D(\alpha_0;\mu)$ is *a measure of the transversality* of the intersection between $W^+$ and $W^-$; in fact, the eigenvalues of $\partial_\alpha\Delta$ are related to the geometric angles between $W^+\cap\{q=\pi\}$ and $W^-\cap\{q=\pi\}$ (of course such geometry depends upon our coordinates). Notice that $\Delta(\alpha_0;0) \equiv\ 0$ by the degeneracy of the unperturbed system.

Theorem 1 and a perturbative study of the linearized (in $\mu$) Hamiltonian equations yield easily the following first–order evaluation of $\Delta$ (compare [P], [Me], [A2]):

**Proposition 1.** *In all the above model* **(M1)** $\div$ **(M3)**, *$\Delta$ is analytic both in $\alpha$ and $\mu$ and if $\Delta \equiv\ \mu\Delta_1(\alpha) + \mu^2\Delta_2(\alpha) + ...$ then:*

$$\Delta_1(\alpha) = \partial_\alpha m_f$$

*with*

$$m_f(\alpha;\sigma,\gamma) \equiv\ -\sum_{0\neq\nu\in Z^{N-1}} e^{i\nu\cdot\alpha} \int_{-\infty}^{+\infty} \frac{e^{i\omega_{\sigma\gamma}\cdot\nu t}}{i\omega_{\sigma\gamma}\cdot\nu} \partial_t[e^{i\nu\cdot\vartheta_t} \hat f_\nu(a_{\sigma\gamma}, p_t, q_t)]dt \tag{5.2}$$

*where $\hat{f}_\nu$ are the Fourier coefficients of $\alpha \to f(a,\alpha,p,q)$, $(p_t,q_t)$ is the unperturbed separatrix motion starting at $q=\pi$ and $\vartheta_t$ is such that $\pi_q\ \phi_0^t(a_{\sigma\gamma},\alpha,p_0,\pi) = \alpha + \omega_{\sigma\gamma} t + \vartheta_t$.*

Thus, for the model **(M1)** (where $\mu$ is small), by the standard Implicit Function Theorem, *homoclinic points* (in the transversal section $\{q=\pi\}$ ) correspond to *non–degenerate critical points* of the periodic function $\alpha \to m_f(\alpha)$. And it is easy to see that, *generically*, $m_f$ has at least $2^{N-1}$ non–degenerate critical points. The determinant $D$ will be (again: in general) of $O(|\mu|)$ and therefore in order to establish *drift along a path* $\bar{\mathcal{L}} \equiv \{\bar{a}_\sigma : \sigma \in [0,1]\}$ one will have to check that such a path has a density of persistent whiskered tori of $o(|\mu|)$; see next section. Here and below $\bar{\mathcal{L}}$ will denote a path $\{a_{\sigma\gamma(\sigma)} : \sigma \in [0,1]\}$ where the value $\gamma(\sigma)$ has been fixed so that $\bar{\mathcal{L}}$ belongs to the same energy level (compare **(i)** in Theorem 1).

For **(M2)** the situation is more delicate as, in general, $\partial_\alpha \Delta_1$ is *exponentially small in* $\varepsilon$. More precisely, assume (for simplicity) that $f$ is a *trigonometric polynomial* in the angular variables $\alpha, q$ and recall the analyticity properties of the unperturbed separatrix motion:

$$p_t = \pm 2\sqrt{\varepsilon}\bar{J}\bar{g}(\cosh\sqrt{\varepsilon\bar{g}}t)^{-1}\ , \quad q_t = 4\arctan e^{\pm\sqrt{\varepsilon\bar{g}}t}\ , \quad \vartheta_t = -2\partial_a(\sqrt{\varepsilon\bar{g}}\bar{J})\tanh\sqrt{\varepsilon\bar{g}}t$$

where $\bar{J} \equiv J_0(\bar{a}_\sigma)$, $\bar{g} \equiv g_0(\bar{a}_\sigma)$. In particular, the integrand in (5.2) is holomorphic in $t$ in a strip $|\ \mathrm{Im}\ t| < \pi/(2\sqrt{\varepsilon\bar{g}})$. Therefore, shifting the contour of integration to

$$\mathrm{Im}\ t = \ \mathrm{sign}(\omega_{\sigma\gamma}\cdot\nu)\big(\frac{\pi}{2\sqrt{\varepsilon\bar{g}}} - r\big)$$

