Arnold diffusion; a compendium of remarks and questions

Pierre Lochak

The problem of the instability of multidimensional, close to integrable Hamiltonian systems is a notoriously difficult one; the underlying phenomenon now usually goes under the name "Arnold diffusion", paying tribute to the oldest and still most significant paper in the subject, namely the four page note by V.I.Arnold ([A]; we call attention to the short section 8 of [A2] in which Arnold recently briefly discussed what he sees as some important questions in the domain). The word "diffusion" is in fact misleading, and perhaps not without a tint of wishful thinking. In [A] and further purely mathematical papers, one studies the more modest question of the existence of a topological instability, without trying to describe it as a "diffusion", a characterization which may be just wrong in a restricted sense (thinking in terms of classical Markov processes), and certainly will require a lot more understanding of the phenomenon before it can can be vindicated.

Learning the subject is not an easy matter, not only because of its intrinsic difficulty, but also because it is not always clear how to make sense and assess the level of rigour of many assertions to be found in the – comparatively scant – literature. This stems perhaps primarily from the discrepancy between the heuristic understanding of the "physicists" and the rigorous, but comparatively very weak theorems of the "mathematicians". It goes without saying that developing an understanding of what is really happening, much beyond the border of what can be actually "proved", is of utmost importance, and in this direction, we refer first to the pioneering report of B.V.Chirikov ([C]), as well as to subsequent work by this and other authors (we briefly return to this point in section 7 below). It should perhaps be added that in this field, numerical experiments are particularly difficult to perform, and are thus not always reliable. Indeed many dimensional dynamical systems are not easy to deal with, and exponentially small quantities – as well as exponentially long times – are difficult to catch and get easily "drowned" in all sorts of "noises", starting with the unavoidable round-off errors.

In the present paper, of slightly unusual format, we have tried to gather remarks and questions pertaining to the purely mathematical side of the subject. The main idea is to gather some material which may help the newcomer enter the subject. Many of the "facts" presented below could be termed "well-known", if sometimes in very restricted (often Russian) circles; some assertions are more adventurous, possibly just wrong. In any case, the whole thing lies obviously wide open to the critics of the experts. We only hope that not too much in it will turn out to be evidently false, and would very much appreciate comments and corrections.

Below, we thus discuss a list of questions which may come to mind, possibly after reading part of the literature on the subject, and again sticking to the purely mathematical side of it. In order to keep this paper to a reasonable length, we skip many formal

definitions of the objects and instead assume a minimal acquaintance with the – by now classical – setting, as defined in the original note by Arnold, which of course we urge the reader to consult. The bibliography has been reduced to a bare minimum, but second generation references (those contained in the bibliographies of the papers we quote) should provide an already much better coverage.

The main character of the story is the general autonomous near integrable Hamiltonian system, with Hamiltonian

$$H(p,q) = h(p) + \epsilon f(p,q), \qquad (p,q) \in {}^n \times {}^n, \tag{1}$$

the study of which Poincaré once called "the fundamental problem of dynamics". Here $n > 2$ is the dimension of the "configuration space" (phase space has dimension $2n$); (p,q) denote the action-angle variables associated with the integrable system governed by h ($^n = (/)^n$ is the n torus); ϵ is a "small" parameter in front of the perturbative term $f(p,q)$. As a – not quite typical and certainly very hard – example, one can think of the planetary problem in celestial mechanics. The unperturbed part h is a sum of uncoupled two-body problems (interactions of the sun with each planet) and the perturbation arises from the mutual interactions between the planets, a typical value of the parameter ϵ being given by the ratio of the mass of the heaviest planet to that of the sun. In our solar system, ϵ turns out to be of order $.5\ 10^{-3}$ (Jupiter being the heaviest planet), unfortunately much to large for most rigorous results to literally apply.

Some basic restrictions (nonlinearity and smoothness) have to be added on the system (1). In any case, by writing the original system this way, we are essentially looking at a region of phase space where no singularities are present and energy surfaces are compact, and assume that there exists a smooth sytem of action-angle coordinates (p,q) for the unperturbed Hamiltonian h. Note that this excludes hyperbolicity in the unperturbed system, which may come in (still preserving integrability) when the foliation by the invariant tori ceases to be regular. Here we shall not need any complicated result on the global geometry of integrable systems and it is enough to keep in mind the familiar example of the ordinary gravitational pendulum, with its stable (elliptic) and unstable (hyperbolic) equilibria. We refer to any book on Hamiltonian dynamics for what is needed to make good sense of the sketch above.

Roughly speaking, there exist at present three types of results describing the behaviour of systems as in (1). The oldest part of the theory (dating back to Lagrange, if not Newton himself), which we can rightly call "classical perturbation theory" is concerned, for reasons which we shall not detail, with the stability of the action variables (the p variables) over long intervals of times. That is, given an initial condition $(p(0), q(0))$, one seeks a bound on the drift $||p(t) - p(0)||$ for $|t| \leq \mathcal{T}(\epsilon)$, with $\mathcal{T}(\epsilon)$ large w.r.t. $1/\epsilon$. This includes for instance the general problem of the stability of the solar system over very long – but finite – intervals of time, since the semiaxes of the planets appear among the action variables for the uncoupled Kepler problems. The recent crowning achievement of the theory is embodied in estimates over exponentially long times ($\mathcal{T}(\epsilon)$ exponential w.r.t. $1/\epsilon$), starting with the paper of N.N.Nekhoroshev ([N]), and it really seems that, for a large class of systems, the theory has now been pushed to its limits (see [L], [L1] and section 7 below).

A second, more recent, line of study of (1), which we can call "geometric perturbation theory", is concerned with the search of geometric objects invariant under the flow of (1). The existence of periodic orbits was the subject of a detailed study by Poincaré and of course KAM theory (starting with Kolmogorov 1954 note) proves the existence of "many" invariant tori with dimension $\leq n$ (for ϵ small enough). Other objects, like the invariant manifolds attached to lower dimensional, partially hyperbolic KAM tori, "Cantori", and others can also be detected. The trajectories sitting on these objects are naturally forever stable in some important cases (invariant tori) or their asymptotic behaviour is more less well understood (invariant manifolds attached to the tori). In particular, the trajectories living on KAM tori of dimension n (the only ones with nonzero Lebesgue measure) experience a drift in action $||p(t)-p(0)||$ which is bounded by the distortion of the tori, namely $O(\epsilon)$ (or in fact $O(\sqrt{\epsilon})$ if one wants a set of tori of asymptotically full measure), for all times.

