EFFECTIVE SPEED OF ARNOLD'S DIFFUSION AND SMALL DENOMINATORS

Pierre LOCHAK
CMA-ENS, 45, rue d'Ulm, F-75005 Paris, France

Received 28 April 1989; revised manuscript received 26 June 1989; accepted for publication 30 October 1989
Communicated by A.P. Fordy

We illustrate the influence of a certain type of resonances on the speed of Arnold's diffusion for Hamiltonian systems; this makes it possible to derive plausible estimates, in view of the upper bound provided by Nekhoroshev's theorem. We also emphasize the inadequacy of the variational (Poincaré–Melnikov) method.

We are interested in the perturbation of an integrable Hamiltonian flow and the ensuing global instability known as "Arnold's diffusion". We shall present here an example which appears as a generalization of Arnold's original example [1] and refer once and for all to ref. [2] for a more general and detailed discussion. In short, the phenomenon which is illustrated below consists in the resonance of high order harmonics of the perturbation with the natural frequency of the motion on the hyperbolic tori ("whiskered tori"). Let us consider the model case of the Hamiltonian

$$H=\tfrac{1}{2}p^2+\tfrac{1}{2}I^2+\epsilon(\cos q-1)[1+\mu F(\varphi)]\,, \tag{1}$$

where

$$(p,q)\in\mathbb{R}\times\mathrm{T}\,,\quad (I,\varphi)=\mathbb{R}^n\times\mathrm{T}^n,\quad \mathrm{T}=\mathbb{R}/\mathbb{Z}\,.$$

The number of dimensions is $N=n+1$, ε and μ are perturbation parameters and F is an analytic real function with an analyticity strip of width $\sigma>0$. Here we take

$$F(\varphi)=\sum_{k\in\mathbb{Z}^n-\{0\}} f_k\cos(k\cdot\varphi)\,, \tag{2}$$

with $k\in\mathbb{Z}^n$; $k\cdot\phi=\sum k_i\varphi_i$ is the usual scalar product, and the Fourier coefficients satisfy the estimate $|f_k|=\mathrm{O}(\mathrm{e}^{-\sigma|k|})$ where $|k|=\sum|k_i|$. In view of possible computer experiments, one can keep in mind the following examples: $f_k=\mathrm{e}^{-\sigma|k|}$ in which case the series can be summed explicitly, and $f_k=\exp(-\sigma\|k\|^2)$ where $\|\ \|$ denotes the Euclidean norm; one is then dealing – almost – with a θ function. We write

$$H=H_1+H_2+\varepsilon\mu\Phi(q,\varphi)\,, \tag{3}$$

with

$$H_1=\tfrac{1}{2}p^2+\epsilon(\cos q-1)\,,\quad H_2=\tfrac{1}{2}I^2=\sum_j\tfrac{1}{2}I_j^2\,.$$

When $\mu=0$, $H_0=H_1+H_2$ has invariant tori of dimension $n=N-1$ defined by $p=q=0$, $I=\omega$, $\varphi\in\mathrm{T}^n$ where we write $\omega=I$ only to underline that this is a frequency ($\omega=\nabla H_2$). With respect to Arnold's example, the generalisation consists in the fact that the dimension is arbitrary (so that one can consider an autonomous Hamiltonian), but much more crucially in that the perturbation term can include arbitrarily high harmonics. As in Arnold's case, hyperbolicity is absent when $\epsilon=0$ and we have introduced *two* parameters to avoid a very difficult singular perturbation problem (see briefly below and in ref. [2]); also, still as in Arnold's case, the perturbation vanishes on the tori, which are thus all conserved, a highly non-generic feature.

We shall now give in detail the computation of the Poincaré–Melnikov integrals for H, in which small denominators arise. When $\mu\geqslant 0$, the stable (+) and unstable (−) manifolds of a torus are defined by equations of the form

$$H_1=\Delta^{\pm}(q,p,I,\varphi)\,,$$

$$\tfrac{1}{2}I_j^2=\tfrac{1}{2}\omega_j^2+\Delta_j^{\pm}(q,p,I,\varphi)\,,\quad j=1,\dots,n\,, \tag{4}$$

where $\Delta^{\pm}$ and the $\Delta_j^{\pm}$ vanish for $\mu=0$. The classical

0375-9601/90/$ 03.50 © Elsevier Science Publishers B.V. (North-Holland)

Poincaré–Melnikov computation consists in evaluating – at least formally – these functions at first order in μ. More precisely, let $(p(t),\ q(t))$ be the solution of the pendulum equation described by H_1, corresponding to the separatrix (say its upper branch) and such that $q(0)=\pi$:

$$q(t)=4\ \mathrm{arctg}\ \mathrm{e}^{\tau}\ , \quad p(t)=\dot{q}(t)=\frac{2\sqrt{\epsilon}}{\mathrm{ch}\ \tau}\ , \quad \tau=\sqrt{\epsilon}\ t\ . \tag{5}$$

