High-precision measurement of separatrix splitting in a nonlinear resonance

V. V. Vecheslavov*) and B. V. Chirikov†)

G. I. Budker Institute of Nuclear Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 11 February 1998)
Zh. Éksp. Teor. Fiz. **114**, 1516–1531 (October 1998)

We present the results of numerical modeling and a theoretical analysis of the splitting of a nonlinear-resonance separatrix in the intermediate asymptotic region for the standard-map model. Direct measurements of the splitting angle $\alpha(K)$, where K is the small parameter of the system, have been carried out over a huge range, $0.1 \gtrsim \alpha \gtrsim 10^{-208}$ ($1 \geqslant K \geqslant 0.0004$), with a relative accuracy greater than one part in 10^{-25} and an average accuracy of roughly one part in 10^{-30}. This made it possible to compare in detail our results with those of the existing asymptotic theory and to detect a number of new effects. We find a relatively simple empirical expression for the α vs. K dependence in the intermediate asymptotic region, and this region proves to be surprisingly broad: $K \lesssim 10^{-2}$. We also study the effect of noise, in particular, errors in measuring the angle, which proved to be much more significant and complicated than expected. Finally, we point out unresolved questions and possible directions of research involving this problem. © *1998 American Institute of Physics.* [S1063-7761(98)02410-X]

1. INTRODUCTION

The dynamics of Hamiltonian (nondissipative) systems is governed by the interaction of nonlinear resonances, which are the basic structural elements of the modern theory of nonlinear oscillations.[1–4] The Hamiltonian of such a system can be written

$$H(I,\theta,t)=H_0(I)+\varepsilon\sum_{n,m} V_{nm}(I)\exp(in\theta+itm\Omega), \quad (1.1)$$

where ε is the small perturbation parameter, I and θ are action–angle variables, Ω is the vector of the frequencies of the external perturbation with harmonics $m\Omega$, and n labels the harmonics of natural oscillations with the unperturbed frequencies

$$\omega(I)=\frac{\partial H_0(I)}{\partial I}. \quad (1.2)$$

Each perturbation term in (1.1) defines a primary resonance:

$$\omega_{nm}\equiv n\omega(I)+m\Omega\approx 0. \quad (1.3)$$

When the oscillations are linear, the frequencies are parameters of the system, which either does or does not wind up in resonance regardless of the initial conditions of motion. The most important feature of nonlinear oscillations is the fact that the oscillation frequencies change in the process of motion because of their dependence on the action variables. Below we examine the case of strong nonlinearity, i.e., when such dependence is present even for the unperturbed frequencies:

$$\frac{\partial\omega(I)}{\partial I}=\frac{\partial^2 H_0}{\partial I^2}\neq 0. \quad (1.4)$$

In this broad class of dynamical systems, the description of nonlinear resonances and the interaction of such resonances proves to be universal and relatively simple.[2]

In numerical models it is more convenient to replace differential equations in terms of the continuous time variable by a discrete map.[1–4] One model of this type that is simple and yet rich in content is the so-called standard map[2] (for a history of the model and related physical applications see Ref. 5):

$$\overline{p}=p+K\sin x, \quad \overline{x}=x+\overline{p}, \quad (1.5)$$

where p and x are action–angle variables. $K \ll 1$ is the sole parameter in the model, which characterizes the effect of the perturbation over a single map period, which is taken to be unity. In terms of the continuous time variable, this model is described by the Hamiltonian (cf. (1.1))

$$H(p,x,t)=\frac{p^2}{2}+K\sum_{m=-\infty}^{\infty}\cos(x-m\Omega t), \quad (1.6)$$

where $\Omega=2\pi$ is the fundamental frequency of the external perturbation. The model comprises an infinite set of highly nonlinear resonances ($\partial^2 H/\partial p^2=1$), with the motion near each of these resonances being identical to within a shift in the action variable: $p-p_m\to p$, where $p_m=m\Omega$ is the resonant value of the action. The frequency of small natural oscillations in any resonance is $\omega_0=\sqrt{K}\ll 1$. Although all the resonances have the same amplitude (K), all except the fundamental, which is governed by the initial conditions, constitute an extremely weak perturbation. This is explained by the high perturbation frequency compared to the natural oscillation frequency at the main resonance:

$$\lambda=\frac{\Omega}{\omega_0}=\frac{2\pi}{\sqrt{K}}\gg 1. \quad (1.7)$$

Interaction among the these resonances is adiabatic, and the effect of such interaction is exponentially small in the large adiabaticity parameter λ.

Neglecting weak nonadiabaticity, we can use the pendulum Hamiltonian

$$H_1(p,x)=\frac{p^2}{2}+K\cos x \tag{1.8}$$

to describe an isolated resonance, say, the one with $m=0$ (see (1.6)). Such a simple form of the resonance Hamiltonian proves to be universal in the case of strong nonlinearity.[2] The most important feature of this problem, which characterizes the nonlinear resonance (1.8), is the presence of a separatrix,

$$p_s=\pm 2\omega_0 \sin(x/2), \quad H_1^{(s)}=\omega_0^2=K, \tag{1.9}$$

a special trajectory that separates phase oscillations (in resonance) from phase rotation (out of resonance). Clearly, near the separatrix the motion of the system specified by (1.8) is highly unstable, since almost any arbitrarily small perturbation changes the nature of the motion dramatically (from rotation to oscillation and vice versa). It is here that chaos can emerge in nonlinear oscillations. As far as we know, this was first observed via numerical modeling,[6] and was subsequently studied by many researchers (see, e.g., Refs. 1–4 and 7).

