Fast Arnold's diffusion in isochronous systems

G. Gallavotti

Università di Roma 1, Fisica
9 settembre 1997

Abstract: an illustration of a mechanism for Arnold's diffusion following a nonvariational approach and finding explicit estimates for the diffusion time.

Keywords: Arnold's diffusion, homoclinic splitting, KAM

§1: Introduction.

Arnold's diffusion has been established for the simple example proposed by Arnold, [A], following a nonvariational method, [CG], and variational methods, [Be], [Br]. The nonvariational method yields estimates that are terribly big; the variational method instead gives better estimates, "*fast*",([Be]) and even very good, "polynomial", ones ([Br]).

Here I illustrate the method of [CG] by developing it with the aim of showing the existence of diffusion, without actually constructing time scales bounds on the diffusing trajectories. This may lead to a clarification of a method which maintains its interest in spite of the better estimates coming from variational methods because it is the only one which, so far, is robust enough to apply to anisochronous systems.

If explicit estimates are avoided one gains enormously in simplicity: this kind of approach was probably the one meant in [A] where the problem was first posed and solved without bothering to give the (fairly obvious, see §5) details. What follows applies also to the Arnold's case, but I prefer to illustrate it in a case that is even simpler.

Furthermore I show that if a new idea is added to the method of [CG] then one can get a "fast" (still exponential) estimate for the drift time.

Here we consider hamiltonians $\mathcal{H}$ with three degrees of freedom described by coordinates $I \in R$, $\underline{A} = (A_1, A_2) \in R^2$ and angles $\varphi \in T^1$, $\underline{\alpha} = (\alpha_1, \alpha_2) \in T^2$:

$$\mathcal{H} = \underline{\omega} \cdot \underline{A} + \frac{I^2}{2} + g^2(\cos\varphi - 1) + \varepsilon f(\varphi, \underline{\alpha}) \tag{1.1}$$

where $\underline{\omega} = (\omega_1, \omega_2) \in R^2$ is a vector with diophantine constants C, τ, *i.e.* such that for all integer components vectors $\underline{\nu} = (\nu_1, \nu_2)$ it is $|\underline{\omega} \cdot \underline{\nu}|^{-1} \le C|\underline{\nu}|^\tau$ if $\underline{\nu} \ne \underline{0}$; the *perturbation* f is supposed to be a (fixed) trigonometric polynomial of degree N: $f(\varphi, \underline{\alpha}) = \sum_{0<|\underline{\nu}|<N, |n|<N} f_{n,\underline{\nu}} \cos(n\varphi + \underline{\nu} \cdot \underline{\alpha})$. The subject being fairly well understood we do not need to be really very careful about units so that some coefficients in (1.1) have been set equal to 1.

One can also use the well known Jacobi's hyperbolic coordinates p_0, q_0 to describe the pendulum $\frac{1}{2}I^2 + g^2(\cos\varphi - 1)$ near the unstable point $I = \varphi = 0$, see appendix A1. In the new coordinates, which we denote p_0, q_0, the pendulum hamiltonian becomes $J(p_0 q_0)$ with $J'(x) \stackrel{def}{=} \frac{dJ(x)}{dx} = g + \sum_{n=1}^\infty g_n x^n \stackrel{def}{=} g(x)$ and the total hamiltonian becomes:

$$\mathcal{H} = \underline{\omega} \cdot \underline{A}_0 + J(p_0 q_0) + \varepsilon f_0(\underline{\alpha}, p_0, q_0) \tag{1.2}$$

where $\underline{A}_0 \equiv \underline{A}$, $\underline{\alpha}_0 = \underline{\alpha}$ and $f_0(\underline{\alpha}, p_0, q_0) = f(\varphi, \underline{\alpha})$.

The function f_0, still a trigonometric polynomial in $\underline{\alpha}$, has the property: $f_0(\underline{\alpha}, p, q) = f_0(-\underline{\alpha}, q, p) = f_0(-\underline{\alpha}, -p, -q)$ and we shall call *parity* the 4–elements group of transformations generated by $P_1 : (\underline{\alpha}, p, q) \longleftrightarrow (-\underline{\alpha}, q, p)$ and $P_2 : (\underline{\alpha}, p, q) \longleftrightarrow (-\underline{\alpha}, -p, -q)$. We call P_0, P_1, P_2, P the group elements; we say that f_0 has "even parity". If $F(\underline{\alpha}, p, q) =$

$-F(P_j(\underline{\alpha},q,p)) \stackrel{def}{=} P_jF(\underline{\alpha},q,p)$, $j=1,2$, we say that F has odd parity: for instance $\partial_{\underline{\alpha}} f_0$ has "odd parity". The p derivative F or the q derivative G of an even function have the property $F=P_1G=-P_2F, G=P_1F=-P_2G$. The Jacobi's map in general transforms functions of φ, $\underline{\alpha}$ with given parity in $\underline{\alpha},\varphi$ in the ordinary sense into functions with the same parity in the $(\underline{\alpha},p,q)$ variables.

§2. Invariant tori and nearby flow.

We look for a change of coordinates $(\underline{A}_0,\underline{\alpha}_0,p_0,q_0) \longleftrightarrow (\underline{A},\underline{\psi},p,q)$ which integrates locally (1.2) near the unstable equilibrium of the pendulum. More precisely so that in the new coordinates the motion is:

$$\underline{A} = const,\ \ \underline{\psi} \to \underline{\psi} + \underline{\omega}\,t,\ \ p \to pe^{-(1+\gamma)gt},\ \ q \to qe^{(1+\gamma)gt} \tag{2.1}$$

where $g=g(pq), \gamma=\gamma(pq)$ and $x \to \gamma(x)$ is a suitable function analytic in x near $x=0$ while $g(x)=J'(x)$ (see (1.2). We shall attempt to write the change of coordinates:

$$\begin{aligned} \underline{A}_0 &= \underline{A} + \underline{H}(\underline{\psi},p,q), \qquad & p_0 &= p + L(\underline{\psi},p,q) \\ \underline{\alpha} &= \underline{\psi}, & q_0 &= q + \tilde{L}(\underline{\psi},p,q) \stackrel{def}{=} q + L(-\underline{\psi},q,p) \end{aligned} \tag{2.2}$$

where $\underline{H}$ has zero $\underline{\psi}$ average and even parity. Setting $\not{\partial} \stackrel{def}{=} q\partial_q - p\partial_p$ and imposing that (2.2) and (2.1) verify the equations of motion one gets the equations:

$$\begin{aligned} \big(g(x)\not{\partial} + \underline{\omega}\cdot\partial_{\underline{\psi}}\big)\,\underline{H} &= -\,\varepsilon\partial_{\underline{\alpha}} f_0(\underline{\psi},p_0,q_0) - g(x)\gamma(x)\not{\partial}\,\underline{H} \\ \big(g(x) + g(x)\not{\partial} + \underline{\omega}\cdot\partial_{\underline{\psi}}\big)\,L &= -\,\varepsilon\partial_{q_0} f_0(\underline{\psi},p_0,q_0) - (g(x_0)-g(x))p_0 + \\ &\quad + \gamma(x)g(x)p - g(x)\gamma(x)\not{\partial}\,L \end{aligned} \tag{2.3}$$

where $x=pq$, $x_0=p_0q_0$, $p_0=p+L$, $q_0=q+\tilde{L}$. In fact the second equation is independent on the first. Both can be shown to admit, for ε small enough as we shall always suppose below, a solution analytic in ε and divisible by ε. Note that the unknown are $\underline{H}, L, \gamma$. A proof is essentially in the basic paper [Ge]: it follows the Eliasson's method, [E], as developed in [G2], [GG]. A "classical" (*i.e.* by quadratic iterations) proof can be derived from §5 in [CG] where the harder anisochronous case is detailed.