(with any $r>0$) one obtains the estimate:

$$\sup_{\alpha\in T^{N-1}} |\partial_\alpha^j m_f| \le k_j \sup_{\nu\neq 0}[e^{-|\omega_{\sigma\gamma}\cdot\nu|\pi/(2\sqrt{\varepsilon\bar{g}})}e^{-|\nu|}] \tag{5.3}$$

where $k_j$ is a suitable constant. Under our Diophantine assumption on the frequencies $\omega_{\sigma\gamma}$ the supremum in the right hand side of (5.3) can be bounded by (or, better: is essentially equal to) $c_1\exp(-c_2/\varepsilon^{\frac{(\tau+1)}{2}})$, for suitable positive constants $c_1, c_2$.

In the case **(M3)** the frequency vector has the form $\omega = (\omega_1, \varepsilon\bar{\omega}_2)$, for a suitable $\bar{\omega}_2(a)$; thus the *supremum in* (5.3) *is of order* 1. However what is really important here is to estimate from below $D \equiv \det\Delta$ and one can show that, even in the case **(M3)**, $D|_{\mu=1}$ is *exponentially small in* $\varepsilon$. But this phenomenon is an *accident*: indeed, in general, a second order (in $\mu$) computation shows that $D|_{\mu=1} = \mathcal{O}(\varepsilon^M)$ for a suitable $M > 1$; see §7 below for a precise statement.

## §6 Density of whiskered tori

The strategy at this point is clear: it remains to find , on the energy surface $\mathcal{E}_E$, "diffusion paths" where the gaps ( $\equiv$ "intervals of non–persistent whiskered tori") are smaller than the size of $D$ (*i.e.* intuitively, of the "smallest" non–trivial angle formed by vectors tangent to the whiskers).

High–density zone of persistent tori may be detected by using the theory of normal forms in the way it is used in Nekhoroshev's Theorem ([N], [BG]). Then, choosing suitably the Diophantine constant $C$ (recall the definition of $\Sigma(C,\tau)$ in §4) as a function of $\mu$ or $\varepsilon$, one can prove the following results concerning **(M1)** and **(M3)**.

Before we need a **definition**: *we say that a path* $\bar{\mathcal{L}} \subset \mathcal{E}_E$ *has "residual measure"* $\rho$ *if the Lebesgue measure of* $[0,1] \setminus \Sigma_{\text{persistent}}$, *with* $\Sigma_{\text{persistent}} \equiv \{\sigma \in [0,1]$ *s.t. to* $a_\sigma$ *it can be associated a persistent whisker (2.1) with the properties described in Theorem 1* $\}$, *is bounded above by* $\rho$. The residual measure gives an upper bound on the maximal size of gaps in $\bar{\mathcal{L}}$.

Now, assume (for simplicity) that $f$ is a trigonometric polynomial in the angle variables $\alpha, q$.

**Proposition 2.** *[Case* **(M1)***] Let* $m > 1$. *There exist a* $\mu_1 < \mu_0$, *a* $\tau_0 \geq N-1$, *an open* $(N-2)$*–dimensional set* $\mathcal{A}_{f,m} \subset \mathcal{E}_E$ *and a function* $C = C_{\mu,m}$ *such that* $\forall\ |\mu| < \mu_1$, $\tau \geq \tau_0$ *any two points of* $\mathcal{A}_{f,m}$ *can be joined by a path* $\bar{\mathcal{L}}$ *with* $\Sigma(C_{\mu,m},\tau_0) \subset \Sigma_{\text{persistent}}$ *and residual measure of* $\mathcal{O}(|\mu|^m)$. *Furthermore, the set* $\mathcal{E}_E \setminus \mathcal{A}_{f,m}$ *itself has* ($N-2$ *dimensional) measure of* $\mathcal{O}(|\mu|^m)$.