A simple example of such weak nonadiabaticity is the interaction of just two nonlinear resonances, say, with $m=0$ and $m=1$, which can be described by the Hamiltonian (see (1.6))

$$H_2(p,x,t)=\frac{p^2}{2}+K\cos x+\epsilon\cos(x-\Omega t), \tag{1.10}$$

where ϵ is the amplitude of the perturbing resonance. Motion near the separatrix of the main resonance ($m=0$) can be approximated by the so-called separatrix map, which was first introduced (implicitly) by Zaslavskiĭ and Filonenko.[7] We write this map in the form used in Ref. 2 (see also Ref. 4):

$$\bar{w}=w+W\sin\phi, \quad \bar{\phi}=\phi+\lambda\ln\frac{32}{|\bar{w}|}. \tag{1.11}$$

Here $w=H_2/K-1$ is the dimensionless (energy) deviation from the unperturbed separatrix, $\phi=\Omega t$ is the perturbation phase at the instant when the ''pendulum'' passes the position of stable equilibrium ($x=\pi$), and W is the nonadiabaticity amplitude. If the perturbing resonance is much weaker than the main resonance ($\epsilon\ll W$), the amplitude W can be calculated relatively simply by the Mel'nikov method[8] (see also Refs. 2–4), and for $\lambda\gg 1$ we can write

$$W\approx 8\pi f\frac{\epsilon}{K}\lambda^2 e^{-\pi\lambda/2}. \tag{1.12}$$

Note that this problem cannot be solved by perturbation techniques, because the dependence of this (or in fact any other) nonadiabatic effect ($W(K)$) on the initial perturbation parameter (K) is not analytic at $K=0$. However, after this effect has been isolated, we can employ perturbation-theory techniques, and this has been successfully done for approximate equations of motion of the separatrix-map type. The first to solve this problem was probably Poincaré.[9] Equation (1.12) also shows that in the given model (1.6) the contribution of other resonances ($|m|>1$) is negligible ($\lambda_m=m\lambda$). In this model, however, all the resonance amplitudes are the same, $\epsilon=K$ (see Eq. (1.6)). This leads to the emergence of an additional factor $f\sim 1$ in (1.12). A qualitative explanation of the emergence of this factor as an effect of higher-order approximations of perturbation theory can be found in Ref. 2, where the value of this factor obtained from numerical models is also given:

$$f=2.15\pm 0.04. \tag{1.13}$$

The low accuracy of the measurements (which in fact turned out to be an overestimate) can be explained by the insufficiently small values of K ($K=0.1-1$) and by the approximate nature of the separatrix map itself.

Considerable progress in solving this problem was made only relatively recently (in 1984) by Lazutkin *et al.*[10,11] The value of the correction factor $f=2.2552\ldots$ was obtained by numerically solving an auxiliary equation from which the exponential factor had been eliminated. In contrast to the general expression (1.12), the factor f is not universal as assumed in Ref. 10 but depends on the specific set of interacting resonances.[12]

In their mathematical work, Mel'nikov,[8] Poincaré,[9] Lazutkin *et al.*,[10] and Gelfreich *et al.*[11,12] and others calculated not the effect of nonadiabaticity in (1.11), which was later studied by physicists,[1–7] but an auxiliary quantity, the separatrix splitting angle α. Although this quantity alone is insufficient for reproducing the detailed dynamics near the separatrix, α is an important characteristic of resonance interaction and, in contrast to W, is rigorously defined and can be calculated to any precision.

Separatrix splitting can be approximately described by the map (1.11) as follows. When there is no perturbation ($W=0$), each of the branches of the separatrix (1.9) is an asymptotic trajectory with an infinitely long period, a trajectory that leaves ($w=0$, $t\to\infty$) the position of unstable equilibrium, $x=0 \bmod 2\pi$, and returns to it ($\bar{w}=0$, $t\to+\infty$). When a perturbation is turned on ($W\neq 0$), two intersecting trajectories emerge: one still originates at $x=p=0$ ($w=0$, $t\to-\infty$) but never returns to that point ($\bar{w}\neq 0$, $t\to\infty$), and the other behaves as the first would under time reversal ($t\to -t$). The free ends of the split separatrix form an infinite number of loops with unboundedly increasing lengths,[8,9] which, however, fill a limited and narrow region along the unperturbed separatrix, forming a so-called chaotic layer.[1–7] One important characteristic of such a layer is the layer half-width $w_s\approx\lambda W\approx 4\alpha/\omega_0$ (see Ref. 2), which is directly related to the separatrix splitting angle (see (1.15)). These layers are the universal, ultimate source of chaos in nonlinear oscillations.

The two branches of the split separatrix intersect, in particular, at $x=\pi$ and a certain $p_s(\pi)\approx p_0=2\omega_0$ (see (1.19)). It is at this intersection that the angle α is usually studied, and this is also true of the present work. The intersection of separatrices corresponds to the value $\phi(\pi)=0$ in (1.11).

Near the intersection point the deviation of the two branches of the separatrix from the unperturbed separatrix is described by a simple approximate formula:

$$q_{\pm}(y)=p_{\pm}(x)-p_s(x)\approx\pm W\frac{\omega_0}{4}\sin\frac{\lambda y}{2}, \tag{1.14}$$

where $y=x-\pi$, and we have employed the fact that $dw/dp=2/\omega_0$ and $\phi=\Omega t=\Omega y/p_0=\Omega y/2\omega_0$. Moreover, the variation of w at the intersection point is half the total variation of w in (1.11), since the latter is symmetric about $x=\pi$ (see Ref. 2). As a result, for the total separatrix splitting angle we have

$$\alpha(\lambda)\approx 2\frac{dq}{dy}\approx\frac{\omega_0\lambda W}{4}=2\pi f\omega_0\lambda^3\frac{\epsilon}{\omega_0^2}e^{-\pi\lambda/2}$$

$$=(2\pi)^4 f\frac{e^{-\pi\lambda/2}}{K}, \tag{1.15}$$

where we have used the standard-map parameter $\epsilon=\omega_0^2=K$ and $\Omega=2\pi$. Note that this simple and important relationship holds only for the separatrix splitting angle at $x=\pi$.

The quantity $\mathscr{S}$ introduced in Refs. 10 and 11, which we call the Lazutkin constant, is related to the correction factor f:

$$\mathscr{S}=16\pi^3 f=1118.82770595\ldots\ . \tag{1.16}$$

According to our data, the most accurate value of $\mathscr{S}$ is given below, in (4.14). Note that the factor $f\ll\mathscr{S}$ more correctly characterizes the order of the effect of the finite amplitude of a perturbing resonance.

The last term in (1.15), with an exact value of $\mathscr{S}$ or f, yields the asymptotic value (as $\lambda\to\infty$) of the separatrix splitting angle: $\alpha_\infty=\alpha(\infty)$. In Ref. 10 there is also an estimate of the correction to α_∞ in the intermediate asymptotic region $0<K\ll 1$:

$$c_\alpha(\lambda)=\frac{\alpha(\lambda)}{\alpha_\infty}-1\gtrsim K^{1/8}. \tag{1.17}$$

Our preliminary numerical models have shown, however, that the correction decreases with K much faster: $c_\alpha\sim K^{1/2}$.

