Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian

P. Lochak
Ecole Normale Superieure, F-75230, Paris, France

A. I. Neishtadt
Space Research Institute, Moscow, 117810, Russia

(Received 20 January 1992; accepted for publication 21 August 1992)

A Hamiltonian system differing from an integrable system by a small perturbation $\simeq\epsilon$ is analyzed. According to the Nekhoroshev theorem, the changes in the perturbed motion of the "action" variables of the unperturbed system are small over a time interval which increases exponentially in length as ϵ decreases linearly. If the unperturbed Hamiltonian is a quasiconvex function of these "actions," the changes in them remain small ($\simeq\epsilon^{1/2n}$) over a time interval on the order of $\exp(\text{const}/\epsilon^{1/2n})$, where n is the number of degrees of freedom of the system.

I. INTRODUCTION

The Nekhoroshev theorem[1] gives an exponential lower estimate of the stability time of analytic Hamiltonian systems which are nearly integrable systems "of general position." According to this theorem, upon a small perturbation, on the order of ϵ, of the Hamiltonian, quantities which were formerly integrals of the unperturbed problem change by no more than ϵ^b over a time $T_*=\exp(c^{-1}/\epsilon^a)$, where a, b, and c are positive constants. The condition "of general position" in this theorem is imposed on the principal (unperturbed) part of the Hamiltonian and is called the "steepness condition."[1] Nonsteep functions are infinitely degenerate.

An important particular case of steep functions is that of quasiconvex functions, i.e., functions whose level surfaces are convex. For example, the principal part of the Hamiltonian of the Solar System is a quasiconvex function of the orbital major semiaxes of the planets.

The primary result of the present study is the following assertion: For systems with a quasiconvex unperturbed Hamiltonian, the Nekhoroshev theorem is valid for $a=1/2n$ and $b=1/2n$, where n is either the number of degrees of freedom of the system, if the unperturbed Hamiltonian is nondegenerate, or the number of fast phases, in the case of an intrinsic degeneracy. This result refines some successively improved estimates of the stability time found previously[1-4] [in Ref. 4, a was $a=1/(2n+2)$].

A universal instability mechanism in systems of this sort is Arnold diffusion.[5,6] Heuristic arguments and numerical estimates of the rate of Arnold diffusion in Refs. 6, 7, and 4 suggests that the estimate $a=1/2n$ is not improvable (but this is not proved rigorously).

The method for deriving estimates which is proposed below is based on an approach which was originally proposed in Ref. 4 and which differs substantially from the original method of Ref. 1. A periodic solution of the unperturbed problem passes by each point in the phase space. Near this periodic solution the overall system has a single rapidly rotating phase, which corresponds to motion along the periodic solution. The standard perturbation-theory procedure furnishes a change of variables which eliminates the dependence of the Hamiltonian on this phase with an exponentially small error.[8] The "action" variable which is the conjugate of this phase (and which we denote by Γ) thus changes at only an exponentially low rate. Consequently, over an exponentially long time [$\simeq\exp(c^{-1}/\epsilon^{1/2n})$, according to calculations], the value of this variable remains constant at exponential accuracy. We can now draw on the geometric discussion of Ref. 2, making use of the quasiconvex nature of the unperturbed Hamiltonian. The phase point in action space should lie near the intersection of the convex level surface of the unperturbed Hamiltonian and the $\Gamma=\text{const}$ plane. This surface and this plane are nearly tangent to each other, so their intersection has a small diameter [$O(\epsilon^{1/2n})$, as can be verified by calculations]. The change in the actions is therefore a bounded quantity on this order.

The estimate $a=b=1/2n$ was also derived by Pöschel,[9] simultaneously, independently, and by a different method, which develops the original method of Refs. 1 and 2.

II. FORMULATION OF THE RESULT

We consider a Hamiltonian system with the Hamiltonian

$$H=H_0(I)+\epsilon H_1(I,\varphi,\epsilon). \qquad (1)$$

Here ϵ is a small parameter $0\leqslant\epsilon<\epsilon_0$; I,φ are conjugate canonical variables; and Hamiltonian (1) is 2π periodic with respect to all components of φ. In the case $\epsilon=0$ the system becomes integrable; I,φ are its action-angle variables, and $\omega(I)=\partial H_0/\partial I$ is the frequency of the unperturbed motion. We assume below that Hamiltonian (1) is a really analytic function of the variables I,φ in a complex δ neighborhood $D+\delta$ of the real region $D=G\{I\}\times T^n\{\varphi\}$, where $G\subset R^n$, n is the number of degrees of freedom of the system, and T^n is an n-dimensional torus $T^n=R^n/(2\pi Z)^n$. We assume that the following estimates hold for $(I,\varphi)\in D+\delta$:

 1054-1500/92/040495-05$001.00

$$|H_0| < c_1, \quad |H_1| < c_1, \tag{2}$$

where c_1 is a positive constant, as are c_i and C_i below.