The study of "Arnold diffusion" essentially covers everything not dealt with by classical and geometric perturbation theory. What is the fate of the trajectories living outside the objects discovered by geometric perturbation theory, in particular outside full dimensional KAM tori, over intervals of times not covered by classical perturbation theory? To be a little more precise, one may say that Arnold diffusion deals with *asymptotic global* instability that is, the global – in space and time – behaviour of the system as the perturbation parameter ϵ goes to zero. One basic question for instance is to find trajectories such that for *some* time t the drift $||p(t) - p(0)||$ is "of order 1", together with some uniformity condition as ϵ goes to 0. Note that such problems were calready discussed by Arnold in [A1], where precise conjecture are made. For the sake of clarity we make a distinction between "Arnold diffusion" which, as outlined above, simply designates the asymptotic global instability of near integrable systems, and "Arnold mechanism", which points to the particular – and probably by no means unique – instability mechanism first described by Arnold in [A].

At this point, as noted above, we allow for a gap in the exposition. From now on, we assume a basic acquaintance with the formal setting of the problem, as introduced by Arnold; all the necessary information can be gained from a careful reading of [A], which perhaps remains the best reference (and certainly by far the shortest).

We will discuss – if only briefly – the following points:
(1) Other settings
(2) Ellipticity versus partial hyperbolicity; singular and non singular perturbation theory
(3) Arnold mechanism: applying KAM theory
(4) Arnold mechanism: diffusion paths and transverse intersections
(5) Simple versus multiple resonance
(6) Arnold mechanism: transition chains
(7) Topology versus quantitative estimates
(8) Genericity, variational and other methods

So let's start with:
(1) Other settings: In the present paper, we shall work with *analytic Hamiltonian flows* (as in [A]) and the purpose of these short remarks is simply to remind the reader that dealing with other types of system does not really alter the main features of the problem.

First, one can of course work with discrete systems i.e. with symplectic maps instead of flows. Then, instead of a near integrable system, one can study the neighbourhood of an equilibrium point, both in the continuous and discrete cases (the distance to the fixed point features the small parameter). For example, [D] deals with the case of a symplectic map near an equilibrium point. There exist more or less standard ways of adapting results from one case to another (see [L] section IV.2 for some more details).

Suffice it to say here that the mutually inverse processes of taking a surface of section (from continuous to discrete) and building a suspension of the system (from discrete to continous) are now rather well understood in the analytic category (by far the hardest to handle as far as suspension is concerned), including from a quantitative viewpoint (see [KP]). Naturally there is always the option of rewriting proofs from scratch when changing the type of system to be studied.

As for smoothness assumptions, the analytic category is in some sense the most natural and most often encountered in practice. Quantitative estimates of the speed of diffusion (cf. section 6 below) *have* to assume analyticity; otherwise, diffusion – or say drift – can be much faster. This can be seen at the level of the elementary theory of averaging and has nothing to do with the symplectic character of the system (see e.g. [LM], Chapter 3). Much of what is done however, can be adapted to systems which are smooth enough (a large enough number of derivatives for the relevant objects), starting of course with KAM type results, and including heuristic reasonings for the speed of diffusion. Purely topological results, such as those considered in [D], are naturally less sensitive to smoothness requirements. Results may be less sensitive, but proofs are. Indeed, working e.g. in the C^∞ category allows for more flexibility: for instance in [D], integrability is broken by using a perturbation term with – small – *compact* support, so that the techniques of that paper do not carry over to the analytic case.

(2) Singular and non singular perturbation theory: This is a very important point; a natural starting point is to consider an analytic integrable Hamiltonian with compact energy surfaces such that the corresponding invariant tori smoothly foliate a domain of phase space. Arnold diffusion, or rather Arnold mechanism for instability, as described in [A], requires partial hyperbolicity, and none is present in the unperturbed system. This leads to a problem of *singular* perturbation theory, which is hard in several respects. In particular, it implies exponentially small splitting of the invariant manifolds (see section 4 below) and it is not even clear whether the mechanism in [A] is generic or not (see sections 3, 4 and 8 below). In [A], the way out was simply to artificially introduce a second parameter μ (in effect, Arnold follows Poincaré in *Méthodes Nouvelles*, including in the naming of the two small parameters). When ϵ is nonzero but μ *is* kept equal to zero, hyperbolicity is "turned on" *but* integrability is retained; μ is then "turned on", and by assuming μ much smaller (exponentially smaller) than ϵ, singular perturbation theory is in effect bypassed. Another essentially equivalent way to phrase this is to set $\epsilon = 1$ and keep only μ (renaming it ϵ...) which amounts to looking at a system such that the unperturbed part is already partially hyperbolic.

Locally, the prototype of such a system is as follows. Consider action variables $p \in {}^{n-m}$ with conjugate angles $q \in {}^{n-m}$; add hyperbolic variables $(x, y) \in {}^{m} \times {}^{m}$ and look at the

integrable Hamiltonian:

$$h_1(p,x,y) = h(p) + x \cdot \Lambda y, \tag{2}$$

where Λ is a diagonal nonsingular matrix and the dot denotes ordinary dot product. We shall essentially restrict ourselves to the perturbations of such systems, and for reasons to be made clear below (in section 5), we shall actually work mostly with $m = 1$, so that $\Lambda = \lambda$ is then simply a positive real number. We stress that this is a substantial simplification w.r.t. to the original question of investigating the stability of perturbations of smooth – fully elliptic – integrable Hamiltonian systems. In [CG], which deals with that kind of systems (with $m = 1$), the authors coined the name "a priori unstable" for such a simplified situation. The fact that hyperbolicity is present in the unperturbed system also completely changes the speed of diffusion which needs not a priori be exponentially slow (see section 7 below). More generally, as soon as partial hyperbolicity is assumed in the perturbed system ("a priori instability"), exponentially small quantities simply do not seem to enter the problem anymore (in an intrinsic way of course).