§3. Stable and unstable manifolds.

From (2.2) we can read the following facts:

(1) Phase space contains a family of invariant tori $\mathcal{T}(\underline{A})$ parameterized by $\underline{A} \in R^2$ and obtained by setting $p=q=0$. The average position of the tori is precisely $\underline{A}$, because $\underline{H}$ has zero average: average with respect to $\underline{\psi}$ or to time. This is a general property of Thirring's models (like (1.1), see [T] and [G2]) "twistless tori".

(2) Given $\underline{A}$ and setting $q=0, p\neq 0$ one obtains a surface whose points are parameterized by $\underline{\psi}, p$ and which is a local piece of the stable manifold $W^s(\underline{A})$ of $\mathcal{T}(\underline{A})$. The quantity $-(1+\gamma(x))g(x)$ is the "Lyapunov exponent" of $W^s(\underline{A})$. Likewise, setting $p=0$, $q\neq 0$ one defines a local piece of the unstable manifold $W^u(\underline{A})$ of $\underline{A}$ and $(1+\gamma(x))g(x)$ is the corresponding exponent.

(3) Given $\underline{A}$ and setting $q\neq 0, p\neq 0$ one parameterizes the rest of phase space *near* $\mathcal{T}(\underline{A})$. In this part of phase space the motion is in some sense very regular.

(4) However the motions, just described locally, are globally more interesting and chaotic. In fact generically $W^u(\underline{A}), W^s(\underline{A}')$ do intersect transversally if $\underline{A}, \underline{A}'$ are close enough (depending on ε) and $\mathcal{T}(\underline{A}), \mathcal{T}(\underline{A}')$ have the same energy. In such cases $W^u(\underline{A}) \cap$

$W^s(\underline{A}\,')$ consists of trajectories, or *heteroclinic intersections*, running asymptotically around $\mathcal{T}(\underline{A}\,)$ as $t\to-\infty$ and around $\mathcal{T}(\underline{A}\,')$ as $t\to+\infty$.

The symmetry of the problem implies, see [CG], that if $\underline{A}\,=\,\underline{A}\,'$ and $\varphi=\pi$ then the point of $W^s(\underline{A}\,)$ with $\varphi=\pi$, $\underline{\alpha}\,=\,\underline{0}$ is *homoclinic, i.e.* it is on the trajectory $W^u(\underline{A}\,)\cap W^s(\underline{A}\,)$. The intersection $W^u(\underline{A}\,)\cap W^s(\underline{A}\,')$ exists for all $\underline{A}\,,\,\underline{A}\,'$ close enough and is generically *transversal* in the sense that if we fix $\varphi=\pi$ (or φ to any other value $\ne 0,2\pi$) then any pair of tangents to $W^u(\underline{A}\,)$ and $W^s(\underline{A}\,')$ at common points form an angle $\ge\mu>0$; the bound μ depends on ε, of course, and it is generically proportional to ε.

(5) Finally a definition: Let $\underline{A}_0,\underline{A}_1,\ldots,\underline{A}_{\mathcal{N}}$ be a sequence such that $|\underline{A}_j-\underline{A}_{j+1}|$ is so small that $W^s(\underline{A}_j)\cap W^u(\underline{A}_{j+1})$ have a transversal heteroclinic intersection, in the above sense, with intersection angles $\ge\mu$ at $\varphi=\pi$. We call such a chain a *heteroclinic chain* or *ladder.* One finds in various simple examples $\mu=O(\varepsilon)=O(\mathcal{N}^{-1})$, see [CG]: for a general theory of the splitting see [GGM] and appended references.

We shall prove the following theorem ("Arnold's diffusion" or "drift"):

Theorem 1: *Let $\underline{A}_0,\underline{A}_1,\ldots,\underline{A}_{\mathcal{N}}$ be a heteroclinic chain: for any $\delta>0$ there are trajectories starting within δ of $\mathcal{T}(\underline{A}_0)$ and arriving after a finite time T within δ of $\mathcal{T}(\underline{A}_{\mathcal{N}})$.*

This theorem is proved in [CG]; I prove it here againg along the lines of [CG]: the purpose being of showing the conceptual difference with respect to the variational approaches, which accounts for the impressive difference in the time scale of T compared with [Br] or with the estimate in theorem 2 below (see (6.5)).

§4. Geometric concepts.

Let $2\,\kappa>0$ be smaller than the radius of the disk in the (p,q) plane where the functions in (2.2) are defined. We call κ a "target parameter".

To visualize the geometry of the problem involving 2–dimensional tori and their 3–dimensional stable and unstable manifolds, in the 5–dimensional energy surface, we shall need the following geometric objects:

•(a) a point X_i, heteroclinic between $\mathcal{T}(\underline{A}_i)$ and $\mathcal{T}(\underline{A}_{i+1})$, which has local coordinates, see (2.2), $X_i=(\underline{A}_i,\underline{\psi}_i,0,\kappa)$.

•(b) the equations, at fixed $q=\kappa$, of the connected part of $W^s(\underline{A}_{i+1})$ containing X_i, in the local coordinates near $\mathcal{T}(\underline{A}_i)$; they will be written as:

$$Y_i(\underline{\psi}\,)=(\underline{A}^{\,s}_{i+1}(\underline{\psi}\,),\,\underline{\psi}\,,p^s_{i+1}(\underline{\psi}\,),\kappa) \tag{4.1}$$

with $|\underline{\psi}-\underline{\psi}_i|<\zeta$ for some $\zeta>0$ (i–independent): it is $\underline{A}^{\,s}_{i+1}(\underline{\psi}_i)=\underline{A}_i$, $p^s_{i+1}(\underline{\psi}_i)=0$ because we require $Y_i(\underline{\psi}_i)=X_i$. There are constants F',F such that $|\underline{A}^{\,s}_{i+1}(\underline{\psi}\,)-\underline{A}^{\,s}_{i+1}(\underline{\psi}_i)|$ and $\max_{|\underline{\psi}-\underline{\psi}_i|=fixed}|p^s_{i+1}(\underline{\psi}\,)|$ are bounded, for ζ small enough, below by $F'|\underline{\psi}-\underline{\psi}_i|$ and above by $F|\underline{\psi}-\underline{\psi}_i|$; the constants F',F have size $O(\mu)$.

Note that $W^s(\underline{A}_{i+1})$ also contains a part with local equations $(\underline{A}_{i+1},\underline{\psi}\,,p,0)$ which is *not* to be confused with the previous one described by the function $\overline{Y_i(\underline{\psi}\,)}$. This is more easily understood by looking at the meaning of the above objects in the original $(\underline{A}\,,\underline{\alpha}\,,I,\varphi)$ coordinates: in a way the first part of $W^s(\underline{A}_{i+1})$ is close to $\varphi=0$ and the second to $\varphi=2\pi$. They can be close because of the periodicity, but they are conceptually quite different.