We assume that the unperturbed Hamiltonian is a quasiconvex function of the variables I in the region $G+\delta$. This assumption means, by definition,[1] that for $I \in G+\delta$ the inequalities,

$$\text{(a)} \ |\partial H_0/\partial I| > c_2^{-1} \quad \text{and} \quad \text{(b)} \ |\xi^T \frac{\partial^2 H_0}{\partial I^2} \xi| > c_3^{-1} |\xi|^2,$$

hold for any ξ from the hyperplane $\{\xi:(\partial H_0/\partial I)\xi=0\}$.

Theorem 1: For the solution $[I(t),\varphi(t)]$ of the system with Hamiltonian (1), under the initial condition $I(0) \in G$, the following holds:

$$|I(t)-I(0)| < C_1 \epsilon^{1/2n} \quad \text{for} \quad 0 \leqslant t \leqslant \exp(C_2^{-1}/\epsilon^{1/2n}).$$

The proof is given in Sec. III.

The case of an intrinsic degeneracy, in which the unperturbed Hamiltonian does not depend on some of the action variables, is important to many applications (including some in celestial mechanics). The perturbed Hamiltonian in this case is

$$H = H_0(I) + \epsilon H_1(I,\varphi,y,x,\epsilon), \tag{1'}$$

where, as before, we have $(I,\varphi) \in D+\delta$, $D = G \times T^n$, and (y,x) are some additional conjugate canonical variables satisfying $(y,x) \in W+\delta$, $W \in R^{2m}$. The number of degrees of freedom is $n+m$. The number of rapidly varying angular variables (phases) is n. We assume that Hamiltonian (1) is really analytic, and satisfies (2) in the complex neighborhood $D'+\delta$ of the real region $D' = G\{I\} \times T^n\{\varphi\} \times W\{y,x\}$. The unperturbed Hamiltonian H_0 is assumed to be a quasiconvex function of the variables I.

Theorem 1': For the solution $[I(t), \varphi(t), y(t), x(t)]$ of the system with Hamiltonian (1'), under the initial condition $I(0) \in G$, $(y(0), x(0)) \in W$, the following assertion holds as long as the condition $(y(t), x(t)) \in W+\delta/2$ holds:

$$|I(t)-I(0)| < C_1 \epsilon^{1/2n} \quad \text{for} \quad 0 \leqslant t \leqslant \exp(C_2^{-1}/\epsilon^{1/2n}).$$

The proof of Theorem 1' is essentially the same as that of Theorem 1, and we will omit it.

III. PROOF OF THEOREM 1

Below we examine the behavior of the solution $I(t),\varphi(t)$ with the initial condition $(I(0),\varphi(0)) = (I_0,\varphi_0) \in G \times T^n$. The quantities $c_4, c_5,...$ are positive constants, i.e., quantities depending on c_1, c_2, c_3, δ and n only. We wish to stress that all the calculations below are valid if ϵ is sufficiently small, specifically, if $\epsilon < \epsilon_1$, where the constant ϵ_1 depends on c_1, c_2, c_3, δ, and ϵ_0. We will not repeat this point.

A. Reduction of the system near the resonance of maximum multiplicity

(1) Let us determine $Q = \epsilon^{-(n-1)/2n}$

Lemma 1: There exists an integer q, $1 \leqslant q < Q$, and a point $I_* \in \mathrm{Re}(G+\delta)$ such that

$$|I_0 - I_*| < \frac{c_4}{qQ^{1/(n-1)}}$$

and such that the ratios of all components of the vector $\omega_* = \partial H_0(I_*)/\partial I$ to the largest of them are rational numbers with a denominator no greater than q.