(3) Applying KAM theory: The first step in following Arnold's scheme in order to vindicate what we call "Arnold mechanism" for instability, is to find *enough* invariant tori in the perturbed system, by applying KAM theory. True, in the original paper, Arnold simply worked with a perturbation which vanishes on the invariant tori which are used in the diffusion path (see section 4 below), so that this step could be skipped altogether. This is however of course not generic, and at this stage, there seems in fact to arise a real problem about the genericity of the mechanism; this should be made clearer from what follows and is briefly discussed again in section 8. For the time being, we shall content ourselves with recalling that one first looks for partially hyperbolic invariant tori. For perturbations of Hamiltonian (2) (any m), the theory was neatly set up in [G], which also discusses the invariant manifolds ("whiskers"). To wit, if $\omega(p) = \nabla h$ is the frequency vector on the unperturbed $d = n - m$ tori, one finds that these are preserved, for small enough perturbations, at the values p such that $\omega(p)$ satisfies a diophantine condition. It should be noticed that this normally hyperbolic situation is significantly easier to deal with than the normally *elliptic* or indeed the mixed situation, i.e. trying to find lower dimension invariant tori with arbitrary spectra for the complementary dimensions. Fortunately, Arnold mechanism requires working in the partially hyperbolic situation only (although the theory could perhaps be generalized to some mixed types, probably at the cost of much additional technical difficulty). An additional nice feature of the hyperbolic situation is that the motion *on* the invariant manifolds remains integrable after perturbation that is, conjugate to a linear flow (see [G]; it is not clear whether this result requires analyticity). As discussed in item 2 above, perturbing Hamiltonian (2) represents a simplification of the problem, avoiding *singular* perturbation theory. One would rather like to keep the small parameter ϵ in front of the hyperbolic term $(x \cdot \Lambda y)$ and add a perturbation of the same size ϵ. This has been worked out in [T] and turns out not to be very different from the non singular case. In short, the non degenerate (i.e. non singular, i.e. hyperbolicity of order 1) and degenerate (i.e. singular, i.e. hyperbolicity of order ϵ) cases of the normally hyperbolic situation are nicely amenable to KAM theory and have both been dealt with neatly.

There remain the questions of finding points with diophantine frequencies, computing the invariant manifolds, their intersections etc. More on this below, and more serious restrictions ahead.

(4) Arnold mechanism: diffusion paths and transverse intersections: The next step in the construction is the determination of possible paths (or "channels" in more physical parlance) of diffusion; these are simply curves in action space, to be eventually followed along by some trajectories of the system (or rather their action coordinates; angles are simply ignored at this point). Looking back at Hamiltonian (2), one sees that considering the $d = n - m$ dimensional variables p amounts to restricting attention to an m-fold resonance of the original system. This expression is not really adequate however, and refers rather to the fully integrable elliptic situation. Namely, start from

$$H(I, \phi) = H_0(I) + \epsilon H_1(I, \phi), \tag{3}$$

set $\Omega(I) = \nabla H_0(I)$, and consider the locus in I space where $\Omega(I)$ is m degenerate; up to a linear symplectic change of coordinates, the model for this is $\Omega = (\omega, 0) \in {}^{n-m} \times {}^{m}$, i.e. a standard $d = n-m$ dimensional resonant plane. Then, along such a plane, prescribing the spectral type of H_1 and with some other provisos to be checked, one can reduce the system to a perturbation of Hamiltonian (2), but again with ϵ in front of the hyperbolic term. It can then be argued as in section 2 above that a good simple model consists in removing this ϵ, so as to get an initially hyperbolic system. So, for now we return to perturbations of (2), with any m, although the next item will explain why imposing $m = 1$ greatly simplifies things (no other case seems to have been treated at the time of this writing).

Consider the $d = n - m$ dimensional p space and curves drawn in it. "How many diophantine points does a curve in ^{d} contain?" We recall the usual diophantine conditions; if ω is a d vector, one requires that there exist $\tau \geq d - 1$ and $\gamma > 0$ such that

$$|\omega \cdot k| \geq \frac{\gamma}{|k|^{\tau}}\,, \text{ for all } k \in {}^{d} \setminus \{0\}. \tag{4}$$

Inasmuch as one assumes, as we do, full nonlinearity of the unperturbed system ($p \mapsto \omega(p)$ is a local diffeomorphism), questions in p or ω spaces are basically equivalent. The question above is then a classical one in geometric number theory; it is discussed in [P] for instance. It is easy to prove that if a curve is "essentially non planar" (has "full torsion"), it contains "many" diophantine points. We refer to [Py]; [CG] contains some simple explicit computations. But how many is "many", and does that suffice for the purpose of rigging up Arnold mechanism?

Here the rough estimates run as follows: for a perturbation of Hamiltonian (2), of size ϵ, one applies KAM at a point satisfying (3) (with $\omega = \omega(p) = \nabla h(p)$). Fix τ (e.g. $\tau = d$) and let γ vary; as γ goes to zero, condition (4) becomes of course less and less stringent. It turns out that the KAM algorithm converges for $\gamma = \gamma(\epsilon)$ of order ϵ^a for some a, $0 < a < 1$, so that we need only "weakly" diophantine points. The simplest geometric result is now roughly as follows; assume the curve $C \subset {}^{d}$ is "essentially non planar" (which means that the tangent vector and its first derivatives span the whole space) and pick a finite portion of it; select γ_m, a minimum value for γ. Then the set of points on C which do *not* satisfy (4) with $\gamma = \gamma_m$ has Lebesgue measure on the order of γ_m^b for

some b, $0 < b < 1$ (explicit estimates are available). The last piece of information is about the invariant manifolds of the partially hyperbolic invariant tori: since we are looking at the *non* singular case (i.e. perturbations of (2)), the splitting of the invariant manifolds is estimated via a *linearized* computation, to wit a multidimensional Poincaré-Melnikov method (see [Gr] and [W]). This is again an important point; in the singular case, we are back to *singular* perturbation theory and evaluating the splitting is a difficult problem which is in fact open in the multidimensional case (the one dimensional – or should one say *two* dimensional ? – case has recently been the object of much study and is now fairly well-understood). In any case the splitting in the singular case (corresponding to the original Hamiltonian (1)) would be exponentially small w.r.t. ϵ (see section 7 below), and that would change things drastically. So we stick to the non singular ("a priori unstable") case (perturbations of (2)). Then the splitting, evaluated by means of a matrix Melnikov integral as noted above, will generically be of order ϵ. Now we need transverse heteroclinic intersections of the invariant manifolds of neighbouring tori, corresponding to points on the curve C. The upshot is that we a priori need these tori to be less than $O(\epsilon)$ apart (so that the splitting ensures transverse intersection). And this is again a priori rather bad news... In fact, from the above, we can fix C (a finite piece of it) and prescribe $\gamma_m \sim \epsilon^a$, so that KAM applies at (τ, γ_m) diophantine points. Now the complement on C of the (τ, γ_m) diophantine points will have relative measure bounded above by $O[(\gamma_m)^b] \sim O(\epsilon^{ab})$, so that any interval of length exceeding $c\epsilon^{ab}$ (c some constant) will indeed contain a "good", i.e. (τ, γ_m) diophantine, point. But... this is not good enough. Indeed, $ab < 1$ and we needed points on C wich are $O(\epsilon)$ apart, so as to ensure heteroclinic transverse intersections. So we are stuck again; note that all this did not happen in [A], where there was no need to apply KAM and no gaps between tori – they all "survived" the onset of the perturbation. The problem does not seem to be easy to solve or get around, but again, there is a way out, as indicated in [CG], and again at the expense of more restrictions. Imagine one could replace ϵ with $\epsilon^{1/c}$ $(0 < c < 1)$ in the above, but keeping the splitting of the same order, namely ϵ. Then of course, for $c < ab$, the obstacle would be removed. But replacing ϵ with $\epsilon^{1/c}$ is precisely what is effected by performing $[1/c] + 1$ steps of perturbation theory ($[x]$ is the integer part of $[x]$). So, *if* we work in a region which is free from resonances of order less than $[1/(ab)] + 1$, we will be in business again, as is the case in [CG]. Here, "free from resonance" is a condition on $\omega(p) = \nabla h(p)$ in (2); but remember that in the context of (3) (which is a rewriting of (1)), we are already working along an m-fold resonant plane, and thus in fact trying to avoid resonances of higher multiplicity.