•(c) a point $P_i=Y_i(\tilde{\underline{\psi}}_i)$ with $|\tilde{\underline{\psi}}_i-\underline{\psi}_i|=r_i$, where $\tilde{\underline{\psi}}_i,r_i$ will be determined recursively, and a neighborhood B_i:

$$B_i = \{|\underline{A} - \underline{A}^s_{i+1}(\underline{\psi})| < \rho_i,\ |\underline{\psi} - \tilde{\underline{\psi}}_i| < \rho_i,\ |p^s_{i+1}(\underline{\psi}) - p| < \rho_i,\ q = \kappa\} \tag{4.2}$$

where $\rho_i < r_i$ is another length to be determined recursively. If $\overline{g}, 2\overline{g}$ are a lower and upper bound to $(1+\gamma(x))g(x)$ for $|x| < 4\kappa^2$, the point P_i evolves in a time $T_i \simeq \overline{g}^{-1}\log\kappa^{-1}$ into a point X'_i near $\mathcal{T}(\underline{A}'_{i+1})$ which has local coordinates $X'_i = (\underline{A}_{i+1}, \underline{\psi}'_i, \kappa, 0)$.

•(d) The points ξ of the set B_i are mapped by the time evolution to points that, at the beginning at least, come close to $\mathcal{T}(\underline{A}_{i+1})$ and in a time $\tau(\xi)$ acquire local coordinates near $\mathcal{T}(\underline{A}_{i+1})$ with $p = \kappa$ exactly: the time $\tau(\xi)$ is of the order of $\overline{g}^{-1}\log\kappa^{-1}$.

If S_t is the time evolution flow for the system (1.1) we write $S\xi = S_{\tau(\xi)}\xi$ (note that S depends also on i). Then S maps the set B_i into a set SB_i containing:

$$B'_i = \{|\underline{A} - \underline{A}_{i+1}| < \frac{1}{E}\rho_i,\ |\underline{\psi} - \underline{\psi}'_i| < \frac{1}{E}\rho_i,\ p = \kappa,\ |q| < \frac{1}{E}\rho_i\} \tag{4.3}$$

because all the points in B_i with $\underline{A} = \underline{A}^s_{i+1}(\underline{\psi})$, $p = p^s_{i+1}(\underline{\psi})$, $q = \kappa$ evolve to points with $\underline{A} = \underline{A}_{i+1}$, $p = \kappa$, $q = 0$ and $\underline{\psi}$ close to $\underline{\psi}'_i$, by the definitions. Here E is a bound on the jacobian matrix of S (which, being essentially a flow over a time $O(\overline{g}^{-1}\log\kappa^{-1})$, has derivatives bounded i–independently: since we suppose that ε is "small enough" we could take $E = 1 + b\varepsilon$ for some $b > 0$ if, as often the case, $|\underline{A}_i - \underline{A}_{i+1}| < O(\varepsilon)$).

§5. The [CG]-method of proof of the theorem.

Consider the points $Y^s_{i+1}(\underline{\psi}) \in W^s(\underline{A}_{i+2})$ with coordinates $(\underline{A}^s_{i+2}(\underline{\psi}), \underline{\psi}, p^s_{i+2}(\underline{\psi}), \kappa)$. They will evolve backwards in time so that $\underline{A}$ stays constant, $\underline{\psi}$ evolves quasiperiodically hence "rigidly", and $p^s_{i+2}(\underline{\psi})$ evolves to κ while the q–coordinate evolves from κ to $q = p^s_{i+2}(\underline{\psi})$ (because pq stays constant, see (2.1)). The time for this evolution is $T_{\underline{\psi}} \simeq \overline{g}^{-1}\log\kappa|p^s_{i+2}(\underline{\psi})|^{-1} \xrightarrow[\underline{\psi}\to\underline{\psi}_{i+1}]{} +\infty$.

Therefore there is a sequence $\underline{\psi}^n \ne \underline{\psi}_{i+1}$ such that $\underline{\psi}^n \to \underline{\psi}_{i+1}$, $p^s_{i+2}(\underline{\psi}^n) \to 0$ $\underline{A}^s_{i+2}(\underline{\psi}^n) \to \underline{A}_{i+1}$ *and* $\underline{\psi}^n - \underline{\omega}T_{\underline{\psi}^n} \xrightarrow[n\to\infty]{} \underline{\psi}'_i$, as a consequence of the diophantine properties of $\underline{\omega}$. So that there is $\tilde{\underline{\psi}}_{i+1} \stackrel{def}{=} \underline{\psi}^n$ with n large enough and a point $P_{i+1} = (\underline{A}^s_{i+2}(\tilde{\underline{\psi}}_{i+1}), \tilde{\underline{\psi}}_{i+1}, p^s_{i+2}(\tilde{\underline{\psi}}_{i+1}), \kappa) \in W^s(\underline{A}_{i+2})$ (actually infinitely many) which evolves, backwards in time, from P_{i+1} to a point of B'_i.

Hence we can define $r_{i+1} = |\tilde{\underline{\psi}}_{i+1} - \underline{\psi}_{i+1}|$ and ρ_{i+1} small enough so that the backward motion of the points in B_{i+1} enters in due time into B'_i. It follows that the set B_i evolves in time so that all the points of B_{i+1} are on trajectories of points of B_i. Hence all points of $B_{\mathcal{N}}$ wil be reached by points starting in B_0.

This completes the proof. All constants can be computed explicitly, even though this is somewhat long and cumbersome, see [CG]. The result is an extremely large diffusion time T (namely the value at $\mathcal{N}$ of a composition of $\mathcal{N}$ exponentials! at least this is the estimate I get after correcting an error in §8 of [CG]: the error is minor but leads to substantially worse bounds).

Nevertheless the estimate that comes out of the above scheme seems essentially optimal. And then the problem is: "how is it possible that by other methods (*e.g.* variational methods of [Be],[Br]) one can get *far better* estimates?

A reason may be that the variational methods are less constructive: less so than the above. The "fast drifting" trajectory exists but there seems to be no algorithm to determine it, not even the sequence of its "close encounters" with the invariant tori that generates drift: which is in fact *preassigned* in the above method. This certainly can account for a difference in the estimates. In fact the above construction is far too rigid: we pretend not only that drift takes place but also that it takes place via a path that visits closely a *prescribed sequence* of tori in an essentially *predetermined* way. In §6 a less constructive method is proposed and used to obtain bounds: which, however, are still far from polynomial.

§6. Fast diffusion: elastic heteroclinic chains.

The following adds a new idea to the method of [CG], exposed in §5, allowing us to improve the superexponential estimate of [CG]. Below ε will be fixed small enough, and $\overline{g}$ will be a lower bound to $g(x)(1+\gamma(x))$, see (2.1).

Let $y \to \underline{A}(y)$, $y \in [0,1]$, $\underline{A}'(y) \stackrel{def}{=} \frac{d\underline{A}}{dy} \neq \underline{0}$ be such that the tori $\mathcal{T}(\underline{A}(y))$ have fixed energy. Then (evaluating the energy at the homoclinic point $\underline{\alpha} = \underline{0}, \varphi = \pi$) one sees that $\underline{\omega} \cdot \underline{A}(y)$ is constant so that the line $y \to \underline{A}(y)$ is parallel to $\underline{\omega}^{\perp} = (\omega_2, -\omega_1)$.