This lemma is an obvious consequence of the Dirichlet theorem in the theory of Diophantine approximations. (We recall an assertion of Dirichlet theorem:[10] For any $\alpha \in R^m$ and $Q \in R^1$, $Q > 1$, there exists an integer q, $1 \leqslant q < Q$, such that each component of the vector $q\alpha$ differs from an integer by no more than $Q^{-1/m}$.) At the point I_* the vector of unperturbed frequencies is proportional to an integer vector: There is a resonance of multiplicity $n-1$, i.e., of the maximum possible multiplicity. The trajectories of the unperturbed system for $I = I_*$ are periodic with a period T such that $c_5^{-1} q < T < c_5 q$.

(2) We wish to determine a constant $c_6 > 10$ such that, for sufficiently small μ, the intersection of the plane $\omega_*(I - I_0) = 0$ and the surface $H_0(I) = H_0(I_0)$ in the space of the variables I, under the condition $|I_0 - I_*| < \mu$, lies in a sphere of radius $c_6 \mu$ centered at the point I_*. The constant c_6 can be chosen in this way since H_0 is quasiconvex. We introduce

$$R = 4c_6 \frac{c_4}{qQ^{1/(n-1)}}.$$

(3) For the function $f = f(I,\varphi)$ we introduce the following notation: $\bar{f} = \bar{f}(I,\varphi)$ is the resonant part of f for $\omega = \omega_*$, and $\tilde{f} = \tilde{f}(I,\varphi)$ is the nonresonant part. In other words, we have

$$\bar{f}(I,\varphi) = \frac{1}{T} \int_0^T f(I,\varphi + \omega_* t) dt,$$

$$\tilde{f} = f - \bar{f}.$$

We write Hamiltonian H in the form

$$H = \omega_*(I - I_*) + F(I - I_*) + \epsilon \bar{H}_1(I,\varphi) + \epsilon \tilde{H}_1(I,\varphi). \tag{3}$$

Here we have discarded an unimportant constant. The expansion of $F(\cdot)$ around the origin begins with the second-order terms. For brevity, we do not specifically indicate the argument ϵ of the functions.

B. Lemma of the annihilation of nonresonant harmonics and derivation of Theorem 1 from it

The following lemma asserts that nonresonant harmonics in a Hamiltonian can be annihilated within an exponentially small error.

Lemma: There exist change of variables $(I,\varphi) \to (J,\psi)$ for $|J - I_*| < 0.5R$, $|\mathrm{Im}\,\psi| < 0.5\delta$, analytic, simplectic, and approximately identical, which puts Hamiltonian (3) in the form

$$H = \omega_*(J - I_*) + F(J - I_*) + \epsilon \bar{\Phi}_1(J,\psi) + \epsilon \tilde{\Phi}_1(J,\psi),$$

$$|\epsilon \tilde{\Phi}_1| < \exp(-c_7^{-1} \epsilon^{-1/2n}),$$

$$|J - I| < c_8 \epsilon T, \quad |\psi - \varphi| < c_8 \epsilon / R.$$

This lemma is proved in Sec. III C. The result of the theorem follows in the standard way from this lemma and from the circumstance that H_0 is quasiconvex.[2] For completeness, we will go through the corresponding discussion.

From this lemma and from Cauchy estimates[11] for $|J-I_*|<0.5R$, $|\text{Im}\,\psi|<0.4\delta$, we find

$$\frac{d}{dt}(\omega_* J)=-\epsilon\omega_*\frac{\partial\widetilde{\Phi}_1}{\partial\psi}=0(\exp(-c_7^{-1}\epsilon^{-1/2n})).$$

Accordingly, for $0\leqslant t\leqslant\exp(\frac{1}{2}c_7^{-1}\epsilon^{-1/2n})$ we have $\omega_*(J(t)-J(0))=O(\exp(-\frac{1}{2}c_7^{-1}\epsilon^{-1/2n}))$ along a trajectory of the system as long as the condition $|J(t)-I_*|<0.5R$ holds. By virtue of conservation of energy along the trajectory we have $H_0(J(t))=H_0(J(0))+O(\epsilon T)$. Here we have $\epsilon T=o(R)$. The point $J(t)$ must therefore lie in an $o(R)$ neighborhood of the intersection of the $\omega_*(J-J(0))=0$ plane and the surface $H_0(J)=H_0(I(0))$. By virtue of the choice of the constant c_6 in the definition of R (Sec. III A), this intersection itself lies in a sphere of radius $c_6|J(0)-I_*|<0.3R$ centered on the point I_*. The solution of the system is thus determined over the exponentially long time interval under consideration, and we have $|I(t)-I(0)|=O(R)=O(\epsilon^{1/2n})$, which is asserted in Theorem 1.