**Proposition 3.** *[Case* **(M3)***] There exist* $a_2$*–intervals* $(a,b)$ *of order* $\mathcal{O}(1)$, *constants*

*$Q_1, \varepsilon_1$ with $Q_1 > Q_0$, $0 < \varepsilon_1 < \varepsilon_0$ and a function $C = C_\varepsilon$ such that $\forall\ 0 < \varepsilon < \varepsilon_1$, $|\mu| \le 2$ the path $\bar{\mathcal{L}} \equiv \{(a_1(\sigma), a + \sigma(b-a))\}$ ($a_1$ being determined by $\bar{\mathcal{L}} \subset \mathcal{E}_E$) has $\Sigma(C_\varepsilon, 1) \subset \Sigma_{\rm persistent}$ and residual measure of $\mathcal{O}(\exp(-c/\sqrt{\varepsilon}))$, $c$ being a suitable positive constant.*

While Proposition 2 is easily obtained by performing a *finite number* of "classical perturbative steps", Proposition 3 involves the use of high order normal forms as in the "analytic part" of Nekhoroshev theory (see [N] and [BG]).

We also remark that the sentence "there exist" in the above propositions really means "one can construct", *i.e.* it is possibly to decide wether a certain region of phase space has a high density of persistent tori.

It is clear that the above results are sufficient in order to get for **(M1)** and **(M3)** the so–called Arnold's diffusion *i.e.* a drift of order 1 in action variables no matter how small the perturbation is (as long as certain non–degeneracies are verified; see also next section). The model **(M2)** is more difficult as indeed the size of $D|_{\mu=1}$ is smaller than any power in $\varepsilon$ (see [L2], [GLT], [HMS], [DS] for lower dimensional discussions) and we refrain to formulate a density result as further analysis is needed.

## §7 Instabilities and Arnold's diffusion

¿From Proposition 2 and the results of §5, it follows easily, for the model **(M1)**, the existence of "diffusion paths" *i.e.* of paths having a density of persistent whiskered tori higher than the measure of the splitting of the whiskers; generiticity being related to the condition:

$$\inf_{\sigma\in\Sigma} |\det \partial_\alpha^2 m_f| \geq \delta > 0 \tag{6.1}$$

If such a condition (easily checked on explicit examples including Arnold's one, at least for subintervals of $\mathcal{O}(1)$) holds, any two points of $\mathcal{A}_f$ can be joined by a path $\bar{a}_\sigma$. Then using the linearization of Theorem 1, one can prove that the time $T_\mu$ needed for the $a$–variable to *drift along the path* $\bar{\mathcal{L}}$ can be *bounded below* by:

$$T_\mu \leq c_3 e^{c_4/|\mu|^2}$$

for suitable positive constants $c_3, c_4$; such constants depends of course on $\varepsilon$ (which for **(M1)** is fixed) and indeed $c_3$ is exponentially small in $\varepsilon$.

We collect now the above results for the model **(M3)** with

$$f \equiv \quad \cos(\alpha_1 + q) + \cos(\alpha_2 + q) \tag{7.1}$$

in the following

**Theorem 2**. *Consider* **(M3)** *and set* $\mu = 1$ *i.e. consider* $H_1 \equiv \ H_0(a_1, a_2, p, q; \varepsilon) + \varepsilon^Q f$ *with* $f$ *as in (7.1). Then the above Theorem 1 and Proposition 3 apply for* $Q > Q_1$ *and* $0 < \varepsilon < \varepsilon_1$. *Furthermore* $\alpha_0 \equiv \ (0,0)$ *(at* $\phi = \pi$*) is a "transversal homoclinic point" i.e. a non–degenerate solution of (5.1) and* $D(\alpha_0; 1) = \mathcal{O}(\varepsilon^m)$ *for a suitable* $m > 1$. *Along any* $\mathcal{O}(1)$*–path* $\bar{\mathcal{L}}$ *as in Proposition 3 "Arnold's drift or diffusion" takes place i.e. there exist initial data* $z_\varepsilon(0)$ *generating a trajectory* $z_\varepsilon(t)$ *which for* $t = 0$ *has* $a_2$*–coordinate* $\varepsilon$*–close to one extreme of* $\bar{\mathcal{L}}$ *while for a suitable* $t = T_\varepsilon$ *it has* $a_2$*–coordinate* $\varepsilon$*–close to the other extreme of* $\bar{\mathcal{L}}$ *[recall that* length *of* $\bar{\mathcal{L}} = \mathcal{O}(1)$*]. The time* $T_\varepsilon$ *can be bounded below by* $c_5 \exp(1/\varepsilon^{c_6})$ *for suitable constants* $c_5, c_6$.