We immediately note that this ''correction'' describes the intermediate asymptotic region of separatrix splitting that is of present interest, and hence the formation of a chaotic layer. Solving this problem is the principal goal of our investigation.

The best-developed theory of the standard map separatrix splitting[11] predicts not only a rapid decrease in this correction, but also makes it possible to calculate the first terms in the power-series expansion by numerically solving auxiliary equations. We believe that to a large extent the success of this theory can be attributed to the felicitous change of variables $(K,\alpha)\to(h,\sigma)$, where

$$h(K)=\ln\left(1+\frac{K}{2}+\sqrt{K+\frac{K^2}{4}}\right)\approx\sqrt{K} \tag{1.18}$$

is the positive characteristic index of the tangential (linearized) map (1.5) at the unstable fixed point $x=p=0$,

$$\sigma(h)=\nu(h)\sin\alpha \tag{1.19}$$

is the simplectic invariant, and $\nu(h)$ is a certain norm of the tangent vectors.

Gelfreich *et al.*[11] found an approximate solution to this problem that can also be written as a correction (similar to (1.17)) to the invariant:

$$c_\sigma(h)=\frac{\sigma(h)}{\sigma_\infty}-1=\sum_{m=1}^{\infty}a_\sigma(m)h^{2m}, \tag{1.20}$$

where $\sigma_\infty=4\alpha_\infty$. Actually, Gelfreich *et al.*[11] calculated only the first four terms in the series (see Table II below). The main limitation of this solution is the implicit function $\sigma(\alpha)$, which can be found only numerically because of the additional unknown function $\nu(h)$.

In the present paper we present the first results of direct measurements of the splitting angle of the separatrix of the standard map (1.5), carried out over a broad range of the parameter K: $1\geqslant K\geqslant 0.0004$ $(1\gtrsim h\gtrsim 0.02)$. Here α is $0.1\gtrsim\alpha\gtrsim 10^{-208}$, with a relative accuracy of better than one part in 10^{-25} and an average accuracy of roughly one part in 10^{-30}. In order to solve this problem we developed a special method for measuring and processing the results that used fully portable software[13] with arbitrary-precision arithmetic. In fact, the accuracy of the numerical model was as high as one part in roughly 10^{-300}.

In addition to measuring the angle α, which is directly related to the nonadiabaticity W, we measured the invariant σ (Eq. (19)) it order to compared to the theoretical value. The function $\nu(h)$, which is needed if we wish to calculate σ but whose analytic expression is not known, was calculated using a special program kindly furnished by V. G. Gelfreich, to whom we are sincerely grateful. All this has made it possible to compare our results with those of the analytic theory of Gelfreich *et al.*[11] in detail, which was fully substantiated, and to detect new effects. We found a relatively simple empirical expression for the K dependence of α in the intermediate asymptotic region, which proved to be surprisingly broad: $K\lesssim 10^{-2}$. We also studied the effect of noise—angle measurement errors, in particular—which turned out to be much more significant and complex than expected.

2. MEASURING THE SEPARATRIX SPLITTING ANGLE

The main difficulty in measuring the separatrix splitting angle α relates to its extremely small magnitude. For instance, when the perturbation parameter K in (1.5) is 0.0004, α amounts to roughly 4.2×10^{-208} rad. As noted in the Introduction, this problem could only be solved thanks to a special-purpose software package[13] that implements all standard FORTRAN capabilities in arbitrary-precision arithmetic, where one can specify the number $\mathscr{N}_c$ of significant figures in the mantissa of the decimal representation of a real number.

We calculated α in the following manner. First and foremost, for any given value of the perturbation parameter K, the ordinate $p_s(\pi)$ of the point of intersection of the separatrix branches was found on the line $x=\pi$. The unperturbed theoretical value $p(\pi)=p_0=2\sqrt{K}$ (Eq. (1.9)) was taken as

the initial ordinate, and the behavior of the phase x (rotation or oscillation) was established for the orbit starting at $(\pi, p(\pi))$. Depending on the result, the value of $p(\pi)$ was successively increased or decreased by a certain step size until the counterpart phase behavior emerged. This made it possible to initially localize the intersection point, with the upper bound $p_r(\pi)$ corresponding to rotation of the phase x, and the lower bound $p_l(\pi)$ to phase oscillations independent of the direction of time. Recall that each split branch of the separatrices remains a boundary between oscillations and rotations of the phase for its own particular direction of time (see the Introduction). The step size in momentum was always chosen larger than the expected width of the chaotic layer ($\approx 4\alpha$), so that the width of this initial interval, $dp(\pi) = p_r(\pi) - p_l(\pi)$, turned out to be much greater than the required value $dp_s(\pi)$, which, in contrast, was always chosen to be much less than the width of the layer (see below). For example, for our smallest value of the perturbation parameter, $K = 0.0004$, the width of the chaotic layer was approximately 2×10^{-207}, while the multiple-precision level $\mathscr{N}_c$ and the precision required to yield $dp_s(\pi) = 10^{-\mathscr{N}_a}$ in locating the ordinate $p_s(\pi)$ were set at $\mathscr{N}_c = 300$ (corresponding to a precision of about one part in roughly 10^{-300}) and $\mathscr{N}_a = 280$, respectively.

After establishing the initial interval $[p_l(\pi), p_r(\pi)]$, we used the bisection method to get to the point $dp(\pi) \leqslant dp_s(\pi)$, and then calculated the desired ordinate of the intersection point, taking $p_s(\pi) = [p_r(\pi) + p_l(\pi)]/2$. We note in passing that this stage of the calculations accounted for most of the computing time (see below): as the interval shrank, the orbit moved closer to the unstable fixed point, and the period of motion (the number of iterations of the map) increased substantially.

After the central point of intersection of the separatrix branches was established, it was adopted as the origin of a new coordinate system ($y = x - \pi$, $q = p - p_s(\pi)$), in which all subsequent steps in calculating the angle α were taken. For two values of the phase, $y = \pm \delta y$, $|\delta y| \ll 1$, to the right and left of the new origin ($y = q = 0$) we found (to the same accuracy $dp_s(\pi) = 10^{-\mathscr{N}_a}$) the boundaries between oscillations and rotations (with time increasing toward the future). This made it possible to approximate a section of the straight branch of the separatrix by a second-degree polynomial $q = Q_2(y)$ and, by determining its coefficients, to find the angle of the slope of the separatrix. Repeating the above procedure with time progressing into the past, we were able to find the slope of the reverse branch of the separatrix and take the difference of the two, yielding the desired intersection angle α_2.