C. Proof of the lemma

The necessary change of variables is constructed as the composition of a large number of successively determined transformations which annihilate nonresonant harmonics in terms of progressively higher order. The estimates are approximately the same as those used in Ref. 8.

1. Successive-transformation procedure

The Hamiltonian of the system obtained after j changes of variables is written in the form

$$H=\omega_*(I-I_*)+F(I-I_*)+\epsilon\bar{\mathscr{H}}_j(I,\varphi)+\epsilon\tilde{\mathscr{H}}_j(I,\varphi). \quad (4)$$

In the $(j+1)$-st step, the change of variables $(I,\varphi)\to(J,\psi)$ is specified by the generating function $J\varphi+\epsilon S_j(J,\varphi)$, where[4]

$$S_j(I,\varphi)=-\frac{1}{T}\int_0^T t\tilde{\mathscr{H}}_j(I,\varphi+\omega_* t)dt.$$

The function S_j satisfies the homology equation

$$\frac{\partial S_j}{\partial\varphi}\omega_*+\tilde{\mathscr{H}}_j=0,$$

which is standard in perturbation theory. The formulas for the change of variables are

$$I=J+\epsilon\frac{\partial S_j}{\partial\varphi},\quad \psi=\varphi+\epsilon\frac{\partial S_j}{\partial J}. \quad (5)$$

Substituting these relations into the Hamiltonian, we find

$$H=\omega_*(J-I_*)+F(J-I_*)+\epsilon\mathscr{H}_{j+1}(J,\psi),$$

$$\mathscr{H}_{j+1}(J,\psi)=\frac{1}{\epsilon}\left[F\left(J-I_*+\epsilon\frac{\partial S_j}{\partial\varphi}\right)-F(J-I_*)\right]$$
$$+\bar{\mathscr{H}}_j\left(J+\epsilon\frac{\partial S_j}{\partial\varphi},\varphi\right)$$
$$+\tilde{\mathscr{H}}_j\left(J+\epsilon\frac{\partial S_j}{\partial\varphi},\varphi\right)-\tilde{\mathscr{H}}_j(J,\varphi). \quad (6)$$

On the right side of Eq. (6) we need to express ψ in terms of φ in accordance with the formulas for the change of variables in Eq. (5).

2. The estimates

The procedure of successive changes of variables of Sec. III C 1 has a total of N steps, where N has not yet been determined. The Hamiltonian obtained after j steps is studied in the region

$$D_j=\{I,\varphi:|I-I_*|<0.8R-(j-1)\rho,$$
$$|\text{Im}\,\varphi|<0.8\delta-(j-1)\sigma\}$$
$$\rho=R/(10N),\quad \sigma=\delta/(10N).$$

Examining the first step of the procedure, we can easily show that for $(J,\psi)\in D_1$ formulas (5) do indeed determine the change of variables, and the Hamiltonian obtained after the first step satisfies the estimates

$$|\epsilon\tilde{\mathscr{H}}_1|<a_1\epsilon TR,\quad |\epsilon\partial\mathscr{H}_1/\partial I|<a_2\epsilon T<R,$$
$$|\epsilon\partial\mathscr{H}_1/\partial\varphi|<a_2\epsilon<a_3R^2.$$

Here and below, a_i are positive constants.

For an argument by mathematical induction, we adopt the following assumption: We have carried out i changes of variables, where $1\leqslant i\leqslant N$, and for all j, $1\leqslant j\leqslant i$, the following relations hold:

$$|\epsilon\tilde{\mathscr{H}}_j|<\eta_j=(1/2)^{j-1}\eta_1,\quad \text{where}\quad \eta_1=a_1\epsilon TR, \quad (7)$$

$$|\partial\epsilon\mathscr{H}_j/\partial I|<2R,\quad |\partial\epsilon\mathscr{H}_j/\partial\varphi|<2a_3R^2. \quad (8)$$