One has

$$\frac{\mathrm{d}}{\mathrm{d}t}\tfrac{1}{2}I_j^2=-I_j\frac{\partial H}{\partial\varphi_j}=\tfrac{1}{2}\mu\omega_j p^2\frac{\partial F}{\partial\varphi_j}\ , \tag{6}$$

since $I_j=\omega_j$ and $H_1=0$ on the separatrix. The linear approximation consists in substituting the unperturbed trajectory for the perturbed one in the integration. One computes the differences $\Delta=\Delta^+-\Delta^-$ and $\Delta_j=\Delta_j^+-\Delta_j^-$ in the plane $q=\pi$ which we denote as $\delta=\delta H_1$ and δ_j to this approximation. These quantities represent, to first order in μ, the distances of the projections of the stable and unstable manifolds in the planes $(p,\ q)$ and $(I_j,\ \varphi_j)$; they are functions of the initial angle φ_0 on the torus (see below), of its frequency ω and of the parameters which describe the perturbation. We thus obtain the version of the Poincaré–Melnikov formula for this example; in particular

$$\delta_j=\tfrac{1}{2}\mu\omega_j\int_{-\infty}^{\infty}p^2\frac{\partial F}{\partial\varphi_j}\,\mathrm{d}t\ . \tag{7}$$

For a harmonic $f_k\cos(k\cdot\varphi)$ of the perturbation, the contribution $\delta_j^{(k)}$ is given by

$$\delta_j^{(k)}=-\tfrac{1}{2}\mu f_k\omega_j k_j\int_{-\infty}^{\infty}p^2(t)\sin[k\cdot\varphi(t)]\,\mathrm{d}t\ , \tag{8}$$

where $\varphi(t)=\varphi_0+\omega t$ describes the unperturbed trajectory. One finds

$$\delta_j^{(k)}=-2\pi\mu f_k(\omega\cdot k)\frac{\sin(k\cdot\varphi_0)}{\mathrm{sh}(\frac{1}{2}\pi\omega\cdot k/\sqrt{\epsilon})}\omega_j k_j\ . \tag{9}$$

After a summation over j, we obtain the contribution of harmonic k to δH_2, denoted as $\delta^{(k)}H_2$:

$$\delta^{(k)}H_2=-2\pi\mu f_k(\omega\cdot k)^2\frac{\sin(k\cdot\varphi_0)}{\mathrm{sh}(\frac{1}{2}\pi\omega\cdot k/\sqrt{\epsilon})}\ . \tag{10}$$

One can compute in a similar way $\delta=\delta H_1$ and find $\delta H_1=-\delta H_2(=-\sum\delta_j)$. This result is not surprising, given the decomposition of H, since the latter is invariant and $\epsilon\mu\delta\Phi$ is negligible at first order in μ (this term is in fact of a still higher order because Φ oscillates).

The reasoning then goes very roughly as follows (see refs. [1], [2] or [3] for instance): in order to construct a "transition chain" between hyperbolic tori, one looks for *heteroclinic* intersections, here in the plane $q=\pi$, between the stable and unstable manifolds of two tori with frequency vectors $\omega^{(1)}$ and $\omega^{(2)}$ respectively. One must then solve the system

$$\Delta=\Delta H_1=0\ ,$$

$$\tfrac{1}{2}\omega_j^{(1)2}+\Delta_j^{(1)-}=\tfrac{1}{2}\omega_j^{(2)2}+\Delta_j^{(2)+}\ , \quad j=1,\ldots,n\ . \tag{11}$$

If the difference between the frequencies is small with respect to μ, the solvability of the system is equivalent, by the implicit function theorem, to that of the following linearized system, where δ and the δ_j are computed at a common value ω lying between $\omega^{(1)}$ and $\omega^{(2)}$:

$$\delta=0\ , \quad \delta_j=\tfrac{1}{2}\omega_j^{(2)2}-\tfrac{1}{2}\omega_j^{(1)2}\ , \quad j=1,\ldots,n\ . \tag{12}$$

Unfortunately, the problem is not so simple, and one will have to assume that, e.g. $\mu=\mathrm{O}(\exp(-c/\sqrt{\epsilon}))$ in order to ensure that the first order (with respect to μ) computation is really conclusive. All this applies equally well to Arnold's example, which, if one is willing to look closer, is far from being completely understood (see ref. [2]).