Special measures were taken to ensure the feasibility of this approach.

For the full range of the parameter K we strove to obtain a relative measurement accuracy in α no worse than one part in 10^{-25}. To this end we were forced to choose values for all the multiple-precision parameters mentioned above: the number of significant figures $\mathscr{N}_c$, the roundoff error $dp_s(\pi) = 10^{-\mathscr{N}_a}$ incurred in locating the ordinate $p_s(\pi)$, and the phase offset δy needed to set up the approximating polynomials, with proper selection of the third parameter being the most critical. According to (1.14), the phase separation Δy between the central and neighboring intersection points is $\Delta y \approx 2\pi/\lambda = \sqrt{K}$, and it is convenient to choose as a phase offset some fraction of this distance,

$$\delta y = \eta \sqrt{K}, \tag{2.1}$$

with the value of η for all numerical alternative usually ranging from 10^{-16} to 10^{-12}.

Above all, it was necessary to understand whether the desired relative accuracy of one part in 10^{-25} in the calculated value of α_2 had been achieved (clearly, this accuracy must in no event be identified with the number of significant digits $\mathscr{N}_c$). To this end we repeated the above calculations for one more pair of double phase values, $y = \pm 2\delta y$, for both senses of time. Using the results for the two pairs of points, we were able to construct fourth-degree approximating polynomials $q = Q_4(y)$ and obtain an improved value of α_4.

Generally, some of the $\mathscr{N}_a$ of the leading significant digits (with allowance for round-off) in the expressions for α_2 and α_4 are the same, but then these values diverge. We always achieved $\mathscr{N}_a \geqslant 25$, which yielded a relative angle-measurement accuracy no worse than one part in 10^{-25}. Calculations with higher-degree polynomials and higher precision only supported this criterion. In actual fact, the relative accuracy turned out to be higher, and on the average amounted to roughly one part in 10^{-30} (see Sec. 3).

Note that there is no need to calculate the points of the separatrix branches time proceeding into the past. The standard map (1.5) is symmetric, and this symmetry makes it possible to reconcile any point (x_f, p_f) on the forward branch of the separatrix with the corresponding point (x_b, p_b) on the backward branch via the simple relationship[11]

$$x_b = 2\pi - x_f, \quad p_b = p_f + K \sin x_f. \tag{2.2}$$

Thus, the approximating polynomials for both branches are constructed simultaneously, so that there is no need for additional computer time. However, to verify our results we did check calculations of some of the values of the perturbation parameter with time progressing into the past. These calculations validated the numerical scheme describe above.

The major difficulty is to find the coefficients of the polynomials that approximate the segments of the separatrices that pass through the intersection point $y = q = 0$:

$$q = a_1 y + a_2 y^2 + a_3 y^3 + a_4 y^4. \tag{2.3}$$

Using four reference points, we obtain a set of linear equations represented in matrix form as

$$\begin{pmatrix} \delta y & (\delta y)^2 & (\delta y)^3 & (\delta y)^4 \\ -\delta y & (\delta y)^2 & -(\delta y)^3 & (\delta y)^4 \\ 2\delta y & 4(\delta y)^2 & 8(\delta y)^3 & 16(\delta y)^4 \\ -2\delta y & 4(\delta y)^2 & -8(\delta y)^3 & -16(\delta y)^4 \end{pmatrix}. \tag{2.4}$$

The elements of this matrix differ by many orders of magnitude (in view of the smallness of the phase offset, δy, according to (2.1)), and its determinant is close to computer zero, which makes it impossible to invert the matrix.

The problem was solved through a change of variables of the form $(y,q)\to(Y=Sy,\ Q=Sq)$. In terms of the new variables the polynomial was

$$Q=a_1Y+b_2Y^2+b_3Y^3+b_4Y^4. \tag{2.5}$$

We see that the only coefficient of interest to us is $a_1=\tan\alpha_4$, which remains constant for any value of the scaling factor S. This is obvious, since our change of variables amounts only to magnifying the entire picture S-fold. We usually set $S=1/\delta y$, which balanced the orders of magnitude of the matrix elements and resolved all difficulties with processing such a matrix.

We used the above scheme to find $\alpha(K)$ for sixty values of the perturbation parameter K in the range $1\geqslant K\geqslant 0.0004$. The angle was represented to 35 significant digits, which provided more than the minimum (one part in 10^{-25}; see above) and average (one part in roughly 10^{-30}; see Sec. 3) relative accuracy needed to determine the angle.

The minimum perturbation parameter that we could work with, $K=0.0004$, was limited by the time per run (about 24 hours). This time rapidly increases as K gets smaller primarily because of an increase in the number of necessary iterations of the map (Eqs. 1.5). The latter is the product of the period of motion near the separatrix $(\sim|\ln dp_s(\pi)|/\sqrt{K}\sim\mathcal{N}_a/\sqrt{K}$; see Ref. 2) and the number of successive approximations at the intersection point $(\sim|\ln dp_s(\pi)|\sim\mathcal{N}_a)$. Moreover, the time it takes to calculate (1.5) is proportional to $\mathcal{N}_c$ (in the range of moderate accuracy $\mathcal{N}_c\lesssim 300$ that we are interested in). As a result, we estimate the total processing time T with this approach to be

$$T\approx T_0\frac{\mathcal{N}_c\mathcal{N}_a^2}{\sqrt{K}}, \tag{2.6}$$

where the empirical value of the parameter T_0 is roughly 2×10^{-8} h.

3. PROCESSING THE MEASUREMENTS

The initial empirical data used our analysis of $\alpha(K)$ dependence consisted of 60 measured values of the unknown function (Sec. 2). In the first stage of processing, to calculate the corrections c_α we used the most accurate value of $\mathscr{L}$ (1.16).[10] In order to compare our result with the theoretical result of Ref. 11, we processed in a similar way the empirical dependence of the correction c_σ to the invariant σ calculated by (1.19).

