SLAC – PUB – 3865
LBL – 21709
January 1986
(A/T)

INVARIANT TORI THROUGH DIRECT SOLUTION OF THE HAMILTON-JACOBI EQUATION*

R. L. WARNOCK

Lawrence Berkeley Laboratory
University of California, Berkeley, California 94720

and

R. D. RUTH

Stanford Linear Accelerator Center
Stanford University, Stanford, California, 94305

Submitted to *Physica D*

* Work supported by the Department of Energy, contracts DE-AC03-76SF00098 and DE-AC03-76SF00515.

ABSTRACT

We explore a method to compute invariant tori in phase space for classical non-integrable Hamiltonian systems. The procedure is to solve the Hamilton-Jacobi equation stated as a system of equations for Fourier coefficients of the generating function. The system is truncated to a finite number of Fourier modes and solved numerically by Newton's method. The resulting canonical transformation serves to reduce greatly the non-integrable part of the Hamiltonian. Successive transformations computed on progressively larger mode sets would lead to exact invariant tori, according to the argument of Kolmogorov, Arnol'd, and Moser (KAM). The procedure accelerates the original KAM algorithm since each truncated Hamilton-Jacobi equation is solved accurately, rather than in lowest order. In examples studied to date the convergence properties of the method are excellent. One can include enough modes at the first stage to get accurate results with only one canonical transformation. The method is effective even on the borders of chaotic regions and on the separatrices of isolated broad resonances. We propose a criterion for the transition to chaos and verify its utility in an example with $1\frac{1}{2}$ degrees of freedom. We anticipate that the criterion will be useful as well in systems of higher dimension.

1. INTRODUCTION

Although studies of chaotic motion have dominated nonlinear mechanics in recent years, the study of regular motion and its stability is still an urgent matter in several fields of research. In particular, problems of a new order of difficulty have emerged in the design of very large particle accelerators and storage rings. In such machines, particles must be held on narrowly confined orbits over stupendous intervals of time. For instance, in the proposed SSC (Superconducting Supercollider) a proton would make 10^9 revolutions in a ring of 100 km circumference and interact 10^{12} times with localized nonlinear magnetic fields. Large deviations from desired orbits due to isolated broad resonances may be a dominant mechanism of particle loss. In regions of phase space where single broad resonances do not dominate, transitions to chaotic behavior associated with overlapping resonances may contribute to beam degradation.

To study these questions, designers of accelerators rely on low-order perturbation theory, isolated resonance models, and above all 'particle tracking'.[1–3] In practice tracking amounts to calculating orbits of single particles in external fields for a few initial conditions, by approximate integration of Hamilton's equations. (Because of the special nature of the problem, the integration methods are not always the usual ones for ordinary differential equations; one uses the 'kick approximation', symplectic mapping methods, etc.) Much effort is devoted to creating integration schemes valid over large time intervals, but inevitable limitations on accuracy and computation time restrict the usefulness of the method. Although steady improvements in technique and computational facilities can be expected, it seems that tracking is not a fully adequate approach for the largest storage rings considered. In such rings, it is difficult to follow orbits over time intervals sufficiently long to judge their stability. Furthermore, one can usually afford to try only a few initial conditions.

Thus we are led to study a possible complementary approach, namely improved methods for direct computation of invariant surfaces in phase space. Such

methods would be particularly useful if they could deal with strong nonlinear perturbations which result in large distortions of invariant tori or large scale chaotic behavior. To an encouraging extent the methods to be described do allow strong perturbations, and they should be useful in a variety of problems in particle beam mechanics, celestial mechanics, semiclassical quantum mechanics, and plasma theory.

The use of perturbation theory for approximate calculation of invariant tori has a long history, associated with names such as Stokes, Delaunay, Lindstedt, Poincaré, von Zeipel, Birkhoff, and Siegel. References 4-6 provide an introduction and bibliography. In recent years new variants of perturbation theory[7-13] and other methods based on infinite expansions[14] have been carried to high orders with the help of computer programs that do symbolic manipulation. Although convergence has not been proved, the results must be taken seriously because of good agreement with direct integration of Hamilton's equations. On the other hand, the one scheme that is known to be convergent, the Kolmogorov-Arnol'd-Moser (KAM) superconvergent perturbation theory,[4,5] has received scant attention as a possible computational device. The point to be emphasized in comparisons to the present work is that most current schemes, being based on series expansions of one type or another, increase rapidly in complexity as they are carried to higher order.