With these considerations, we have ideally completed the construction of the geometric skeleton in Arnold mechanism, namely transversally hyperbolic ("whiskered") invariant tori, invariant manifolds and heteroclinic intersections. There remains to "flesh" it, i.e. find trajectories which run along the bones. We should add that all the ingredients were already present in the original paper by Arnold and the potential difficulties already apparent; there the discussion is conducted on an example which was chosen so as to avoid the difficulties mentioned above and in order to exemplify the – then original – mechanism which allowed for instability of the action variables. Admittedly, explanations were rather terse, or at times cryptic or even implicit, and it took a long time to work out the "details" (such as KAM theory for lower dimensional hyperbolic tori). For the untiring reader, we note that things are redone from scratch in [CG].

(5) Simple versus multiple resonances: In principle, diffusion, or rather global instability should take place all over a given energy surface of a system with a Hamiltonian function as in (1) (see the conjectures in [A1]). By looking instead at perturbations of (2), one introduces a major simplification, namely again avoiding to tackle the *singular* perturbation problem, but also a major restriction; perturbations of (2) model the study of (1) along a given m-fold resonance. So, one implicitly gives up the idea of exploring the whole energy surface, and focusses instead on a given $d = n - m$ dimensional resonant plane. Unfortunately this *still* seems to be at present an almost too difficult situation to deal with, and one has to restrict the object of study even more; in practice, as far as we know, all authors study the case of a resonant hyperplane ($d = n - 1$) that is, of a *simple* resonance ($m = 1$). The reason for this is rather clear and goes as follows. In order to study trajectories which will hopefully trace along the skeleton which has been – potentially – constructed at this point, it is certainly useful to achieve the best possible normal form in the vicinity of the conserved tori; and in this respect, the cases $m = 1$ and $m > 1$ are markedly different. When $m > 1$, the best that can be achieved in the generic case is to arrive (locally near a given torus) at a normal form which displays the existence of the torus, the existence of the stable and unstable manifolds attached to it, and the motion *on* these manifolds, which is surprisingly simple (up to conjugation of course). This is explained and implemented quite clearly in [G], as noted above (see also [T] for the singular case). What is lacking is a simple description of the motion in a full neighbourhood of the torus (outside the local invariant manifolds), and this is simply because this motion may well be very complicated in the generic case. If one thinks in terms of the $2m$ hyperbolic degrees of freedom only (forgetting about the "elliptic" part), we have to deal with the neighbourhood of an hyperbolic fixed point in a $2m$ dimensional space. But then it has long been known that the plane ($m = 1$) case is in fact "integrable", essentially because *no* small divisors arise in the normalising series, which are thus convergent, as proved in [M].

For the full system, one may first apply a KAM algorithm (which is *symplectic* in essence), and then prove the existence of the invariant manifolds (which is *not* a symplectic, but simply hyperbolic phenomenon) as is done in [G]. One can also do everything at one stroke, including further normalisation (assuming $d = 1$), as in [CG]. The point is that in applying a perturbative algorithm, the – potentially small – divisors one comes across will look like $[\lambda\,(\ell - m) + i\,\omega \cdot k]$ where, refering to Hamiltonian (2), one has $(\ell, m) \in {}^{2}, k \in {}^{n-1}, \omega = \nabla h, \Lambda = \lambda > 0$. The fact that the torus is normally hyperbolic is reflected in that the divisors are composed of a purely real part and a purely imaginary one; now, because $m = 1$, it is obvious that this divisor can be small only when $\ell = m$ and one may thus cook a resonant normal form which contains only the corresponding terms, up to a flat remainder. Again, this is analogous to the plane purely hyperbolic case as first discussed in [M], although of course the technical "details" are much more intricate. We refer to [CG], in which however the explicit quantitative normal form one arrives at is not so easy to dig out (see (8.16) and backtrack).

The upshot of the above is that in the generic case, a good normal form in the vicinity of the invariant tori, which is subsequently required in the geometric reasoning leading to the existence of well-behaved trajectories, can be achieved in the case of a simple resonance only. And thus, one – reluctantly – confines oneself to drifts along such simply resonant

hyperplanes. Note that in [D], perturbations with compact supports are used; since they are supported away from the tori (in fact near the heteroclinic intersections), the system remains unperturbed near the tori and transitions in their vicinity can be conveniently analysed (see the conjugacy lemma in [D], III.13); but this is of course "very" non generic – and non analytic as well.

As we shall see below, the above should be qualified, and the need for very accurate normal forms largely comes from the rather rudimentary methods which have been used until recently in the construction of transition chains. More geometrically rooted approaches do require a less detailed knowledge of the motion, as outlined in the sequel.

(6) Arnold mechanism: transition chains: Assuming the skeleton has been constructed, how do we add the flesh? This is described terribly quickly in [A], and "details" really have to be supplied. We first mention two places, namely [D] and [CG], where they have been worked out and written down, albeit in particular cases, and the results as well as the limitations of these papers also serve to illustrate the difficulties that arise in the "traditional" approach (see also [X]).

There are two types of "transitions" to be accounted for: one is the "fast" transition from one torus to the next, sliding along the respective invariant manifolds in a neighbourhood of the heteroclinic trajectory; the other one, the "slow" one (for these denominations, see also section 7 below), is the shift from a neighbourhood of the stable manifold to a neighbourhood of the unstable manifold of the *same* torus. Although it is of a more local nature, this second type is the more difficult to deal with. It is at this point that a good normal form for the system is a priori quite helpful; a sensible question to ask is: what exactly do we need to know about the flow in the vicinity of the torus in order to prove the existence of trajectories which experience the transition we need? This will be discussed a little further below, but in any case, we have already mentioned that in [D], the system is in fact unperturbed in the vicinity of the tori; more precisely, it looks like a skew product, which can be recast in a simple form (cf. [D], III.13.6), from which the necessary properties are immediately read off. In [CG], the good normal form which can be obtained in the neighbourhood of a simply resonant plane allows in principle for a comparatively simple and direct geometric reasoning (it is unfortunate that the normal form is not really explicit, making the subsequent estimates in §8 difficult to check). The technique in [D] is topological in nature, in an essentially C^∞ (not analytic) setting; it also uses the full power of the integrability in the neighbourhood of the tori. In [CG], the approach is quite direct, and eventually results in a superexponential estimate for the speed of diffusion, which seems hard to avoid, if working along these lines (see section 7 below).