Define $y \to \underline{A}(y)$, $y \in [0,1]$, to be a *elastic heteroclinic chain* with flexibility parameters $\beta, \vartheta > 0$ and splitting μ if:
(i) for all $|y-y'| < \vartheta\mu$ there is a heteroclinic intersection between the stable and unstable manifolds of $\mathcal{T}(\underline{A}(y))$ and $\mathcal{T}(\underline{A}(y'))$ with splitting angles $\geq \mu$ at $\varphi = \pi$.
(ii) the intersection matrix $D \stackrel{def}{=} \mu D_o$ at $\varphi = \pi$, $\underline{\alpha} = \underline{0}$ verifies:

$$(\underline{w}^{\perp} \cdot D_o^{-1}\underline{w}^{\perp}) \stackrel{def}{=} \beta \neq 0, \qquad \underline{w}^{\perp} \stackrel{def}{=} \frac{\underline{\omega}^{\perp}}{|\underline{\omega}|} \tag{6.1}$$

where D, D_o are y–independent (because of isochrony).
(iii) a heteroclinic intersection at $\varphi = \pi$ between $W^s(\underline{A}(y))$ and $W^u(\underline{A}(y+\delta))$ takes place at $\underline{\alpha}_y(\delta) = D_o^{-1}\underline{w}^{\perp}|\underline{A}'|\vartheta' + O(\vartheta'^2)$ for $\delta = \mu\vartheta'$, $|\vartheta'| \leq \vartheta$, $|\underline{A}'| = |\underline{A}'(y)|$ and:

$$\frac{1}{2}|\underline{A}'|\beta|\tilde{\vartheta}| < |(\underline{\alpha}_y(\delta') - \underline{\alpha}_y(\delta'')) \cdot \underline{w}^{\perp}| < 2|\underline{A}'|\beta|\tilde{\vartheta}| \tag{6.2}$$

for all $\delta' = \mu\vartheta', \delta'' = \mu\vartheta''$ and $|\vartheta'|, |\vartheta''| < \vartheta$ with $\tilde{\vartheta} = \vartheta' - \vartheta''$.

Remarks:
(a) thus every sequence $y_0, y_1, \ldots, y_{\mathcal{N}}$ with $|y_i - y_{i+1}| < \vartheta\mu$ is a heteroclinic chain in the sense of §3, and the theorem proved in §5 applies to it. A elastic heteroclinic chain with parameter ϑ is also elastic with parameter $\vartheta' < \vartheta$. Hence it is not restrictive to suppose that ϑ is as small as needed.
(b) condition (6.1) is a transversality property while (6.2) is just saying that ϑ is so small that the involved first order Taylor's expansions are "good" approximations (hence it is a weak condition and it follows from (ii) provided ϑ is small enough). The geometrical meaning of (6.1), (6.2) is that when y varies by δ (so that $\underline{A}(y)$ varies in R^2 *orthogonally* to $\underline{\omega}$ by $O(\delta)$), then the heteroclinic intersection $\underline{\alpha}_y(\delta)$ between $W^s(\underline{A}(y))$ and $W^u(\underline{A}(y+\delta))$ is away from $\underline{0}$ *in the direction orthogonal to* $\underline{\omega}$ by $O(\delta\mu^{-1})$ provided $\delta\mu^{-1} = \vartheta'$ is small enough.
The same remains true if one looks at the displacement of the heteroclinic intersection at any other section located away from the tori by a fixed distance $\kappa > 0$, if ε is small enough. In fact consider the intersection matrix $D(t)$ evaluated along the heteroclinic trajectory at a time t after the passage through $\varphi = \pi$. From the equations of motion its evolution is $D(t) = D - \int_0^t \partial_{\underline{\alpha}\varphi} f(\varphi(t), \underline{\omega}\tau)\partial_{\underline{\alpha}}\Delta(\tau)\,d\tau$ where $\Delta(t)$ denotes the splitting in the φ–coordinates and $\varphi(t)$ the heteroclinic evolution of φ: *i.e.* $D(t) = D + O(\varepsilon^2)$ (while $D = O(\varepsilon)$), see (5.5) in [GGM].
In particular if we look at the heteroclinic intersection point $\underline{\psi}_y(\delta)$ at $q = \kappa$, on the same heteroclinic trajectory, and compare it with the position of the homoclinic point $\underline{\psi}_y(0)$ of $\mathcal{T}(\underline{A}(y))$ at $q = \kappa$ then we can say that, for some constants $2b_1, 2b_0$ (the factor 2 is just convenient) it is $|\underline{w}^{\perp} \cdot (\underline{\psi}_y(\delta') - \underline{\psi}_y(\delta''))| \in [2b_1\tilde{\vartheta}, 2b_0\tilde{\vartheta}]$, with $\tilde{\vartheta} = (\delta' - \delta'')\mu^{-1}$. The constants b_0, b_1 depend on κ, which is however prefixed, and on β.
(c) examples exist; and generically $\mu = O(\varepsilon)$.
(d) the above definition is a special case of a natural more general definition relevant for higher dimensions and for anisochronous systems. For instance in the case

of anisochronous systems in which a term $\underline{A}^2/2J$, with $J>0$ constant, is added to (1.1) one has to require that $y\to \underline{A}(y)$ is a simple rectifiable curve and that, uniformly in $y\in[0,1]$, (6.1) holds with D replaced by the intersection matrix D_y, and $\underline{\omega}$ replaced by $\underline{\omega}(\underline{A}(y))=\underline{\omega}+\underline{A}(y)J^{-1}$. In higher dimensions one replaces (6.1) by requiring that the determinant of the matrix $P_y^\perp D_y^{-1}P_y^\perp$, with $P_y^\perp$ the projection on the plane orthogonal to $\underline{\omega}(\underline{A}(y))$, be non zero uniformly in y. But in the anisochronous cases the condition that for all y there is the torus $\mathcal{T}(\underline{A}(y))$, called "no gap condition", is strongly restrictive and quite artificial (although it is verified in the example in [A]).

(e) One can free completely what follows from the parameter ε and discuss everything in terms of μ only. For simplicity we leave the general formulation to the reader and stay with ε small enough. Also for simplicity we shall take $|\underline{A}'|=1$, as this is not restrictive.

Theorem 2: *Suppose that $y\to\underline{A}(y)$ is elastic in the above sense, then fixed a,b there exist heteroclinic chains $\underline{A}_0=\underline{A}(y_0),\underline{A}_1=\underline{A}(y_1),\ldots,\underline{A}_{\mathcal{N}}=\underline{A}(y_{\mathcal{N}})$ with $y_0=a,y_{\mathcal{N}}=b$ along which the drift time is $O(\mu^{-2})$.*

The estimates proceed by performing the construction of §5 *without* fixing *a priori* the heteroclinic chain: we construct it inductively.