Our concern in the present paper is to introduce iterative methods in which the algebraic complexity at the n^{th} iteration does not increase with n. Our approach stems from a point of view that is elementary in functional analysis and modern numerical analysis; namely, to state the problem as a fixed point problem on an appropriate vector space, and then apply techniques suggested by the contraction mapping principle, the implicit function theorem, and Newton's method.[15,16] Such methods are generally more tractable than series expansions, both in theoretical studies and in numerical realizations.

Our fixed point problem is equivalent to the Hamilton-Jacobi equation for

the generating function G of that canonical transformation which makes the Hamiltonian a function of action variables alone. By separating the linear part of the equation and taking a Fourier transform, we are able to write it as

$$g = A(g) \quad , \tag{1.1}$$

where g is a vector made up of the Fourier amplitudes of G regarded as a function of angle variables. If the equation (1.1) is truncated so as to include only a finite number of Fourier modes, then one may show both theoretically and numerically that it can be solved by simple iteration (starting with $g = 0$) under appropriate restrictions on the Hamiltonian and the region of phase space considered. For faster convergence in a wider domain, one can instead solve (1.1) by Newton's method. Newton's method provides a deeper view of the problem, since it entails an examination of the Jacobian matrix of $g - A(g)$ and its singularities. In the present exposition we emphasize Newton's method, but we observe that plain iteration also works very well in many cases of interest.

We emphasize that computer programs to implement our method are short and simple, and the computation time required is modest. Unlike high-order perturbation theory there is no need for symbolic manipulation programs. Accuracy can be enhanced merely by increasing the number of Fourier modes, at least up to a certain limit. In a region of parameter space far from the transition to chaos, our experiments are consistent with the hypothesis that one can include any number of modes, to achieve arbitrary accuracy. To expand the mode set, we use a solution achieved with one set of modes as the starting point for an iteration involving a larger set. By contrast, in a region very close to transition it is difficult to extend the mode set beyond a certain range. Beyond that range convergence suffers, and accuracy increases slowly if at all. Nevertheless, one can include enough modes to get impressive accuracy, as judged by comparison with numerical integration of Hamilton's equations.

Faced with a limitation on the size of the mode set, which might in principle apply in parameter regions far from transition as well as near, how can one

achieve arbitrary accuracy? The answer we propose is to embed our algorithm in a KAM iteration, which entails a sequence of canonical transformations chosen to make the non-integrable part of the Hamiltonian progressively smaller.[5]

Each step of the KAM iteration invokes an approximate solution of the Hamilton-Jacobi equation, the approximation being taken to lowest order in the perturbation that survives at that step. Each step involves a finite number of Fourier modes of the generating function, the number being increased at the next step. Our proposal is to modify the KAM algorithm by replacing the lowest order solution of the Hamilton-Jacobi equation in a finite mode space by a precise Newton solution of that equation. The standard KAM method is retrieved by taking just the lowest order solution of (1.1) by plain iteration, namely $g = A(0)$, for the generator of a KAM step. Since we get a vastly better solution of the Hamilton-Jacobi equation by solving the truncated (1.1) precisely, we can be sure that the modified KAM sequence will converge under the original conditions of the KAM proof, but better yet we can expect a great acceleration in the rate of convergence, most probably under weaker conditions. In order to test this idea in practice, we have carried out two KAM steps numerically for a simple nonintegrable Hamiltonian. Using an accurate solution of the Hamilton-Jacobi equation at the first step, and a lowest-order solution at the second, we find in general very small corrections to the results of the first step.

A great practical advantage of the emerging scheme over ordinary perturbation theory or standard KAM theory is that we usually have to deal (modulo tiny corrections) with only one canonical transformation. Consequently, the invariant torus is given explicitly as a Fourier sum, and one avoids the usual awkward step of solving a chain of nonlinear equations for successive canonical transformations.