If one wants to go to more general situations (in particular higher dimensional resonances and switching from one resonant plane to another), and also get more reasonable quantitative estimates for the speed of diffusion, it seems that new techniques are called for. One way of approaching the problem is via "windowing" (cf. [E], [E1]). We will very informally and briefly discuss it from a more general viewpoint in section 8 below. Suffice it to say, as for now, that it may provide a more flexible and general tool to detect trajectories which run along (shadow) a prescribed path in action space. The results contained in [E] are rather eloquent in that respect. They have been taken up and clarified in [Mar], which also adapt them to the heteroclinic case (i.e. that of Arnold mechanism properly speaking,

while [E] deals with *homo*clinic transitions). It is seen there how a rather crude normal form provides enough information on e.g. obstruction properties, to ensure the existence of connecting orbits and study some of their properties.

We also note that in a recent paper ([B]), working on the original example of Arnold, U.Bessi successfully applied a variational technique to solve the problem in that case, obtaining a close to optimal value for the "speed of diffusion" (see section 7 below) and we shall say somewhat more about this in section 8. In any case, it shows how the variational approach also provides a very promising circle of ideas to study these global instability problems.

(7) Topology versus quantitative estimates: The first concern of mathematicians has been to prove the existence of trajectories with some prescribed behaviour; this leads to essentially topological questions, which leave aside measure theoretic questions or the problem of estimating the "speed" of "diffusion", whatever definitions may be adopted for these notions. One would first like to prove that *topological transitivity on the energy surface* (see [A1]) generically takes place for these systems (this presupposes that a good notion of genericity has been defined in the analytic category, which is not too much of a problem). The above may have given an idea of how far we seem to stand from proving this, provided a) it is true, b) one sticks to Arnold mechanism as a mean to describe instability (see section 8 below). Summarizing again, we have seen that perturbations of (1) lead to singular perturbation theory which has been hitherto rather avoided than dealt with, and that, *independently* of this first difficulty, studies have been confined (for good enough reasons – see above) to diffusion along simply resonant planes; treating the case of multiple resonances (in the analytic framework) and a fortiori the problem of switching from one resonance to another is yet another difficult task ahead of us – at least again if one insists on following this scheme... We refer to [A2] for a concise and enlightening discussion of this geometric problem.

Beyond purely topological problems, tackling for instance questions linked with metric entropy properties appears to be really hard indeed; recall for instance that the existence of metric entropy in the – plane – standard map or in the – plane – "stochastic strip" for a perturbed pendulum are still unsolved problems at the time of this writing (V.F.Lazutkin has announced a proof for the standard map with a large value of the parameter).

Another type of problem, possibly closer to physical concerns, has to do with the speed of diffusion. Classical perturbation theory investigates the stability of the action variables over long times and Arnold diffusion lives over timescales on which *by definition*, perturbation theory breaks up. Here one simply asks, for Hamiltonian (1), how long it will take for the action variables p to experience a drift of order 1. Classical perturbation theory has its crowning achievement in the study of stability over exponentially long times (w.r.t. the inverse of the perturbation parameter ϵ) which was initiated by N.N.Nekhoroshev in [N]. Using a different way of approach, these results were improved in [L] (see also [L1] and [P]), to a point which is probably optimal, so that we do witness the "end" of classical perturbation theory, in the sense that it has been pushed to its limit. More precisely, assuming analyticity and that the integrable part h is convex one proves that the action variables of (1) are stable over time intervals on the order of $\exp(c\epsilon^{-1/(2n)})$ ($c > 0$ some constant). These rigorous results fit quite well with a heuristic reasoning of B.V.Chirikov

in [C] which predicts basically the same quantity for the inverse of the speed of Arnold diffusion; i.e. over longer timescales, instability does take place and perturbation theory does break up. This argument has been rewritten and discussed in [L] (§V.2), to which we refer. Besides, we point out that the use of simultaneous diophantine approximation makes it crystal clear why the exponent $1/(2n)$ arises and why it should be optimal (this is in fact how it was discovered, from the viewpoint of perturbation theory; see [L1] for a discussion). But one should also mention that Chirikov's reasoning, however suggestive, would probably be very hard to turn into a rigorous proof.

Let us briefly recall how one gets an idea about the speed of diffusion, at least of course if one believes in Arnold mechanism, as this will again exemplify the difference between the singular (Hamiltonian (1)) and non singular (perturbations of (2)) cases. Returning to Arnold mechanism, trajectories following along the skeleton of tori and heteroclinic manifolds can be cut into pieces. First there is a certain number $N(\epsilon)$ of steps, each step corresponding to the transition from one torus to the next ($N(\epsilon)$ is thus also the number of tori that are used). Each step may then be divided into two pieces, one "fast" and one "slow", as alluded to above: the fast part is the one which occurs when the trajectory is relatively far away from the tori and basically follows the heteroclinic orbit between two neighbouring tori; the slow part corresponds to the winding of the trajectory in the vicinity of a given torus and its switching from the stable to the unstable manifold attached to this torus. Let $\mathcal{T}(\epsilon)$ be (an upper bound for) the time it takes to achieve one full step; then of course $\mathcal{T}(\epsilon) = \mathcal{T}_f(\epsilon) + \mathcal{T}_s(\epsilon)$ (f and s for fast and slow) and $\mathcal{T}_f(\epsilon) = O(1)$ is negligible w.r.t. $\mathcal{T}_s(\epsilon)$. So, the speed of diffusion can be estimated as $(N(\epsilon)\ \mathcal{T}_s(\epsilon))^{-1}$.