Using the notations of §4 assume that y_j have been constructed for $j\le i+1$ together with $\tilde{\underline{\psi}}_j,r_j,\rho_j,B_j,B_j'$ for $j\le i$. We must define $y_{i+2},\tilde{\underline{\psi}}_{i+1},r_{i+1},\rho_{i+1}$ and, as a consequence, B_{i+1},B_{i+1}'.

Let E be as in §5 and let E' be so large that if $T^0=\overline{g}^{-1}\log E'E^{-1}$ the points $\underline{\omega}t$, $t\in[0,T^0]$, fill the torus within $\frac12 b_1\vartheta$ (see (b) above for the definition of b_1). This means that E' is very big: $E'=\exp O(C\vartheta^{-\tau-1})$ if one uses the estimate that the time needed to a quasi periodic rotation of the torus with vector $\underline{\omega}$, diophantine with constants C,τ, to fill within δ the whole torus T^2 is $O(C\delta^{-\tau-1})$ (for completeness see appendix A2).

Let $X_{i+1}(y)$ be heteroclinic between $\mathcal{T}(\underline{A}_{i+1})$ and $\mathcal{T}(\underline{A}(y))$ for $y\in[y_{i+1}+\frac12\mu\vartheta,y_{i+1}+\mu\vartheta]$. Let the local coordinates of $X_{i+1}(y)$ be $(\underline{A}_{i+1},\underline{\psi}_{i+1}(y),0,\kappa)$ (see §4 for the notations). Let, see (4.1):

$$Y_{i+1,y}(\underline{\psi})=(\underline{A}^s_{i+2,y}(\underline{\psi}),\underline{\psi},p^s_{i+2,y}(\underline{\psi}),\kappa)\tag{6.3}$$

be the equation of $W^s(\underline{A}(y))$ in the local coordinates around the torus $\mathcal{T}(\underline{A}_{i+1})$ near $X_{i+1}(y)$. We may suppose that $|\underline{A}^s_{i+2,y}(\underline{\psi})-\underline{A}_{i+1}|,|p^s_{i+2,y}(\underline{\psi})|<b_3\mu|\underline{\psi}-\underline{\psi}_{i+1}(y)|$, for some b_3 of $O(1)$ and we may suppose $b_3\ge1$, for simplicity.

Suppose r small: a first approximation to $\tilde{\underline{\psi}}_{i+1}$ will be a point $\underline{\psi}_{i+1,y,r}$ at distance r from $\underline{\psi}_{i+1}(y)$ such that:

$$\lambda\stackrel{def}{=}|p^s_{i+2,y}(\underline{\psi}_{i+1,y,r})|=\frac12\max_{|\underline{\psi}-\underline{\psi}_{i+1}(y)|=r}|p^s_{i+2,y}(\underline{\psi})|\tag{6.4}$$

and $\lambda\in[b_2\mu r,b_3\mu r]$, for some $b_2=O(1)>0$, by the assumption on the splitting. The constants b_2,b_3 are large or small with the "target" parameter κ, fixed once and for all, see beginning of §4). Let $d=b_2/b_3$. *Note that $\underline{\psi}_{i+1,y,r}$ is defined non constructively.*

As r varies in the range $d\frac{\rho_i}{4b_3E'}<r<\frac{\rho_i}{4b_3E}$ the point $\underline{\psi}_{i+1,y,r}$ varies and λ varies by a factor not smaller than E'/E by our definition of d. Hence the time $T(r)$ necessary in order that the backward evolution of the point $Y_{i+1,y}(\underline{\psi}_{i+1,y,r})=(\underline{A}^s_{i+2,y}(\underline{\psi}_{i+1,y,r}),\underline{\psi}_{i+1,y,r}$, $p^s_{i+2,y}(\underline{\psi}_{i+1,y,r}),\kappa)$ *interchanges* the last two coordinates will vary by an amount $\ge T^0=O(const\,\vartheta^{-\tau-1})$, see the first lines of this proof and (2.1).

This implies, by continuity, that there will be a value $r(y)$ such that the backward motion of duration $T(r(y))$ of $Y_{i+1,y}(\underline{\psi}_{i+2,y,r(y)})$ has $\underline{\psi}$–coordinate $\hat{\underline{\psi}}_{i+1,r(y),y}$ close within $\frac12 b_1\vartheta+\frac{\rho_i}{4b_3E}$ to the coordinate $\underline{\psi}'_i$ of the point $X_i'=(\underline{A}_{i+1},\underline{\psi}'_i,\kappa,0)$ (around which the already known set B_i' is constructed (see (4.3)); *i.e.* closer than $b_1\vartheta$ (as we suppose that $\frac{\rho_i}{4b_3E} < \frac{1}{2}b_1\vartheta$ assuming that $\rho_i \le \rho_1$ and ρ_1 is small enough, which will turn out to be not restrictive). We can even obtain that $\hat{\underline{\psi}}_{i+1,r(y),y}$ is on a chosen side of the line through $\underline{\psi}'_i$ parallel to $\underline{\omega}$.

Note that this is just a continuity statement: hence it is non constructive; furthermore other continuity statements will follow. In this sense the analysis is quite close in spirit to the variational approaches: nothing is really constructive.

We now vary $y \in [y_{i+1} + \frac{1}{2}\mu\vartheta, y_{i+1} + \mu\vartheta]$: the point $\underline{\psi}_{i+1}(y)$ varies in the direction orthogonal to $\underline{\omega}$ by $b_1\vartheta$ at least (see remark (b) and condition (6.1)).

This means that there is y^*, in the considered interval, for which $\hat{\underline{\psi}}_{i+1,r^*,y^*}$, having set $r^* = r(y^*)$, *is on the line parallel to* $\underline{\omega}$ *off* $\underline{\psi}'_i$ and within a distance $b_1\vartheta\mu$ of it.

Fixed y^*, r^*, let $\underline{\psi}$ rotate on the circle of radius r^* around $\underline{\psi}_{i+1}(y^*)$. By the definition of $p^s_{i+2,y^*}(\underline{\psi})$ its modulus λ will vary at least by a factor 2 with respect to its value λ^* at $\underline{\psi}_{i+1,y^*,r^*}$, see (6.4), (in fact it will vary between $2\lambda^*$ and 0, extremes included).

Hence by suitably adjusting $\underline{\psi}$ on the circle we can find a point $\tilde{\underline{\psi}}_{i+1}$ such that the point $Y_{i+1,y^*} = (\underline{A}^s_{i+2,y^*}(\tilde{\underline{\psi}}_{i+1}), \tilde{\underline{\psi}}_{i+1}, p^s_{i+2,y^*}(\tilde{\underline{\psi}}_{i+1}), \kappa)$ *evolving backwards in time* exchanges the p, q coordinates in a time $T_{i+1} \simeq T(r^*)$, and $\underline{\psi}_{i+1,y^*,r^*} - \underline{\omega}T_{i+1} \equiv \underline{\psi}'_i$ (below we shall worry about the difference $\underline{\psi}_{i+1,y^*,r^*} \ne \tilde{\underline{\psi}}_{i+1}$).

In fact in order to obtain this we only have to change the time $T(r^*)$ by an amount $O(b_1\vartheta)$, and this is achieved by varying λ off λ^* by a factor $e^{O(b_1\vartheta)} \in [\frac{1}{2}, 2]$, as it is not restrictive to take $b_1\vartheta$ small.