Pursuing Arnold's reasoning, the maximal step in a transition chain, i.e. $\|\omega^{(2)}-\omega^{(1)}\|$, is essentially determined, according to (12), by the size of the Poincaré–Melnikov integrals, and, *if* one can neglect the time necessary to achieve one step, with respect to the number of steps, then the "diffusion speed", i.e. the inverse of the time necessary to produce a drift of the action variables on the order of unity, will be approximately equal to that same number. In Arnold's example, one thus estimates the speed as $\exp(-c/\sqrt{\epsilon})$ (again a rigorous proof of this fact seems to be very difficult).

The situation for the Hamiltonian H is still more complicated... The new fact comes from the appearance, in formulas (9) and (10), of the classical expression $\omega\cdot k$, or rather the combination $(\omega\cdot k)/\sqrt{\epsilon}$, which produces an intrication of hyperbolicity (heteroclinic manifolds), singular perturbation (factor $1/\sqrt{\epsilon}$), and ellipticity ("small divisor" $\omega\cdot k$; see below the estimate (13)), this last ingredient being new with respect to Arnold's example (and of course generic).

Let us repeat that, because there is no hyperbolicity in the unperturbed system, one obtains, for ϵ and μ non-zero, a pendulum with frequency $\sqrt{\epsilon}$ perturbed by a term of frequency 1 (and amplitude μ), or equivalently, after rescaling time, a pendulum of frequency 1 with a fast perturbation (or frequency $1/\sqrt{\epsilon}$). We underline that this problem of nonlinear singular perturbation theory is far from being understood even in simpler instances (see ref. [4] and especially ref. [5]). To avoid it, Arnold uses *two* parameters: ϵ introduces hyperbolicity without destroying integrability and μ breaks the integrability but it has to be very small (in fact *exponentially* small with respect to ϵ). The Poincaré–Melnikov integral in ref. [1] looks like (10), with only $\mathrm{sh}(\pi\omega/2\sqrt{\epsilon})$ in the denominator, where ω is a scalar (and $k=1$); this is *always* exponentially small with respect to ϵ, which allows a prediction (at least formally). Here on the contrary, the factor $(\omega\cdot k)/\sqrt{\epsilon}$ can a priori assume any value (including zero), which indicates a possible resonance of the frequency vector ω with the relative frequency $k/\sqrt{\epsilon}$ of some harmonic of the perturbation.

Up to now, the only condition has been $\delta=\delta H_1=0$, which constrains the vector I (or ω) to move on a sphere ($\|I\|=\mathrm{const}$). Examining (9) with $\sin(k\cdot\varphi_0)=\pm 1$ because one must extremize with respect to the initial angle, it is easy to see that one could predict any speed, including one in contradiction with Nekhoroshev's theorem [6]: because of the small divisors, the Poincaré–Melnikov computation, which is linear in nature, loses all validity, even a formal one.

Now, *suppose* that Diophantine frequencies are privileged for some reason (see below): one can then perform the following estimate, which meets with some rather obscure indications in Chirikov's inexhaustible article [7], where they are formulated in probabilistic terms. Let $\omega(\gamma,\tau)$ be Diophantine, k such that

$$\sqrt{\epsilon}\ll|\omega\cdot k|\approx\gamma/|k|^{\tau}\ll 1\,,\quad \gamma>0\,,\quad \tau\geqslant n-1\,, \tag{13}$$

and $|f_k|$ is of the order of $\mathrm{e}^{-\sigma|k|}$. According to formula (9), this harmonic will contribute to the splitting of the separatrices by a term of order

$$\delta_j^k\sim|k|^{-\tau}\,\mathrm{e}^{-\sigma|k|}\exp(-c/\sqrt{\epsilon}\,|k|^{\tau})\,, \tag{14}$$

where we have not put μ in front; one would really like to set $\mu=\epsilon$, but, as has already been mentioned, this would lead to a singular problem. One may choose instead μ exponentially small with respect to $1/\sqrt{\epsilon}$, which only modifies the constant c (apart from the contribution of μ, one has $c=\pi\gamma/2$). Let us now extremize (maximize) this contribution with respect to $|k|$, ϵ being fixed. In this way, we shall find the maximal allowed step in a transition chain, and this will give us the speed of diffusion, according to Arnold's reasoning recalled above. One finds that one should take $|k|\sim\epsilon^{-1/2(\tau+1)}$ so that

$$\delta_j^k\sim\epsilon^{\tau/2(\tau+1)}\exp[-\sigma(1+1/\tau)\epsilon^{1/2(\tau+1)}]\,. \tag{15}$$

The most important feature in this formula is the exponent $1/2(\tau+1)$ of ϵ in the iterated exponential. In particular, when $\tau=n$, one finds that the speed of diffusion goes as $\exp(-c/\epsilon^{1/2n})$ which exhibits a dependence with respect to the number of degrees of freedom which is rather close to that given in Nekhoroshev's theorem (the optimality of the latter is still far from being understood; see ref. [2]).