In accordance with the theory developed in Ref. 11, both corrections were sought in the form of a finite series expansion in even powers of h.

$$\tilde{c}(h)=a(0)+\sum_{m=1}^{M}a(m)h^{2m}, \tag{3.1}$$

with least-squares interpolation (see, e.g., Ref. 14). Although formally the coefficient $a(0)$ is zero (see (1.17) and (1.20)), incorporating it into (3.1) makes possible a considerable improvement in the parameter $\mathscr{L}$ in comparison to its known value (1.16).

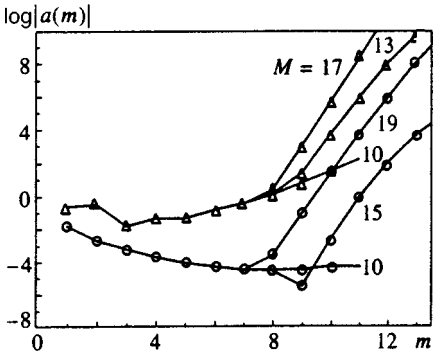

FIG. 1. An example of the dependence of the expansion coefficients in (3.1), found by interpolation of the empirical data, on the total number M of terms in the series for the angle (open triangles) and the invariant (open circles).

The main difficulty in the interpolation of the empirical data via the series (3.1) lies in the fact that the various terms of this series differ by many orders of magnitude. For example, in the typical case where $h\sim0.05$ and $M=10$, the ratio $(a_{10}/a_1)h^{18}$ is of order 10^{-27}. Hence, despite the special methods used in processing the data, even in the initial stage the processing was done with quadruple precision, and in the final stages the required precision amounted to one part in roughly 10^{-100}. But even this was not enough. As in the interpolation of the separatrix branches (Sec. 2), the matrix of the system of linear equations that was so highly inhomogeneous usually became singular during numerical processing. The solution to this problem was found in the same way as in Sec. 2, i.e., by scaling the variables of the problem: $(h,\tilde{c})\to(H=Sh,\ C=S\tilde{c})$, with the scaling factor S much larger than unity. In contrast to the problem of Sec. 2, the coefficients of the polynomial (3.1) are not invariants under such a transformation, and must be restored after interpolation in the new variables: $A(m)\to a(m)=A(m)S^{2m-1}$.

The number of terms in the series (3.1), $M\sim10$, is bounded above by errors in the calculation of the coefficients $a(m)$ primarily because of ''noise'' resulting from the finite precision of the empirical data on $\alpha(h)$. If M is increased beyond 10, we obtain no new coefficients—we even lose some of the old ones. This is especially evident in Fig. 1 from the sharp break in $a(m)$ dependence. We chose the optimum value $M=10$ by trial and error (see also Fig. 3 below).

The accuracy of the empirical formula (3.1) can be characterized by the root-mean-square (rms) error

$$\Delta c=\langle[c(h)-\tilde{c}(h)]^2\rangle^{1/2}, \tag{3.2}$$

where the angle brackets denote averaging over the entire interpolation interval. The latter does not necessarily include all 60 values of $c(h)$. Furthermore, attempts to use the entire empirical interval have revealed the extremely low accuracy of such ''global'' interpolation: $\Delta c(h)\gtrsim10^{-6}$ (cf. Fig. 2). This is quite natural, since the theoretical power-law dependence[11] (3.1) characterizes only the intermediate asymptotic region. For this reason, the deviation

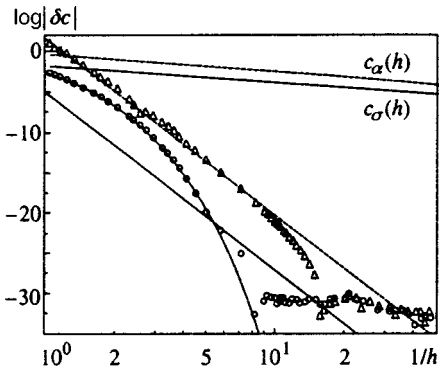

FIG. 2. Interpolation of the empirical data on the separatrix splitting angle (open triangles) and the invariant (open circles); $\delta c(h)$ is the deviation from the intermediate asymptotic behavior specified by (3.1). The oblique straight lines represent the first term of the remainder (4.7), and the curves represent the exponential deviation (4.6). The lines at the top represent the total corrections $c(h)$ for the angle and invariant, respectively.

$$\delta c(h)=c(h)-\tilde{c}(h) \tag{3.3}$$

contains the most valuable and interesting information about additional nonadiabatic effects not in the theory. Thus, in the given set of empirical data, it was also necessary to choose an interpolation range (h_1-h_2), where $h_1=0.02$ is the minimum value in the original data. The basic criteria here were the smallest value of Δc (see Eq. (3.2)) and the precision of the derived coefficients of the series (3.1) (see Tables I and II).

The quality of interpolation decreases not only as h_2 increases (as noted earlier), but also as it decreases, because of the small contribution of high powers of h for small values of h, and also due to the decreasing number of points N_p participating in the interpolation. The interpolation is best when $h_2\approx 0.063$ $(N_p=19)$ and $\Delta c_\alpha\approx 2\times 10^{-31}$, and when $h_2\approx 0.12$ $(N_p=36)$ and $\Delta c_\sigma\approx 2\times 10^{-31}$, but other values of N_p close to those just mentioned were also used (see Sec. 4). Note that the eventual interpolation accuracy yields a rough estimate of the average accuracy of the measured values of $\alpha(K)$.

The basic results were obtained through standard interpolation by minimizing the variance $(\Delta c)^2$ (see (3.2) and (3.1)). To control the procedure, we used forward interpolation, in which for the zeroth-order unknown coefficient $a(0)$ in expressions of type (3.1) we took the total value of the Lazutkin constant. Relative interpolation was also carried out, and with it we minimized the relative variance

$$\frac{(\Delta c)^2}{c^2}=\left\langle\left[1-\frac{\tilde{c}(h)}{c(h)}\right]^2\right\rangle. \tag{3.4}$$

The results of all three interpolation schemes are in good agreement with one another.

The results of processing the data are depicted in Fig. 2 and listed in Tables I and II.