Let us examine the $(i+1)$-st change of variables. This change is specified by the generating function $J\varphi+\epsilon S_i(J,\varphi)$, where $|\epsilon S_i|<a_4T\eta_i$. It follows from Eq. (5) and Cauchy estimates[11] that if

$$\frac{\eta_i T}{\rho\sigma}<a_5^{-1}, \quad (9)$$

then the change of variables is determined for

$$(J,\psi)\in D_i'=\{J,\psi:|J-I_*|=0.8R-(i-0.5)\rho,$$
$$|\text{Im}\,\psi|<0.8\delta-(i-0.5)\sigma\},$$

and the following relations hold:

$$|I-J|<a_6\eta_i T/\sigma,\quad |\varphi-\psi|<a_6\eta_i T/\rho. \quad (10)$$

It follows from the definition of ρ, σ, and η_i that in order to satisfy Eq. (9) it is sufficient to satisfy the inequality

$$\epsilon T^2 N^2 < a_7^{-1}. \tag{11}$$

For the Hamiltonian in terms of the new variables we find, using Eqs. (6), (8), and (10),

$$\tilde{\mathscr{H}}_{i+1} < a_8 \left[\frac{R\eta_i T}{\sigma} + \frac{R^2 \eta_i T}{\rho} \right] < a_9 RTN\eta_i.$$

We choose $N=[K^{-1}/(RT)]$ with a constant K which is undetermined at this point. We then write

$$\tilde{\mathscr{H}}_{i+1} < \left[\frac{a_9}{K}\right]\eta_i.$$

Choosing $K>2a_9$, we find that Eq. (7) holds for $j=i+1$: $|\epsilon\tilde{\mathscr{H}}_{i+1}| < (1/2)\eta_i = \eta_{i+1}$. We also have $|\epsilon\mathscr{H}_{j+1} - \epsilon\mathscr{H}_j| < a_{10}\eta_j$, where $j=1, 2,..., i$. We thus have $|\epsilon\mathscr{H}_{i+1} - \epsilon\mathscr{H}_1| < a_{11}\eta_1$. Correspondingly, for $(I,\varphi)\in D_{i+1}$ we have

$$\left| \frac{\partial}{\partial I} [\epsilon\mathscr{H}_{i+1}(I,\varphi) - \epsilon\mathscr{H}_1(I,\varphi)] \right|$$

$$< a_{12} \frac{\eta_1}{\rho} < a_{13}\epsilon TN < a_{14} K^{-1}\epsilon/R < a_{15}K^{-1}R.$$

We choose $K>a_{15}$. Then we have $|(\partial/\partial I)\epsilon\mathscr{H}_{i+1}| < 2R$. Therefore, the first inductive estimate in Eq. (8) holds for $j=i+1$. In a similar way, we can prove that the second estimate in Eq. (8) holds for $j=i+1$.

Let us test Eq. (11),

$$\epsilon T^2 N^2 < K^{-2}\epsilon/R^2 < a_{16}K^{-2} < a_7^{-1},$$

for the choice $K > (a_{16}a_7)^{1/2}$. For the K chosen we can thus carry out N changes of variables where

$$N = [K^{-1}/(RT)] > a_{17}^{-1} q Q^{1/(n-1)}/T > a_{18}^{-1} Q^{1/(n-1)}$$

$$> a_{18}^{-1} \epsilon^{1/2n},$$

in accordance with the choice of Q and R in Sec. II A. For the Hamiltonian $\Phi_1 = \mathscr{H}_N$ obtained after N steps of the procedure the following relation thus holds:

$$|\epsilon\tilde{\Phi}_1| < a_{19}(1/2)^N < \exp(-c_7^{-1}\epsilon^{-1/2n}).$$

We find the following result for the difference between the original variables I, φ and the variables J, ψ introduced in step N:

$$|J-I| < a_{20}(\epsilon T + \epsilon TR/\sigma) < a_{21}\epsilon T,$$

$$|\psi - \varphi| < a_{20}(\epsilon T + \epsilon TR/\rho) < a_{21}\epsilon/R.$$

The lemma is thus proved.

IV. ESTIMATES OF THE STABILITY TIME OF THE MOTION AT THE RESONANCE OF MAXIMUM MULTIPLICITY

If, under the assumptions of Theorem 1, the initial point is the point of the resonance of maximum multiplicity, or if it is sufficiently close to this resonance, then the estimate of the stability time can be improved. Below we formulate the corresponding assertion.

We fix the point $I_* \in G$ such that the vector $\omega_* = \omega(I_*)$ is proportional to a vector with integer components.

Theorem 2: For the solution $(I(t),\varphi(t))$ of the system with Hamiltonian (1) under an initial condition $(I(0), \varphi(0))$ such that we have $|I(0)-I_*| < C_0\sqrt{\epsilon}$, the following holds:

$$|I(t)-I(0)| < C_1\sqrt{\epsilon} \quad \text{for} \quad 0 \leqslant t \leqslant \exp(C_2^{-1}/\sqrt{\epsilon}).$$

[The constant $C_0>0$ can be chosen arbitrarily and fixed. The constants $C_{1,2}>0$ depend on c_1, c_2, c_3, δ, n, $\omega(I_*)$, and C_0.]