Since $\tilde{\underline{\psi}}_{i+1}$ differs from $\underline{\psi}_{i+1,y^*,r^*}$ by a small amount $2r^* \le \frac{\rho_i}{2b_3E}$ (at most, by construction), this means that the point $\tilde{\underline{\psi}}_{i+1} - \underline{\omega}T_{i+1}$ differs from $\underline{\psi}'_i$ by at most $\frac{\rho_i}{2E}$, recalling that $b_3 > 1$. Since $r^* < \frac{\rho_i}{4b_3E}$ the other coordinates verify $|p^s_{i+1,y^*}(\tilde{\underline{\psi}}_{i+1})| < b_3\mu\frac{\rho_i}{2b_3E}$ and $|\underline{A}^s_{i+2,y^*}(\tilde{\underline{\psi}}_{i+1}) - \underline{A}_{i+1}| < b_3\mu\frac{\rho_i}{2b_3E}$ (recall that $\underline{A}^s_{i+2,y^*}(\underline{\psi}_{i+1}(y^*)) \equiv \underline{A}_{i+1}$). Hence the point Y_{i+1,y^*} evolves, backward, in time T_{i+1} to a point well inside B'_i: so do, in a time which differs suitably from T_{i+1} by a factor of $O(1)$, all the points close enough to it, say within $\rho_{i+1} = r^*/2$.

Hence if we set $r_{i+1} = r^* \ge d\frac{\rho_i}{4b_3E'}$ and $\rho_{i+1} = r_{i+1}\frac{\rho_i}{4^2E}\frac{\overline{g}}{|\underline{\omega}|}$, $y_{i+2} = y^*$ and $\underline{A}_{i+2} = \underline{A}(y_{i+2})$ we see that all points of B_{i+1} are on the forward evolution of points in B'_i. The time needed for the passage through B_{i+1} of the points of B'_i, which visit it, is bounded proportionally to T_{i+1}. The radius ρ_i has to be chosen so small so that the ratio between the variation of the time to exchange p and q is small enough (*i.e.* $O(\frac{\rho_i}{E})$).

It follows that within a time $T = const\sum_{i=0}^{\mathcal{N}} T_{i+1}$ the whole chain will be run by some trajectories. Here $\mathcal{N} \ge 2(b-a)(\vartheta\mu)^{-1}$ and $T_i \le O(\overline{g}^{-1}\log(E')^i)$, so that drift between $\underline{A}(a)$ and $\underline{A}(b)$ takes place in a time $O(\overline{g}^{-1}2^{\mathcal{N}})$:

$$T \le const\, C\,\overline{g}^{-1}\, 2^{\mu^{-1}\vartheta^{-1}} \tag{6.5}$$

and, recalling that ϑ is fixed, if $\mu = O(\varepsilon)$ this is $const\, 2^{const\varepsilon^{-1}}$.

§7. Concluding remarks. Very fast diffusion?

For a review on diffusion see [L]: in this paper the possibility of estimates of size of an inverse power of ε is proposed and discussed.

(1) The above nonvariational proof gives results not directly comparable to the best known, [Be], [Br], based on a variational method and giving (in [Br]) a polynomial drift time of $O(\mu^{-2})$.

The papers [Be],[Br], deal with Arnold's example, [A], *i.e.* with a different case. They use in an essential way the structure of the model, implying existence of a "gapless" system of local coordinates in which the motion is "trivial" (*i.e.* given by (2.1)). Although such coordinate system does not appear explicitly in the proofs in [Br], it nevertheless exists under Arnold's assumptions as shown by [P].

Therefore this difference between the present paper and [Br] is not so important: the above proofs apply also to the model in [A] and [Br] (in fact in absence of gaps also anisochronous systems admit coordinates $(\underline{A}, \underline{\psi}, p, q)$ with the properties of the ones in §2). Only the constants may be affected (in particular the ones in (6.5)), although I do not think that (6.5) changes. It is hard to see how to improve the bounds of §6, which are already quite non constructive. Hence the difference between the size of the bounds remains a puzzle that I do not understand.

(2) It is worth stressing that the above methods apply every time there is a heteroclinic chain and "no gaps" around resonant tori: therefore they apply to the case in [A] with, in the notations of [A], $\mu = \varepsilon^c$ and c large enough.

In the isochronous models they apply, immediately, to a variety of cases: a non trivial one is the hamiltonian (1.1) with $\underline{\omega} = (\eta^a, \eta^{-1/2})$, $a \geq 0$, $\varepsilon = \mu\eta^c$ with c large enough and, *possibly, even a further "monochromatic, strong and rapid"* perturbation $\beta f_0(\varphi, \lambda)$ like $\beta \cos(\lambda + \varphi)$ with $\beta = O(1)$. Suppose that we consider only values of η such that $|\underline{\omega} \cdot \underline{\nu}| > C\eta^d |\underline{\nu}|^{-\tau}$ for all $\underline{0} \neq \underline{\nu} \in Z^2$, see §2 in [GGM]. Then *by using the results of* [GGM] (§8) we see that if η is fixed small enough the homoclinic splitting is analytic in β for $|\beta| < O(\eta^{-1/2})$, while it *does not vanish* for β small (*i.e.* $\beta = O(\eta^c)$), generically in f (but it is very small, see [GGM], §6). Hence it is not 0 for all $\beta < 2$ (say) *except possibly finitely many values of β*. This means that in such *strongly perturbed systems* ($\beta = O(1)$) one still has elastic heteroclinic chains of arbitrary length, see §8 of [GGM], and therefore there is diffusion (provable by the methods of §5,§6). Furthermore the $\underline{A}$–independent (because of isochrony, see [GGM]) homoclinic angles can become large when β, μ approach their convergence radii and this gives us the possibility of "very fast" drift on time scales of $\sim O(1)$. In fact I think that the homoclinic splitting might be a monotonic function of ε, β for interesting classes of perturbations.

A similar analysis can be made for the model in [A] (which is also without gaps).

(3) An advantage of the technique of §5 is its flexibility which makes it immediately applicable, essentially without change, to anisochronous systems, see [CG].

(4) Constructivity, even partial (see comments in §5), seems the key to understanding the huge difference between the results of §5 and the variational results, or those of §6 above: diffusion time bounds in an inverse power of ε (in [Br] and §6) versus an exponential in the more constructive proposal in §5. A hint in this direction is provided by the bound in §6: by adding a new idea to the method of §5, *i.e.* of [CG], one can get a drift time estimate of $O(\varepsilon^{-2})$ instead of the exponential of [CG], and §5. *But the theory becomes now less constructive*: not even the sequence of close encounters with invariant tori is determined.

(5) It seems possible that the construction of §6 might be rendered constructive without losing the bounds: this is certainly an interesting problem.

(6) Finally we discussed only drift in phase space: but it is clear that heteroclinic chains do not need to "advance" at each step (*e.g.* a A–coordinate needs not to increase systematically): we can use heteroclinic chains that advance and back up at our prefixed wish (*e.g.* randomly). Hence, in this sense, there is no difference between drift and diffusion.

Acknowledgements: I am indebted to P. Lochak stimulating comments and, in particular, to G. Gentile and V. Mastropietro for many discussions and help in revising the manuscript. This work is part of the research program of the European Network on:

"Stability and Universality in Classical Mechanics", # ERBCHRXCT940460.