We note however that the above reasoning does *not* immediately correspond to what happens in practice, that is when one tries to derive rigorous upper bounds for the transit time along a transition chain. Then, if one adopts a rather straightforward way of attack (such as in [D] and [CG], which does not however treat the a priori stable case, i.e. perturbations of (1)), things look much more as follows: one gets a *different* estimate $\mathcal{T}_i(\epsilon)$ for each of the steps i corresponding to the transition between tori i and $i+1$ ($i = 1, \ldots, N(\epsilon) - 1$)). Moreover, the estimates deteriorate *at least* geometrically, i.e. $\mathcal{T}_{i+1} \geq c\mathcal{T}_i$ for some constant $c > 1$ (possibly numbering the tori backward). The total transit time is then bounded from above by $\sum_i \mathcal{T}_i(\epsilon)$ and, as the partial sum of a series whose general term increases faster than geometrically is essentially given by the last term, one estimates the speed of diffusion as $(\mathcal{T}_N(\epsilon))^{-1}$ where $N = N(\epsilon)$ is the last step. All this looks rather artificial however and is due to the direct kind of methods that are used; in fact, beating this (super)geometric increase of the estimates for the transit times at every step is difficult but necessary if one wants to ever derive reasonable estimates for the speed of diffusion. This is done in [Mar] using windowing, under some reasonable assumptions which should allow to treat realistic problems, of the type considered in [L]. As for Arnold's original Hamiltonian (considered in [A]), U.Bessi proves in [B] that the transit time is as expected and this again raises hope that this variational approach can be adapted to more general situations.

Returning to the original heuristic argument, we know that the distance between two successive tori is given by the splitting of the respective invariant manifolds (call it $\sigma(\epsilon)$) so that, since one looks for drifts in action on the order of 1 (i.e. a "real" instability, on a finite scale), one has that $N(\epsilon) \sim \sigma(\epsilon)^{-1}$ and the estimate for the speed reads

$(N(\epsilon)\ \mathcal{T}_s(\epsilon))^{-1} \sim \sigma(\epsilon) \times \mathcal{T}_s(\epsilon)^{-1}$. The second factor is a priori algebraic in ϵ (see below), whereas the splitting, *in the singular case* corresponding to (1) is exponentially small. So, in that case, which corresponds to the original problem, the transition time $\mathcal{T}(\epsilon) \sim \mathcal{T}_s(\epsilon)$ can be ignored in the rough computation (it contributes an algebraic prefactor), and one is left with the very simple fact that the speed of diffusion should be on the order of the splitting $\sigma(\epsilon)$ of the invariant manifolds. One can then try and estimate the latter, at least formally (see [L], §V.2) and predict a speed of diffusion of order $\exp(-c\epsilon^{-1/(2n)})$, in agreement with the rigorous upper bound derived from classical perturbation theory. This is where Chirikov's argument and the results of [L], [L1] and [P] very nicely match.

As can be seen from the above reasoning, the evaluation of the splitting between the invariant manifolds of the tori is only part of the difficulty which one meets when trying to rigorize this approach to the problem. We also warn the reader that this is perhaps not even necessary if one can further develop variational (see [B]) or geometric paths (see [Mar]). In any case, the splitting problem is interesting by itself, as part of what is sometimes called "asymptotics beyond all order". It has attracted a lot of attention during the last ten years for problems with two degrees of freedom, starting with pioneering work of V.F.Lazutkin in 1984 on the standard map. We shall only very briefly mention a few facts which seem to be valid at the time of this writing. There has appeared many works dealing with the two dimensional situation and the important problem of the rapidly forced pendulum is basically understood; we refer to [DS], which also contains a short history of this circle of problems, the connection with works on the standard map and further references. It is not perhaps so well-known that the "Méthodes nouvelles de la mécanique céleste" ([Po], end of volume II) contain a lot about this problem, at least from a formal viewpoint; this has been pursued in [S], to which we refer for more details.

As for the many dimensional problem, it seems that it should be considered as completely open. In particular, it would be very interesting to prove the estimate which is heuristically derived in [L] (§V.2) for a Hamiltonian which is introduced there in analogy with Arnold's original example but presents some new features. We only note that it is quite essential to work with an infinite number of harmonics in the perturbation (contrary to most model problems which have been considered in two dimensions) and that there may not exist an asymptotic estimate properly speaking: it would be enough to assert that there exist arbitrarily small values of the perturbation parameter ϵ such that the splitting distance $\sigma(\epsilon)$ is larger than $\exp(-c\epsilon^{-1/(2n)})$ for some constant $c > 0$. Recently, direct proofs of KAM type theorems have been developed, using combinatorial methods and making use of subtle cancellations which occur in the series to prove their convergence. These methods may well be useful to attack the splitting problem; as far as we understand, G.Gallavotti has been able to reprove in this way (see [Ga], Theorem 2 p. 348 and section 8) the flatness of the intersection (i.e. the fact that the splitting distance is smaller than any power of the perturbation parameter) and get at least part of the exact asymtotic results in dimension 2 (see [Ga] p.394, Comment 1 and Footnote 6).

We now briefly turn to the so-called "a priori unstable" case, i.e. perturbations of (2), where these delicate questions of singular perturbation theory do not arise. The splitting $\sigma(\epsilon)$ is easily estimated by the Melnikov method (non singular case) to be of order ϵ, and the transition time $\mathcal{T}_s(\epsilon)$ should then be properly estimated. Now, this quantity is essentially governed by ergodic properties of a linear flow on the d torus along a diophantine vector

satisfying (4). Picking a small δ, the question is: how long does it take for a trajectory of the linear flow along ω to fill the torus ^{d} within δ; in other words, one looks for $T(\delta)$ such that for any $x \in {}^d$, there exists $t,\ 0 \leq t \leq T(\delta)$ such that the point $x + t\omega$ lies within δ of the integer lattice ^{d} ($dist(x + t\omega, {}^d) \leq \delta$); in a broader perspective, this is a kind of transfer result between simultaneous and linear diophantine approximations (see [L], Appendix 1). These "near ergodization times" have been estimated to be on the order of $\delta^{-c(\tau)}$, where τ is the diophantine constant in (4): indeed, an elementary argument shows that $c(\tau) \leq \tau + d$ (G.Gallavotti; private communication), one finds the estimate $c(\tau) \leq \tau + d/2$ in [Du] and J.Bourgain recently proved that the – optimal – value $c(\tau) = \tau$ holds (to appear in a joint paper). In any case, one has to use it for $\delta = \delta(\epsilon)$ following again a power law, so that $\mathcal{T}_s(\epsilon) \sim \delta(\epsilon)^{-c(\tau)}$ ends up being bounded by a power of ϵ.

Finally, the speed of diffuson, which goes like $\epsilon\,\mathcal{T}(\epsilon)^{-1}$, should also be estimated as a power of ϵ. The upshot is that, if one believes in the above, which is of course conjectural in several respects, one tends to predict a *polynomially* slow instability for perturbations of (2), that is in the non singular ("a priori hyperbolic") case (this is to be compared with the superexponential bound derived in [CG]). This is of course in marked contrast with the exponential slowness of Arnold diffusion properly speaking, i.e. the instability developing in systems with Hamiltonians of type (1).