It goes without saying that any approximation based on a finite mode set cannot reveal the full structure of phase space for a non-integrable system. In reality the invariant tori are defined only on a Cantor set in frequency space, while the finite mode approximations are locally smooth functions of frequency on open

sets. Smooth interpolations of the invariant tori (physical only on a Cantor set!) have been constructed by Pöschel[17] and Chierchia and Gallavotti;[18] undoubtedly the finite mode model follows such interpolations to some approximation.

An essential feature of the KAM iteration is the control of small divisors by varying actions in the course of iteration, so as to hold small divisors constant. In application of the Newton method to (1.1) the issue of small divisors is altered, since zeros of the determinant of the Jacobian are crucial, rather than zeros of the traditional simple divisors $\omega \cdot \mathbf{m}$. Consequently, a somewhat different technique for controlling frequencies is called for. There are different possibilities which we outline. For the numerical calculations we choose a method which is probably not optimum regarding rate of convergence, but it works well enough for the present purposes. The choice of an optimum method is still under study.

In Section 2 we recall the Hamilton-Jacobi formalism, and give two different statements of the Hamilton-Jacobi equation in terms of Fourier amplitudes of the generating function. In Section 3 we derive explicit formulae for Newton's method applied to the truncated Hamilton-Jacobi equation. Section 3 includes a discussion of methods for iterative determination of the action parameter so that the final frequencies will have the desired non-resonant values. Section 4 is concerned with successive canonical transformations.

In Section 5 we begin numerical considerations by reviewing some issues regarding numerical Fourier transforms and other points of computer programming for the task at hand. Section 6 is concerned with testing the method in a soluble example, namely the 4^{th} order isolated resonance model in $1\frac{1}{2}$ degrees of freedom, which is locally integrable. We find a large domain of convergence for Newton's method in this example, and report the interesting discovery that the domain can be made still larger by artificially introducing an initial frequency shift in the Jacobian matrix. In Section 7 we give numerical results for a non-integrable example, the two-resonance model in $1\frac{1}{2}$ degrees of freedom. This is mathematically similar to the problem of a particle interacting with two waves.[19]

We take the two resonance frequencies to be on either side of the golden mean and calculate the KAM surface with frequency at the golden mean as a function of perturbation strengths. We then explore convergence of the Newton iteration for a sequence of perturbation strengths leading up to the critical value at which the KAM surface breaks ('transition to chaos'). We propose a criterion for the transition, intrinsic to the Hamilton-Jacobi method, and check it against numerical integration of Hamilton's equations. Section 7 ends with consideration of a second canonical transformation.

In Section 8 we show how Greene's residue criterion fits into the Hamilton-Jacobi formalism.[20,21] Although the discussion digresses from our main theme, it may help to relate our viewpoint to current ideas. In Section 9 we comment on the outlook for further work and briefly review related studies by other authors.

2. THE HAMILTON-JACOBI EQUATION

We discuss a system with d degrees of freedom having a time-dependent Hamiltonian periodic in time (in colloquial nomenclature, a system of '$d + 1/2$ degrees of freedom'). In the angle-action variables of the unperturbed problem $(\mathbf{\Phi}, \mathbf{J})$, the Hamiltonian is written as

$$H(\mathbf{\Phi}, \mathbf{J}, \theta) = H_0(\mathbf{J}) + V(\mathbf{\Phi}, \mathbf{J}, \theta) \quad , \tag{2.1}$$

where bold-faced quantities are d-component vectors and θ is the time. The perturbation V is periodic in θ with period 2π,

$$V(\mathbf{\Phi}, \mathbf{J}, \theta) = V(\mathbf{\Phi}, \mathbf{J}, \theta + 2\pi) \quad . \tag{2.2}$$

In a circular accelerator, θ would be the azimuthal position of a particle rather than the actual time. In accord with the definition of angle-action variables, V is also periodic in each component of $\mathbf{\Phi}$ with period 2π. It is trivial to specialize the following work to the case of a time-independent Hamiltonian.