The accuracy of the coefficients was established in two different ways. First, we calculated the standard rms interpolation error[14] (columns headed $\langle\delta\rangle$ in Tables I and II). This quantity characterizes the expected error in calculations of the coefficients for a random rms error in the empirical data. Actually, however, the total error is almost never purely random but contains a systematic error, which leads to a shift (additional variation) in the values of the coefficients. This is clear in Fig. 2, where the accuracy of interpolation is much greater than the remainder in the series (1.20), which is not included in the interpolation of (3.1). This is why another method was used. The values of the coefficients were determined by averaging over several interpolations with various amounts of initial data: $N_p=14-20$ for the angle and $N_p=33-38$ for the invariant. These values are listed in Tables I and II (columns headed $a_\alpha(m)$ and $b_\sigma(m)$, respectively). The relatively weak dependence of the averages on N_p served as the main criterion in selecting these two groups. For the error we took the rms errors of the coefficients in a group (columns headed Δ in the tables). We see that the two estimates are of the same order, although the error in a group is the greatest, and hence dominant, in all cases except the one with $b(10)$. The difference between the two rms errors, which is especially appreciable for the angle, definitely suggests that there is a systematic error. The values of the rms error in a group determine the number of valid significant figures in the coefficients (in our opinion). In the columns headed $a_\alpha(m)$ and $b_\sigma(m)$, we have left two or three ''superfluous'' digits for comparison with future more accurate empirical and/or theoretical values.

TABLE I. Coefficients $a_\alpha(m)$ in the series (3.1) for the angle.

m	$a_\alpha(m)$	Δ	$\langle\delta\rangle$
1	−0.23337 64288 64381 61062 76396 19	0.239×10^{-24}	0.118×10^{-24}
2	−0.29081 81551 24688 86036 776	0.101×10^{-20}	0.453×10^{-21}
3	−0.01482 49555 34894 05088 4	0.240×10^{-17}	0.973×10^{-18}
4	0.04318 21901 48643 630	0.357×10^{-14}	0.130×10^{-14}
5	−0.04151 92394 77208	0.348×10^{-11}	0.115×10^{-11}
5	−0.13137 33101 9	0.227×10^{-08}	0.673×10^{-09}
7	−0.31916 9504	0.983×10^{-06}	0.261×10^{-06}
8	−1.06063 5	0.273×10^{-03}	0.647×10^{-04}
9	−4.3613	0.439×10^{-01}	0.923×10^{-02}
10	−24.02	$0.312\times 10^{+01}$	$0.579\times 10^{+00}$

Note: Here and in Table II, Δ is the rms error in the group, and $\langle\delta\rangle$ is the average rms error in an individual interpolation.

TABLE II. Coefficients $b(m)$ in the series (4.1) for the invariant.

m	$b_\sigma(m)$	Δ	$\langle\delta\rangle$
1	18.59891 19582 09297 35881 71520	0.101×10^{-22}	0.343×10^{-23}
2	−4.34114 12705 68162 53678 60	0.369×10^{-19}	0.125×10^{-19}
3	−4.18326 37590 91894 112	0.971×10^{-16}	0.345×10^{-16}
4	−4.93413 95907 30929	0.186×10^{-12}	0.735×10^{-13}
5	−10.64548 64427 41	0.263×10^{-09}	0.121×10^{-09}
6	−35.86008 1765	0.276×10^{-06}	0.151×10^{-06}
7	−177.60356	0.212×10^{-03}	0.139×10^{-03}
8	−1239.507	$0.114\times10^{+00}$	0.894×10^{-01}
9	−11766.0	$0.386\times10^{+02}$	$0.362\times10^{+02}$
10	−163000	$0.627\times10^{+04}$	$0.693\times10^{+04}$

Note: The underlined figures are the values of the coefficients obtained in Ref. 11.

4. DISCUSSION

We start by examining the behavior of the expansion coefficients in (3.1) on the basis of the data in Tables I and II. We use the representation of the coefficient in the simplest form (3.1) except when we compare our results directly with those of Ref. 11, where the coefficients $b(m)$ are represented by the Taylor series (cf. (1.20))

$$\frac{\sigma}{\sigma_\infty}\mathscr{L}=\mathscr{L}+\sum_{m=1}^{\infty}\frac{b(m)}{m!}h^{2m},\quad b(m)=a(m)m!\,\mathscr{L}. \tag{4.1}$$

First and foremost, we were able to find a relatively simple extrapolation of $a(m)$ outside the range of direct measurements of the separatrix splitting angle. The results of this extrapolation are depicted in Fig. 3 and can be described by the approximate expressions

$$a_\alpha(m)\approx A_\alpha\frac{e^{\gamma m}}{m^p} \tag{4.2a}$$

for the angle, and

$$a_\sigma(m)\approx\frac{A_\sigma}{m^p} \tag{4.2b}$$

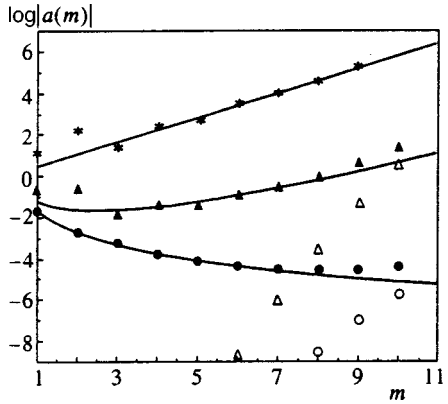

FIG. 3. Variation of the coefficients in the intermediate asymptotic region specified by (3.1) for the angle $a_\alpha(m)$ (filled triangles), the invariant (filled triangles), the invariant $a_\sigma/(m)$ (filled triangles), and their a_α/a_σ ratio (asterisks). The solid curves represent the empirical formulas (4.2a) and (4.2b). The open triangles and circles represent the errors in the coefficients.

for the invariant.

These empirical relationships were built according to the following reasoning. From Fig. 2 and Tables I and II we see that the positive correction $c_\sigma(h)>0$ is much smaller in absolute value than $c_\alpha(h)<0$ over the entire range of h studied here, and the two corrections have opposite signs. This is a consequence of the behavior of the norm $\nu(h)=\nu_\infty(1+c_\nu(h))$ ($\nu_\infty=4$ and $c_\nu(h)>0$) in (1.19). This yields the following relationship between the coefficients:

$$a_\sigma(m)=a_\nu(m)+\sum_{k=1}^{m}a_\alpha(k)a_\nu(m-k). \tag{4.3}$$

All the norm coefficients $a_\nu(m)$ are positive, while almost all $a_\alpha(m)$ and $a_\sigma(m)$ are negative (with the exception of $a_\alpha(4)$ and $a_\sigma(1)$; see Tables I and II). On the other hand, Fig. 3 shows that the coefficients for the angle and invariant behave quite differently. The latter decrease relatively slowly with increasing m, approximately as power-law functions, while the former increase very rapidly, almost exponentially, as the norm coefficients do. This is corroborated by forward interpolation and directly follows from the fact that the norm balances the large correction for the angles almost perfectly: $c_\sigma\ll|c_\alpha|$.