As a test case, one may go back to Arnold's original example, set $\epsilon = 1$, so as to obtain an initially unstable (hyperbolic) system and try to estimate the "speed of diffusion" as a function of the remaining parameter μ (this parameter has the meaning of ϵ above, i.e. the size of the perturbation of a Hamiltonian of type (2)). Note that for dimensional reasons the subtle properties of the ergodization times discussed above do not enter. This problem has been treated by P.Bernard (in [Be]) adapting the variational method of [B]; the result is that the speed is at least on the order of μ^2, which should be close to optimal (may one venture to propose $\mu|\log(\mu)|$ as a possible candidate for the maximal possible speed of drift?).

(8) Genericity, variational and other methods: From the arguments developed above, one does *not* derive the conviction that Arnold mechanism should generically apply in the analytic category for perturbations of (1); it is indeed not at all clear whether one we will be able to build – or rather detect the existence of – the needed skeleton, that is find enough invariant tori along a path in action space, so that the corresponding invariant manifolds will nicely arrange as the treillis of heteroclinic intersecting manifolds we need. Again, one important point here is the respective sizes of the gaps between the surviving invariant tori and the splitting of the invariant manifolds. For a perturbation of size ϵ, we have seen in section 4 that the sizes of the gaps could a priori be of order ϵ^a for some a, $0 < a \leq 1$, whereas one expects an exponentially small splitting, as $\exp(c\epsilon^{-b})$ for some b, $0 < b \leq 1$ (see [L], §V.2; beware of the fact that the constant a and b considered there are not the same as those used here); the invariant manifolds of consecutive tori would then just simply miss each other. This again is typical of perturbations of (1), not of (2) which gives rise to an algebraic splitting. And all this already happens for simple resonances and has nothing to do with the problems discussed above of going to multiple resonances, switching from one resonant plane to another etc. Of course, even if Arnold mechanism turns out to be non generic for analytic Hamiltonian of type (1), this would not

a priori invalidate the conjecture of generic topological transitivity on the energy surface, which may well hold true for other reasons. We would also like to mention (see [L] for some details) that one could imagine studying Arnold mechanism for a skeleton based on periodic orbits only; this case is indeed particular because many more closed orbits (alias tori of dimension 1) survive than tori of other dimensions. Note that these closed orbits are normally invariant sets inside the energy surface, which is not the case for the other higher dimensional invariant tori. The persistence of closed orbits with a threshold in the perturbation which is uniform in the period of the orbits constitutes a highly non trivial fact connected with the Conley-Zehnder theorem and was developed in particular by D.Bernstein and A.Katok (see [BK];convexity of the integrable part h is required). So it could be a priori that enough hyperbolic periodic orbits survive in the generic case (say with strictly convex unperturbed integrable Hamiltonian) so as to justify Arnold mechanism. To our knowledge, this has never been studied – it should also be very hard because even the asymptotic distribution of the closed orbits for an *integrable* Hamiltonian reduces to a problem in simultaneous approximation which is basically unsolved (see [L] again for some details, §IV. 4 and Appendix 1).

Until now, most studies of the instability properties of many dimensional Hamiltonian systems have basically followed the initial paper by Arnold; that is, they tried to justify completely what we call Arnold mechanism for instability. We would like to close with a very informal discussion of possible tracks for further studies, namely the use of variational methods and that of (Conley) index type techniques. There is in fact not one but several (i.e. at least two) variational viewpoints with different but overlapping goals in mind. One has been developed by J.Mather and other mathematicians and a fairly recent state of the art can be found in [Ma]. Another circle of ideas is presented in [CES] and it has now found a direct and striking application to the diffusion problem in [B]. We shall say little about this recent way of attack, although again the paper by U.Bessi shows how promising it is. Trying to find connecting orbits between (the neighbourhoods of) two given points, or orbits with some prescribed behaviour, via a variational method, a priori differs in a decisive way from trying to vindicate Arnold mechanism: in fact no invariant geometric objects are required in principle. One should carefully define a space of admissible paths and a good functional on it, of necessity a variant of the traditional "action functional", so that the trajectories one is looking for appear as extrema of this functional. It often so happens that the functionals one uses are Lipschitz but *not* differentiable, which precludes the use of *minimax* points; only *extrema* are correctly defined and one thus seeks – say – minima of the functional. The only remark we would like to make is that this approach may not in fact be as disjoint as it looks from the index type techniques to be very briefly discussed below. This stems from the fact that one has to impose geometric constraints on the admissible trajectories in order to get a nice space of admissible paths; and of course one does not want the minima to occur on the boundary, as they would not give rise to actual trajectories. So that one will typically look at paths wich are confined in some – broken – tubes of phase space; in other words, one first prescribes a rough behaviour for the admissible paths, which for instance should not leave a certain tubular region (see [B] for an illustration, where the traditional "skeleton" consisting of hyperbolic tori, heteroclinic intersections etc. is used). As will be seen shortly, this is not without resemblance with what happens in the practice of "windowing" or similar techniques, so

that although variational calculus in the large seems a priori almost disjoint from some more direct geometric approaches in phase space, geometry in phase space does somehow reappear through the back door, making the flavour perhaps less different.

In closing, we indulge in a brief piece of advertisement for what we termed above "index type" techniques. Perhaps things could be somehow phrased as follows; fully hyperbolic systems (starting with geodesic flows with negative sectional curvatures) were studied in a detailed way during the sixties, with the gradual emergence of the pervasive property of "shadowing", a very deep global stability property of fully hyperbolic systems. It had somehow been recognized by Hadamard and Morse in particular contexts and came into full light especially in the work of R.Bowen, which made very intensive use of it as a mean to understand the global structure of the trajectories of hyperbolic systems, construct Markov partitions, the so-called BRS (Bowen, Sinai, Ruelle) invariant measure etc. Shadowing comes from full hyperbolicity and can consequently be traced back to the basic fixed point theorem for a contracting map in a Banach space (defined e.g. via pseudo orbits for a hyperbolic map). Relaxing the requirement of hyperbolicity is well-known to be very hard indeed... Symbolic dynamics for Hamiltonian sytems of mixed type was studied in the late sixties by V.M.Alekseev (see [Al]), whose work seems to have remained largely "unread" (at least in "western" countries). This was not of course the case of the work of Conley; it seems that, as far as Arnold mechanism is concerned, one needs not use the full machinery of the Conley index (even the proof of "Poincaré's last theorem" by C.Conley and E.Zehnder uses only part of it) and only some key notions, like that of "isolating block" have to be fully exploited. This leads to the notion of "windowing", developed by R.Easton in a series of papers (see [E], [E1] and references therein). This notion is analogous to shadowing but does not in principle require full hyperbolicity; the aim is very analogous to that of shadowing and the techniques consists in first finding "pseudo-orbits" and then developing the machinery which ensures that if these pseudo-orbits enjoy some prescribed properties (pass through "windows" which properly "line up"), then a *true* orbit exists, close to the given pseudo-orbit. It has been applied to a problem closely connected with Arnold diffusion in [E] and we have already mention that [Mar] constitutes a clarification and direct continuation of this work. We hope these methods can be pushed further in order to overcome some of the obstacles which were discussed above, staying in the framework of Arnold mechanism but improving on the answer to the question: "When we have the skeleton, how do we get the flesh?", in a qualitative and also quantitative way.