As already mentioned above this theorem is a consequence of Theorem 1, Proposition 3 and of a second order (in $\mu$) analysis of the Melnikov integrals showing that the "homoclinic angles" (better: $D$) are *polynomially small in* $\varepsilon$ (the higher $\mu$–orders being easily controlled because of the large radius of $\mu$–convergence).

**References**


[A1] Arnold, V., I.: *Mathematical methods of classical mechanics*, Springer Verlag, 1978.

[A2] Arnold, V.: *Instability of dynamical sistems with several degrees of freedom*, Sov. Mathematical Dokl., 5, 581-585, 1966.

[BG] Benettin, G., Gallavotti, G.: *Stability of motionions near resonances in quasi-integrable hamiltonian systems*, J. Statistical Physics, 44, 293-338, 1986.

[CG1] Chierchia, L., Gallavotti, G.: *Smooth prime integrals for quasi-integrable Hamiltonian systems* Il Nuovo Cimento, 67 B, 277-295, 1982.

[CG2] Chierchia, L., Gallavotti, G.: *Drift and Diffusion in Phase Space* Preprint 1992, 1–135

[DS] Delshams, A., Seara, M.T.: *An asymptotic expression for the splitting of separatrices of rapidly forced pendulum*, preprint 1991.

[G] Gallavotti, G.: *The elements of Mechanics*, Springer, 1983.

[GLT] Gelfreich, V. G., Lazutkin, V.F., Tabanov, M.B.:*Exponentially small splitting in Hamiltonian systems*, Chaos, 1 (2), 1991.

[Gr] Graff, S.M.: *On the conservation for hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations 15, 1-69, 1974.

[HMS] Holmes, P., Marsden, J., Scheurle,J: *Exponentially Small Splittings of Separatrices in KAM Theory and Degenerate Bifurcations*, Preprint, 1989.

[La1] Lazutkin, V. F.: *The existence of caustics for a billiard problem in a convex domain*, Izv. Akad. Nauk. SSSR, 37 (1), 1973

[La2] Lazutkin, V.F.: *Separatrices splitting for standard and semistandard mappings*, Preprint, 1989.

[LW] de la Llave, R., Wayne, E.: *Whiskered Tori*, preprint 1990

[M] Moser, J.: *Convergent series expansions for quasi-periodic motions*, Mathematische Annalen, 169, 136-176, 1967.

[Me] Melnikov, V.K.: *On the stability of the center for time periodic perturbations*, Trans. Moscow Math Math. Soc., 12, 1-57, 1963.

[N] Nekhorossev, N.: An exponential estimate of the time of stability of nearly integrable hamiltonian systems, Russian Mathematical Surveys, 32, 1-65, 1975.

[P] Poincarè, H.: *Les Méthodes nouvelles de la mécanique céleste*, 1892, reprinted by Blanchard, Paris, 1987.

[Pö] Pöschel, J.: *Integrability of Hamiltonian systems on Cantor sets*, Communications Pure Appl. Math, 35, 653-696, 1982.

[Sv] Svanidze, N.V.: *Small perturbations of an integrable dynamical system with an integral invariant*, Proceed. Steklov Institute of Math., 2, 1981

[Z] Zehnder, E.: *Generalized implicit function theorems with applications to some small divisors problems* I,II, Communications Pure Applied Mathematics, 28, 91-140, 1975.