This, then, is the key to building the empirical relations (4.2a) and (4.2b). Indeed, such balance does not occur for a purely exponential dependence of $a_\alpha(m)\approx a_\nu(m)$. One can easily verify, however, that this becomes possible for the combined dependence (4.2a), provided that $p>1$ and $m\gg1$. Figure 3 also depicts the dependence of a_α/a_σ on m, which is perfectly fit by an exponential (see (4.4) below). This is all the more remarkable since the last two or even three coefficients of the invariant clearly deviate from the simple power law (4.2b). The same is true of the combined dependence (4.2a) for the angle.

The nature of these anomalies remains an open question, and requires further study. Note that anomalies appear only in the highest-order coefficients, which cannot be increased in number without a catastrophic increase in the errors (Fig. 1). Furthermore, the ratio of the coefficients of the angle and invariant in the intermediate asymptotic region contains no obvious anomalies. Of course, the exact dependence of the coefficients, provided that it can be expressed explicitly, hardly has the simple form of (4.2a) and (4.2b) even when

$m \gg 1$. This is clear, if only from the fact that the accuracy of the empirical relations can in no way be compared to the precision of the coefficients themselves (see Tables I and II and Eqs. (4.4) below). Nevertheless, even the approximate formulas (4.2a) and (4.2b) help in interpreting the characteristic features of the intermediate asymptotic region and in further studies of this problem. In this paper, however, we have limited ourselves to this approximation.

The parameters of the empirical relations can be obtained by interpolating data on the ratio of the coefficients with $m=7-10$ on a semilogarithmic scale, and data on the invariant with $m=4-6$ on a log–log scale. The results are as follows:

$$\gamma = 1.358 \pm 0.0059, \quad p = 3.51 \pm 0.17,$$
$$A_\alpha = -0.0174(1 \pm 0.28), \quad A_\sigma = -0.0233(1 \pm 0.27), \tag{4.4}$$

where the accuracy is given in terms of absolute and relative rms errors. Note that interpolation with just three coefficients $(m=2,3,4)$, obtained in Ref. 11, yields similar values of the parameters in (4.2b):

$$A_\sigma = -0.0216(1 \pm 0.45), \quad p = 3.38 \pm 0.41. \tag{4.5}$$

The accuracy of the empirical relations (4.2a) and (4.2b) can also be characterized by the relative rms error of the extrapolation itself ($\Delta a/|a|$ for the values of m specified above), which amounts to ± 0.029 for the angle and ± 0.028 for the invariant. Note that the relative smallness of the rms errors themselves, compared to the rms errors in the coefficients (4.4), can be explained in terms of the strong correlation of the latter. In both cases the anomalies of the coefficients are much larger than the reduced errors in the values of (4.2a) and (4.2b) and the two errors of the coefficients themselves in Tables I and II (see Fig. 3).

Using (4.2a) and (4.2b), we can set up an approximate model of the intermediate asymptotic behavior, taking for the initial coefficients, which clearly do not obey (4.2a) and (4.2b), their exact values from Tables I and II. Preliminary experiments along these lines show that the model indeed reproduces the shift of all (and especially the last) coefficients. However, this shift lies within the limits of error (see Tables I and II and Fig. 3) and does not explain the anomalies discussed earlier. Furthermore, even turning on additional noise that is uniform in h and simulates the errors in angle measurements does not help.

The approximate relationships (4.2a) and (4.2b) also make it possible to perceive global behavior in the intermediate asymptotic region. Above all, the series (4.2b) for the invariant is convergent over the entire range $h \leqslant 1$ up to global chaos limit, although it does not describe the actual behavior of the invariant for $h \gtrsim 0.14$. Here the presence of an exponential is clear (see Fig. 2):

$$|\delta c(h)| \approx 63 e^{-\pi^2/h}. \tag{4.6}$$

This function exceeds the remainder (4.7) and describes the perturbation of the separatrix by a distant resonance with a frequency $2\Omega = 4\pi$ (see (1.6)). The simple theory in Ref. 2 predicts a numerical factor of 8, i.e., smaller by a factor of almost 10. Such a discrepancy can easily be explained by another (unknown) value of the factor f in (1.12) and a system of resonances of higher-order approximations, which is much more complicated.[15]

The situation for the separatrix splitting angle is much more interesting, since the series (4.2a) diverges when $h > h_{\rm cr} \approx \exp(-\gamma/2) \approx 0.507$, i.e., within the range under investigation (15 leftmost points in Fig. 2). At the same time, no singularities or anomalies in the behavior of the function $c_\alpha(h)$ or its deviation $\delta c_\alpha(h)$ from interpolation has been observed in this range. Furthermore, the deviation can be described perfectly well by the remainder term in (4.7), which, incidentally, is represented in Fig. 2 only by its first term. This clearly shows that there is a significant change in the behavior of both the angle and the invariant at large $h \sim 1$ as compared with the intermediate asymptotic region.

The approximate relations (4.2a) and (4.2b) make it possible to more accurately (than in Ref. 11) estimate the remainder in the series (1.20), which is not included in the interpolation (3.1):

$$R(h,M) = \sum_{m=M+1}^{\infty} a(m) h^{2m} \approx a(M+1) h^{2M+2}. \tag{4.7}$$

Figure 2 shows that even the first term in R provides a fairly good description of the behavior of the angle deviation over the entire range under investigation. The same can be said of the exponential function (4.6) for the invariant. However, an attempt to add both expressions to the polynomial (3.1) catastrophically reduces the interpolation accuracy: $\Delta c \sim 10^{-6}$. This again demonstrates that the structure of the region where $h \sim 1$ is extremely complicated. In view of the importance of this region in many applications, this problem deserves further study.