Appendix A1. Jacobi's map.

This appendix is standard: here it is taken from A9 of [CG] with small changes, to use it for future references.

The theory of jacobian elliptic functions shows how to perform a complete calculation of the functions, below denoted R, S, in terms of which the canonical Jacobi's coordinates are defined, see [GR] (9.198),(9.153), (9.146), (9.128), (9.197). The result, reported for completeness, is discussed in terms of the pendulum energy:

$$\frac{\dot\varphi^2}{2} + g^2(1-\cos\varphi) = E \tag{A1.1}$$

where the origin in φ is set at the stable equilibrium, to adhere to the notations in the theory of elliptic functions. Setting $u = t(E/2)^{1/2} \equiv \varepsilon^{1/2} gt$, $k^2 = 2g^2/E = \varepsilon^{-1}$ where ε is the *dimensionless* energy so that $\varepsilon = 1$ is the separatrix, let:

$$\mathbf{K}(k) = \int_0^{\pi/2} \frac{d\alpha}{(1-k^2\sin^2\alpha)^{1/2}} \tag{A1.2}$$

We shall use the "standard" notations (*i.e.* those in [GR]) for the jacobian elliptic integrals *except* for $x(.)$, which is usually denoted $q(.)$, but which we would confuse with the variable q that we want to construct:

$$\begin{aligned} k' =&(1-k^2)^{1/2}, \qquad g_J = g\frac{\pi}{2k\mathbf{K}(k')} = \varepsilon^{1/2} g, \qquad \lambda \equiv \frac{1}{2}\frac{1-k^{1/2}}{1+k^{1/2}} \\ x(k') =& e^{-\pi\mathbf{K}(k)/\mathbf{K}(k')} = \lambda + 2\lambda^5 + 15\lambda^9 + 150\lambda^{13} + 1707\lambda^{17} + \ldots \end{aligned} \tag{A1.3}$$

In terms of the above notations we have, directly from the definitions (*i.e.* from the equations of motion):

$$I(t) = \dot\varphi = -2g\varepsilon^{1/2}\,\mathrm{dn}\,(u,k), \qquad \varphi(t) = 2\,\mathrm{am}\,(tg\varepsilon^{1/2}) \tag{A1.4}$$

which yield, changing the origin for φ to the unstable point to conform with our notations (*i.e.* obtaining $\varphi(t) = 2(\,\mathrm{am}\,(tg\varepsilon^{1/2}) + \pi/2))$, for $I(t) = R(p(t),q(t)), \varphi(t) = S(p(t),q(t))$:

$$R = -2g\varepsilon^{1/2}\frac{\mathrm{dn}\,(iu,k')}{\mathrm{cn}\,(iu,k')}, \quad \sin\frac{S}{2} = \frac{1}{\mathrm{cn}\,(iu,k')}, \quad \cos\frac{S}{2} = i\frac{\mathrm{sn}\,(iu,k')}{\mathrm{cn}\,(iu,k')} \tag{A1.5}$$

Setting $p = e^{-g_J t}, q = x(k')e^{g_J t}$, see [GR], and using $R(p,q) = g_J(-p\partial_p + q\partial_q)S(p,q)$ to evaluate S from R, the quoted basic relations between elliptic integrals imply immediately that the $I(t) = R(p(t),q(t)), \varphi(t) = S(p(t),q(t))$, solve the pendulum equations if:

$$\begin{aligned} R(p,q) =& -2g_J\Big[\tfrac{p}{1+p^2} + \tfrac{q}{1+q^2} - \textstyle\sum_{n=1}^\infty (-1)^n \tfrac{1+x^{2n-1}}{1-x^{2n-1}}(p^{2n-1}+q^{2n-1})\Big] \\ S(p,q) =& 2\Big[\,\mathrm{arctg}\,p - \,\mathrm{arctg}\,q - \textstyle\sum_{n=1}^\infty (-1)^n \tfrac{1+x^{2n-1}}{1-x^{2n-1}}\tfrac{(p^{2n-1}-q^{2n-1})}{2n-1}\Big] \\ \sin\tfrac{S(p,q)}{2} =& \tfrac{g_J}{g}\Big[\tfrac{p}{1+p^2} - \tfrac{q}{1+q^2} - \textstyle\sum_{n=1}^\infty (-1)^n \tfrac{1-x^{2n-1}}{1+x^{2n-1}}(p^{2n-1}-q^{2n-1})\Big] \\ \cos\tfrac{S(p,q)}{2} =& -\tfrac{g_J}{2g}\Big[\tfrac{1-p^2}{1+p^2} + \tfrac{1-q^2}{1+q^2} + 2\textstyle\sum_{n=1}^\infty (-1)^n \tfrac{1-x^{2n}}{1+x^{2n}}(p^{2n}+q^{2n})\Big] \end{aligned} \tag{A1.6}$$

with $x \equiv pq$. Note that g_J depends on x, and so do k', k: hence the coefficients of the first and of the last two of (A1.6) are also functions of $x = pq$. Furthermore the (dimensionless) energy ε becomes a function of $\xi = pq$ defined by inverting the map:

$$\xi = x(k') \equiv x((1-\varepsilon^{-1})^{1/2}) \tag{A1.7}$$

and the point corresponding to $\varphi=\pi$ and to a dimensionless energy ε, has coordinates:

$$p\equiv 1,\quad q\equiv x(k') \tag{A1.8}$$

(a rearrangement of (A1.6) showing convergence for $p=1$ and $|x|<1$ is exhibited below).

The variables (p,q) defined above are nice and natural: however they are not canonically conjugated to (I,φ): the jacobian determinant of the map $(p,q)\longleftrightarrow(I,\varphi)$ is not 1. But the jacobian determinant must be a function $D(x)=\frac{\partial(p,q)}{\partial(I,\varphi)}$ of x alone (*i.e.* of the product pq); then (A1.8) and the equations of motion imply that $D(x)^{-1}=g_J^{-1}\frac{2g^2\,d\varepsilon(x)}{dx}=4g\frac{d\varepsilon^{1/2}}{dx}$.

Therefore one can modify the variables p,q into new variables $(p_J,q_J)=(pF(x),qF(x))$ with F such that the jacobian $\frac{\partial(p_J,q_J)}{\partial(I,\varphi)}=\frac{\partial(p_J,q_J)}{\partial(p,q)}D(x)$, which is $D(x)\cdot\partial_x(xF^2(x))$, is identically 1. One finds: $F(x)=(4g)^{1/2}(\frac{\varepsilon^{1/2}-1}{x})^{1/2}$.