Proof: We assume that all components of the vector ω_* except the last are zero. This assumption does not restrict the generality of the discussion, since we can use a unimodular change of angular variables to arrange a situation such that the frequency vector is of this form at the point of the resonance of maximum multiplicity. We introduce the following notation: q and p are vectors constructed from the first $n-1$ components of the vectors φ and I, respectively; χ and Γ are the last components of φ and I, respectively; and p_* is the value of p at the point I_*. We introduce $P=(p-p_*)/\sqrt{\epsilon}$. For $P=O(1)$ the system of equations for the new variables has the standard form of a system of equations with one rotating phase χ:

$$\dot{P}=O(\sqrt{\epsilon}), \quad \dot{q}=O(\sqrt{\epsilon}), \quad \dot{\Gamma}=O(\sqrt{\epsilon}), \quad \dot{\chi}=\Omega+O(\sqrt{\epsilon}),$$

$\Omega=\text{const}\neq 0$. This system is a Hamiltonian system with a simplectic structure $\sqrt{\epsilon}dP\wedge dq + d\Gamma\wedge d\chi$. According to Ref. 8, the change of variables $(P,q,\Gamma,\chi)\to(\bar{P},\bar{q},\bar{\Gamma},\bar{\chi})$, which is $O(\sqrt{\epsilon})$ close to identical and which preserves the simplectic structure, can transfer the dependence of the right sides of the system on the phase $\bar{\chi}$ to terms $O(\exp(-c_4^{-1}/\sqrt{\epsilon}))$. The equation for the variable $\bar{\Gamma}$, which is the canonical conjugate of the phase $\bar{\chi}$, is

$$\dot{\bar{\Gamma}}=O(\exp(-c_4^{-1}/\sqrt{\epsilon})).$$

In other words, the motion across the plane $\bar{\Gamma}=\text{const}$ is exponentially slow. Working from this circumstance, the quasiconvex nature of the unperturbed Hamiltonian, and the discussion in Ref. 2 (which was repeated in Sec. III B), we conclude that the slow variables I change by only $O(\sqrt{\epsilon})$ over a time interval $0\leqslant t\leqslant\exp(c_2^{-1}/\sqrt{\epsilon})$. This was the assertion above.

ACKNOWLEDGMENT

This article is based on work presented at the 1991 Soviet-American Conference on Chaos, sponsored by the American Institute of Physics and the Soviet Academy of Sciences and supported in part by the Alfred P. Sloan Foundation.

[1] N. N. Nekhoroshev, Usp. Mat. Nauk 32, 5 (1977) [Russian Mat. Surveys 32, 1 (1977)]; Part II, Tr. Semin. im. I. G. Petrovskogo 5, 5 (1979) [in Russian].
[2] N. N. Nekhoroshev, *The method of the successive transformation of variables*, in Lectures on Hamiltonian Systems, edited by J. Moser (Mir, Moscow, 1973) [in Russian].
[3] G. Benettin, L. Galgani, and A. Giorgilli, Celestial Mechanics 37, 1 (1985).

[4] P. Lochak, "Canonical perturbation theory via simultaneous approximation," Preprint, Ecole Normale Superieure, 1991.
[5] V. I. Arnold, Dokl. Akad. Nauk SSSR **156**, 11 (1964) [Sov. Math. Dokl. **5**, 581 (1964)].
[6] B. V. Chirikov, Phys. Reports **52**, 263 (1979).
[7] V. V. Vecheslavov and B. V. Chirikov, "How fast is the Arnold's diffusion?" Preprint 89-72, Inst. of Nuclear Phys., Novosibirsk, 1989 [in Russian].
[8] A. I. Neishtadt, Prikl. Mat. Mekh. **48**, 197 (1984) [J. Appl. Math. Mech. **48**, 133 (1984)].
[9] J. Pöschel, "On Nekhorochev's estimate for quasi-convex Hamiltonians," Preprint, Forschunginsitut für mathematik, ETH, Zurich, 1991.
[10] J. W. S. Cassels, in *An Introduction to Diophantine Approximation* (Cambridge University, Cambridge, England, 1957).
[11] V. I. Arnold, Usp. Mat. Nauk **18**, 91 (1963) [Russian Math. Surv. **18**, 85 (1963)].