Finally, we focus on the most precise and accurate way of finding the Lazutkin constant $\mathscr{L}$. We introduce the correction

$$\delta \mathscr{L}_i = (\mathscr{L}_i - \mathscr{L}_0) \times 10^{23}, \tag{4.8}$$

where i labels the various ways of obtaining $\mathscr{L}$, and

$$\mathscr{L}_0 = 1118.82770\ 59409\ 00778\ 41514\ 639 \tag{4.9}$$

is the zeroth-order value, which we already obtained in preliminary numerical modeling (cf. (1.16)). Formally, the highest accuracy is achieved in individual interpolation with $N_p = 14$ and $M = 10$:

$$\delta \mathscr{L}_\alpha = 0.323560 \pm 0.000017,$$
$$\delta \mathscr{L}_\sigma = 0.323572 \pm 0.000017. \tag{4.10}$$

Since the intermediate asymptotic series (4.2) differ substantially in these two cases, such good agreement is a serious argument favoring the reality of this accuracy.

In view of the importance of this constant, we also used other methods to determine it. First we isolated a group of variants of interpolation schemes with the same value $M = 10$ but different values $N_p = 12-20$ and $N_p = 13-38$ for the angle and invariant, respectively (35 cases in all). The group was chosen to be as broad as possible, the only limitation being that $\delta \mathscr{L}_i$ was supposed to decrease rapidly and

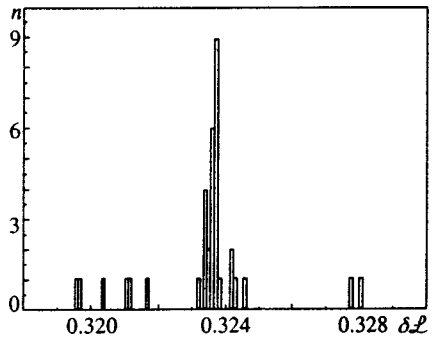

FIG. 4. Bar diagram of the correction (4.8) of the value of the Lazutkin parameter (4.9): n is the number of different values of the correction in a bar with a width of 10^{-4}.

monotonically for reasons discussed earlier. For the average value and the rms error in this group we found

$$\langle \delta\mathscr{L}_i\rangle_{35} = 0.323368 \pm 0.0017. \tag{4.11}$$

We see that the error has increased significantly. However, a histogram of the distribution of $\delta\mathscr{L}_i$ in the group (Fig. 4) shows that the main peak is much narrower. If we discard the wings of the distribution (two cases to the right and six to the left), we obtain

$$\langle \delta\mathscr{L}_i\rangle_{27} = 0.323654 \pm 0.00029. \tag{4.12}$$

Finally, leaving only the nine cases in the rightmost cell of the histogram, we find that

$$\langle \delta\mathscr{L}_i\rangle_{9} = 0.323660 \pm 0.000021. \tag{4.13}$$

Weighing the pros and cons, we conclude that the most accurate value of the Lazutkin constant and the error are

$$\mathscr{L} = \underline{1118.82770\ 59409\ 00778\ 41514\ 63932\ 3566} \pm 3\times 10^{-27}. \tag{4.14}$$

We have underline the value of $\mathscr{L}$ previously obtained with the approximate theory of Lazutkin *et al.*[10]

The results of the present work corroborate the theory of Gelfreich *et al.*[11] both qualitatively (the form of the intermediate asymptotic series (3.1)) and quantitatively (Table II). Furthermore, we have found the intermediate asymptotic region directly for the separatrix splitting angle, which is important in and of itself, and which has made it possible, when combined with the data on the invariant, to obtain approximate empirical relations (4.2) in this region that are not limited by the number of directly derived coefficients.

We are grateful to N. S. Dikanskiĭ for the opportunity to use the ALPHA-4100 computer, to V. G. Gelfreich for useful discussions and for the assistance rendered in the process of comparing our results with the theoretical results, to A. G. Grozin for providing the arbitrary-precision software[13] and helping to master the program, and to L. F. Khaĭlo for considerable assistance with numerical modeling. This work was partially supported by a grant from the Russian Fund for Fundamental Research (Grant No. 97-01-00865).

*)E-mail: vecheslavov@inp.nsk.su
†)E-mail: chirikov@inp.nsk.su

[1] G. M. Zaslavskiĭ and B. V. Chirikov, Usp. Fiz. Nauk **105**, 3 (1971) [Sov. Phys. Usp. **14**, 549 (1972)].
[2] V. V. Chirikov, Phys. Rep. **52**, 263 (1979).
[3] R. Z. Sagdeev, D. A. Usilov, and G. M. Zaslavskiĭ, *Nonlinear Physics: From the Pendulum to Turbulence and Chaos*, Harwood Academic, New York (1988).
[4] A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, Springer-Verlag, Berlin (1992).
[5] B. V. Chirikov, Zh. Éksp. Teor. Fiz. **110**, 1174 (1996) [JETP **83**, 646 (1996)].
[6] N. N. Filonenko, R. Z. Sagdeev, and G. M. Zaslavsky, Nucl. Fusion **7**, 253 (1967).
[7] G. M. Zaslavskiĭ and N. N. Filonenko, Zh. Éksp. Teor. Fiz. **54**, 1590 (1968) [Sov. Phys. JETP **27**, 851 (1968)].
[8] V. K. Mel'nikov, Trudy Mosk. Mat. Obsch. **12**, 3 (1963) [Trans. Moscow Math. Soc. **12**, 1 (1965)].
[9] A. Poincaré, *Les méthodes nouvelles de la méchanique céleste*, Gauthier-Villars, Paris (1892), p. 226.
[10] V. F. Lazutkin, I. G. Schachmansky, and M. B. Tabanov, Physica D **40**, 235 (1989).
[11] V. G. Gelfreich, V. F. Lazutkin, and N. V. Svanidze, Physica D **71**, 82 (1994).
[12] V. G. Gelfreich, *Reference System for Splitting of Separatrices*, Universidad de Barcelona, Mathematics Preprint Series No. 163 (1990).
[13] D. H. Bailey, ACM Trans. Math. Softw. **19**, 288 (1993).
[14] B. M. Shchigolev, *Mathematical Analysis of Observations*, Elsevier, New York (1965); D. Hudson, *Statistics*, CERN, Geneva (1964).
[15] V. V. Vecheslavov, Zh. Eksp. Teor. Fiz. **109**, 2208 (1996) [JETP **82**, 1190 (1996)].

Translated by Eugene Yankovsky