Even if this hope turns out to be short lived, we do hope to have somehow convinced the nonexpert reader that the global instability properties of near integrable Hamiltonian systems, thirty years after the pioneering work of V.I.Arnold, are far from well-understood. It could almost be said that little progress has been made, and new ideas are definitely called for. It is an obviously difficult problem, but a fascinating one too, and which naturally stands at the top of the waiting list, especially now that classical perturbation theory and KAM theory have reached a certain maturity.

References

[A] V.I.Arnold, Instability of dynamical systems with several degrees of freedom, Soviet Math. Doklady **5** (1964), 581-585.
[A1] V.I.Arnold, Problems of stability and ergodic properties of classical dynamical systems, Proceedings of the Moscow 1966 ICM, Mir Publ. (1968), 387-392 (in Russian).
[A2] V.I.Arnold, Mathematical problems in classical physics, in *Trends and perspective in Applied Mathematics*, Appl. Math. Sc. Series **100** (1992), Springer Verlag.
[Al] V.M.Alekseev, Quasirandom dynamical systems I, II, III, Math. of the USSR Sbornik **5** (1968), 73-128, **6** (1968), 505-560 and **7** (1969), 1-43.
[BK] D.Bernstein and A.Katok, Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians, Invent. Math. **88** (1987), 225-241.
[B] U.Bessi, An approach to Arnold diffusion through the calculus of variations, Nonlinear Analysis, T.M.A., 1995.
[Be] P.Bernard, Perturbation d'un Hamiltonien partiellement hyperbolique, Note submitted to the CRAS (Comptes Rendus de l'Académie des Sciences).
[C] B.V.Chirikov, A universal instability of many-dimensional oscillator systems, Physics
Reports **52** (1979), 263-379.
[CG] L.Chierchia and G.Gallavotti, Drift and diffusion in phase space, Annales de l'IHP, Section Physique Théorique **60** (1994), 1-144.
[CES] V.Coti Zelati, I.Ekeland, E.Séré, A variational approach to homoclinic orbits in Hamiltonian systems, Math. Annalen **288** (1990), 133-160.
[D] R.Douady, Stabilité ou instabilité des points fixes elliptiques, Annales Scientifiques ENS **21** (1988), 1-46.
[Du] S.Dumas, Ergodization rates for linear flows on the torus, J. of Dynamics and Diff. Equ. **3** (1991), 593-610.
[DS] A.Delshams and T.Seara, An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum, Commun. Math. Phys. **150** (1992), 433-463.
[E] R.Easton, Orbit structure near trajectories biasymptotic to invariant tori, in *Classical mechanics and dynamical systems*, L.Devaney and Z.H.Nitecki Eds, Marcel Dekker Publ. (1981), 55-67.
[E1] R.Easton, Homoclinic phenomena in Hamiltonian systems with several degrees of freedom, J. Diff. Eq. **29** (1978), 241-252.
[G] S.M.Graff, On the conservation of hyperbolic invariant tori for Hamiltonian systems, J. of Diff. Eq. **15** (1974), 1-69.
[Ga] G.Gallavotti, Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable systems; a review, Reviews in Mathematical Physics **6** (1994), 343-411.
[Gr] J.Gründler, The existence of homoclinic orbits and the method of Melnikov for systems in ^{n}, SIAM J. Math. Anal. **16** (1985), 907-931.
[KP] S.Kuksin and J.Pöschel, On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications, in *Seminar on Dynamical Systems, St. Petersburg 1991*, S.Kuksin, V.Lazutkin, J.Pöschel Eds, PNDLE **12**, Birkhäuser Publ. (1994), 96-116.
[L] P.Lochak, Canonical perturbation theory via simultaneous approximation, Russian Math. Surveys **47** (1992), 57-133.
[L1] P.Lochak, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency, Nonlinearity **6** (1993), 885-904.
[LM] P.Lochak and C.Meunier, Multiphase averaging for classical systems, Appl. Math. Sciences Series **72**, Springer Verlag, 1988.
[Ma] J.Mather, Variational construction of connecting orbits, Ann. Inst. Fourier **43** (1993), 1349-1386.
[Mar] J.-P.Marco, Transition le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques, to appear in Annales IHP, Section Physique Théorique.
[M] J.Moser, The analytic invariants of an area-preserving mapping near a hyperbolic fixed point, Commun. in Pure and Applied Math. **9** (1956), 673-692.
[N] N.N.Nekhoroshev, An exponential estimate for the time of stability of nearly integrable Hamiltonian systems, Russian Math. Surveys **32** (1977), 1-65.
[P] J.Pöschel, Nekhoroshev estimates for quasi-convex Hamiltonian sytems, Math. Z. **213** (1993), 187-216.

[Po] H.Poincaré, Méthodes nouvelles de la mécanique céleste, Volume II, Gauthier Villars, 1893.
[Py] A.S.Pyartli, Diophantine approximations on submanifolds of Euclidean space, Functional Anal. and its Applications **3** (1969), 303-306.
[S] D.Sauzin, Résurgence paramétrique et exponentielle petitesse de l'écart des séparatrices du pendule rapidement forcé, Annales de l'Institut Fourier **45** (1995), 453-511.
[T] D.V.Treshtshev, The mechanism of destruction of resonant tori of Hamiltonian systems, Math. Sbornik **68** (1991), 181-203.
[W] S.Wiggins, *Global bifurcations and chaos*, Applied Math. Sc. Series **73**, Springer Verlag (1988).
[X] Z.Xia., Arnold diffusion in the elliptic 3 body problem, Journal of Dynamics and Diff. Eqations **5** (1993), 219-240.

Warning:
As things may turn out to evolve rather quickly, it seems fit to point out that the above text can at best reflect the situation at the time of its writing, namely **September 1995**.

address: Département de Mathématiques, ENS, 45 rue d'Ulm, 75230 Paris Cedex 05
e-mail: lochak@dmi.ens.fr