If one wants now to justify the coming into play of arithmetic conditions, it is useful to remind oneself of the necessary steps in the proof of existence – and possibly the evaluation – of Arnold's diffusion; we shall distinguish three steps:

(i) Existence of hyperbolic tori and local invariant manifolds.

(ii) Semi-global control of the extension of these local manifolds and proof of the existence of heteroclinic intersections.

(iii) Construction of a transition chain and reduction of the problem to a version of the λ-lemma.

It seems that the small denominators could come into play at *any* of these levels.

(i) This is qualitatively rather well understood, at least in the nonsingular case, i.e. when hyperbolicity is present from the start, in the unperturbed Hamiltonian. One can then show that many hyperbolic

Diophantine tori persist after perturbation [8–10] by using a conceptually simple variant of the KAM techniques. The intrication of hyperbolicity and ellipticity is relatively clear at this level – although technically cumbersome. Of course, no realistic estimate for the threshold of breaking of the tori is available to date. The singular problem is more difficult [11] but one should be able to use versions of Kolmogorov's theorem adapted to degenerate Hamiltonians [12].

Several remarks may be in order: first of all, in the above example, as in Arnold's, this first part does not exist because the perturbation vanishes on the tori; second, we have privileged tori of dimension $N-1$ but all intermediate dimensions are possible and can typically be modellised using analogous examples consisting of the sum of m penduli and $n=N-m$ rotators; third, the usual Diophantine condition is rather artificial as a measure of irrationality and this is also difficult to remedy.

(ii) It is precisely the aim of this Letter to evidence the role of the small denominators at this level. This seems to invalidate the usual variational linear method, even as a formal means of prediction, and to call for a really nonlinear method.

(iii) To our knowledge, this step is rigorously treated only in ref. [3], but only qualitatively and for a very unrealistic example. One then only demands, as is the case in ref. [1], that the components of ω be rationally independent, i.e. that the resonance be effectively a simple one, or still that the flow on the $N-1$ torus be transitive. In the above example, it seems likely that small denominators will come into play also at this level. We only recall in this respect that, in the course of the computation, we assumed fulfilled the condition $\sin(k\cdot\varphi_0)=\pm 1$ (or more generally $|\sin(k\cdot\varphi_0)|>c>0$); but when $|k|$ increases, this "initial condition" on the torus becomes of course ever more unstable.

To conclude, we have tried to show that the speed of diffusion for the Hamiltonian H could well be of the same order as the estimate from above guaranteed by Nekhoroshev's result. A proof of this, however difficult it may appear, would be of fundamental importance, because it would make estimates from above and from below for the diffusion in Hamiltonian systems – roughly – touch each other.

The example presented above is generic, in the sense that the sum of a pendulum and a rotator with irrational frequency vector provides a "universal" model for a simple integrable resonance (modulo linear symplectic transformations); one then destroys integrability using a term with an arbitrary number of harmonics, while keeping the requirement of analyticity, which is necessary if one wants to compare the result with Nekhoroshev type estimates, which are of an essentially analytic nature.

On the other hand, the example is particular and simpler than the generic case because all tori are preserved and one looks at simple resonances only; as we already mentioned, it is not difficult to write similar examples for resonances of any order, in which case one should use, at the linear level, a vector Poincaré–Melnikov computation [13]. Lastly the introduction of *two* perturbation parameters is also artificial, just as in ref. [1] and its aim is only to avoid *singular* perturbation theory, as far as this is possible.

References

[1] V.I. Arnold, Sov. Mat. Dokl. 5 (1964) 581.
[2] P. Lochak, in preparation.
[3] R. Douady, Ann. Sc. ENS (1988).
[4] P. Holmes, J. Marsden and J. Scheurle, preprint (1987).
[5] V.F. Lazutkin, Vestn. Leningr. Univ. Fiz. (1984).
[6] N.N. Nekhoroshev, Russ. Math. Surv. 32 (1977) 1.
[7] B.V. Chirikov, Phys. Rep. 52 (1979) 263.
[8] J. Moser, Math. Ann. 169 (1967) 136.
[9] S.M. Graff, J. Diff. Eq. 15 (1974) 1.
[10] H. Rüssman, preprint (1988).
[11] R. de la Llave and C.E. Wayne, preprint (1988).
[12] A.I. Neistadt, J. Appl. Math. Mech. 45 (1981) 766.
[13] J. Gründler, SIAM J. Math. Anal. 16 (1985) 907.