To invert the map $(p_J,q_J)=(pF(x),qF(x))$ define $x_J\stackrel{def}{=}p_Jq_J$ and $G(x_J)\stackrel{def}{=}F(x)^{-1}$ then: $p=p_JG(x_J)$ and $q=q_JG(x_J)$, $x=x_JG^2(x_J)$. The final result is a local canonical map between Jacobi's coordinates (p_J,q_J) and global (I,φ) coordinates:

$$I=R(p_JG(x_J),q_JG(x_J)),\qquad \varphi=S(p_JG(x_J),q_JG(x_J)) \tag{A1.9}$$

where R,S are defined above, see (A1.6) which are written in a form easily recognized in the elliptic functions tables. The functions R,S can be rewritten in the following form:

$$\begin{aligned}
R(p,q)=&-4g\left[\sum_{m=0}^{\infty}\left(\frac{x^mp}{1+x^{2m}p^2}+\frac{x^mq}{1+x^{2m}q^2}\right)\right]\\
S(p,q)=&4\left[\sum_{m=0}^{\infty}\left(\operatorname{arctg}x^mp-\operatorname{arctg}x^mq\right)\right]\\
\sin\frac{S(p,q)}{2}=&\frac{2g_J(x)}{g}\left[\sum_{m=0}^{\infty}(-1)^m\left(\frac{x^mp}{1+x^{2m}p^2}-\frac{x^mq}{1+x^{2m}q^2}\right)\right]\\
\cos\frac{S(p,q)}{2}=&\frac{g_J(x)}{2g}\left[1-2\sum_{m=0}^{\infty}(-1)^m\left(\frac{x^{2m}p^2}{1+x^{2m}p^2}+\frac{x^{2m}q^2}{1+x^{2m}q^2}\right)\right]
\end{aligned} \tag{A1.10}$$

exhibiting some of the properties of the Jacobi map in a better way.

One checks that in the (p_J,q_J) variables the pendulum hamiltonian, (A1.1) has become a function $J(p_Jq_J)=2g^2+gx_J+O(x_J^2)$. The domain of definition of the map is given by the properties of the elliptic functions or, more restrictively, by the domain of convergence of the above series. It inludes a disk of some radius $\rho_J>0$ around the origin.

The important symmetry $R(p,q)=R(q,p)$ and $S(p,q)=-S(q,p)$ is manifest.

Appendix A2. Filling times of quasi periodic motions.

Let $(\omega_1,\ldots,\omega_d)=\underline{\omega}\in R^d$ be such that $|\underline{\omega}\cdot\underline{\nu}|^{-1}\le C|\underline{\nu}|^\tau$. Let $\chi(x),\chi_\perp(x)$ be C^∞–functions even and strictly positive for $|x|<\frac{1}{2}\pi$, vanishing elsewhere and with integral 1. Let $\underline{\psi},\underline{\psi}_0\in T^d$ and $x(\underline{\psi})\stackrel{def}{=}\varepsilon^{-(d-1)}\chi(\underline{\omega}\cdot(\underline{\psi}-\underline{\psi}_0)/|\underline{\omega}|)\cdot\chi(\varepsilon^{-1}|P^\perp(\underline{\psi}-\underline{\psi}_0)|)$, $P^\perp=$ orthogonal projection on the plane orthogonal to $\underline{\omega}$.

The function x can be naturally regarded as defined and periodic on T^d: if $\hat\chi(\sigma)$ is the Fourier transform of χ as a function on R then the Fourier transform of x is $\hat\chi(\nu^\parallel)\hat\chi(\varepsilon|\underline{\nu}^\perp|)$, $\underline{\nu}$ integer components vector, $\nu^\parallel=\underline{\omega}\cdot\underline{\nu}/|\underline{\omega}|$, $\underline{\nu}^\perp=P^\perp\underline{\nu}$. The average $\overline{T}^{-1}\int_0^{\overline{T}}x(\underline{\omega}t)\,dt$ is:

$$X=1+\sum_{\underline{\nu}\neq\underline{0}}\hat x(\underline{\nu})e^{-i\underline{\psi}_0\cdot\underline{\nu}}\frac{1}{\overline{T}}\frac{e^{i\underline{\omega}\cdot\underline{\nu}\overline{T}}-1}{i\underline{\omega}\cdot\underline{\nu}}\ge 1-\frac{2C}{\overline{T}}\sum_{\underline{\nu}\neq\underline{0}}\left|\hat\chi(\nu^\parallel)\hat\chi(\varepsilon|\underline{\nu}^\perp|)\right||\underline{\nu}|^\tau \tag{A2.1}$$

Since the last sum is bounded above by $b\varepsilon^{-(\tau+d-1)}$ the average X is positive for all $\underline{\psi}_0$ if $\overline{T}>2bC\varepsilon^{-(\tau+d-1)}$. This means that for $T>2bC\varepsilon^{-(\tau+d-1)}+\pi/|\underline{\omega}|$, hence for

$T > BC\varepsilon^{-(\tau+d-1)}$ with B a suitable constant depending only on d, the torus will have been filled by the trajectory of any point within a distance ε. This proof is taken from (5), p. 111, of [G1], see [BGL] for an alternative proof.

References

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

[Be] Bessi, U.: it An approach to Arnold's diffusion through the Calculus of Variations, Nonlinear Analysis, 1995.

[Br] Bernard, P.: *Perturbation d'un hamiltonien partiellement hyperbolique*, C.R. Academie des Sciences de Paris, **323**, I, 189–194, 1996.

[BGL] Bourgain, J., Golze, ., Lochak, P.: , preprint, in print in J. Statistical Physics.

[CG] Chierchia, L., Gallavotti, G.: *Drift and diffusion in phase space*, Annales de l' Institut Poincarè, B, **60**, 1–144, 1994.

[E] Eliasson, L.H.: *Absolutely convergent series expansions for quasi-periodic motions*, Mathematical Physics Elctronic Journal, MPJE, **2**, 1996.

[G1] Gallavotti, G.: The elements of mechanics, Springer, 1983.

[G2] Gallavotti, G.: *Twistless KAM tori*, Communications in Mathematical Physics **164**, 145–156, (1994).

[Ge] Gentile, G.: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable systems*, Forum Mathematicum, **7**, 709–753, 1995. See also: *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, Dynamics and Stability of Systems, **10**, 269–308, 1995.

[GG] G. Gallavotti, G. Gentile: Majorant series convergence for twistless KAM tori, *Ergodic theory and dynamical systems* **15**, 857–869, (1995).

[GGM] G. Gallavotti, G. Gentile, V. Mastropietro: *Pendulum: separatrix splitting*, preprint, chao-dyn@xyz. lanl. gov 9709004. And G. Gallavotti, G. Gentile, V. Mastropietro: *Arnold's diffusion is anisochronous systems*, in preparation.

[GR] Gradshteyn, I.S., Ryzhik, I.M.: *Table of integrals series and products*, Academic Press, 1965.

[L] Lochak, P.: *Arnold's diffusion: a compendium of remarks and questions*, Proceedings of 3DHAM, s'Agaro, 1995, in print.

[P] Perfetti, P.: *Fixed point theorems in the Arnol'd model about instability of the action–variables in phase space*, mp_arc@math.utexas.edu, #97-478, 1997, in print in Discrete and continuous dynamical systems.

[T] Thirring, W.: *Course in Mathematical Physics*, vol. 1, p. 133, Springer, Wien, 1983.

Internet: Author's preprints downloadable (latest version) at:

`http://chimera.roma1.infn.it`

`http://www.math.rutgers.edu/~giovanni`

Mathematical Physics Preprints (mirror) pages.

e-mail: giovanni@ipparco.roma1.infn.it

Archived also in:
mp_arc@math.utexas.edu # 97-??? and
chao-dyn@xyz.lanl.gov # 9709011