We seek a canonical transformation $(\mathbf{\Phi}, \mathbf{J}) \mapsto (\mathbf{\Psi}, \mathbf{K})$ in the form

$$\mathbf{J} = \mathbf{K} + G_{\mathbf{\Phi}}(\mathbf{\Phi}, \mathbf{K}, \theta) \quad , \tag{2.3}$$

$$\mathbf{\Psi} = \mathbf{\Phi} + G_{\mathbf{K}}(\mathbf{\Phi}, \mathbf{K}, \theta) \quad , \tag{2.4}$$

such that the Hamiltonian becomes a function of $\mathbf{K}$ alone. Subscripts denote partial differentiation. The Hamilton-Jacobi equation to determine G is the requirement that the new Hamiltonian H_1 indeed depend only on $\mathbf{K}$; namely,

$$H_0(\mathbf{K} + G_{\mathbf{\Phi}}) + V(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) + G_\theta = H_1(\mathbf{K}) \quad . \tag{2.5}$$

By Hamilton's equations in the new variables, $\mathbf{K}$ will be invariant, and $\mathbf{\Psi}$ will advance linearly with the time:

$$\begin{aligned} \mathbf{K} &= \text{constant} \quad , \\ \mathbf{\Psi} &= \boldsymbol{\omega}\theta + \mathbf{\Psi_0} \quad , \end{aligned} \tag{2.6}$$

where

$$\boldsymbol{\omega} = \frac{\partial H_1}{\partial \mathbf{K}} \tag{2.7}$$

is the perturbed frequency. Once a G satisfying (2.5) is known, the invariant surfaces $\mathbf{J} = \mathbf{J}(\mathbf{\Phi}, \theta)$ are given explicitly by (2.3), with $\mathbf{K}$ as a parameter. On the other hand, to find the orbit $(\mathbf{J}(\theta), \mathbf{\Phi}(\theta))$ for a given initial condition one must solve the nonlinear equation (2.4) for

$$\mathbf{\Phi} = \mathbf{\Phi}(\mathbf{\Psi}, \mathbf{K}, \theta) \quad . \tag{2.8}$$

Then the orbit is obtained through (2.8), (2.6), and (2.3), with initial conditions determined by the constants $(\mathbf{K}, \mathbf{\Psi_0})$. A method to solve (2.4) is given at the end of this section.

By (2.3) and the physical meaning of angle-action variables it is clear that we want a G which is periodic with period 2π, in each component of $\mathbf{\Phi}$. Also, because the Hamiltonian is periodic in θ, the section of the invariant surface at $\theta = \theta_0$ must be identical to that for $\theta = \theta_0 + 2\pi n$. Therefore we look for a G that is periodic in θ as well as in $\mathbf{\Phi}$. (There are also solutions of the Hamilton-Jacobi equation that are not periodic in θ, which are useful for finding time evolution over a finite interval, but those do not concern us here.) In view of the periodicity a Fourier analysis of G is appropriate:

$$G(\mathbf{\Phi}, \mathbf{K}, \theta) = \sum_{\mathbf{m},n} g_{\mathbf{m}n}(\mathbf{K}) e^{i(\mathbf{m}\cdot\mathbf{\Phi} - n\theta)} \quad , \tag{2.9}$$

where n and the components of $\mathbf{m}$ run over all integers. To state the Hamilton-Jacobi equation in terms of the coefficients $g_{\mathbf{m}n}$ we first subtract and add terms to write it as follows:

$$\begin{aligned} \omega_0(\mathbf{K}) \cdot G_{\mathbf{\Phi}} + G_\theta = &- [H_0(\mathbf{K} + G_{\mathbf{\Phi}}) - H_0(\mathbf{K}) - \omega_0(\mathbf{K}) \cdot G_{\mathbf{\Phi}} + V(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta)] \\ &+ [H_1(\mathbf{K}) - H_0(\mathbf{K})] \quad . \end{aligned} \tag{2.10}$$

The components of ω_0 are the unperturbed frequencies,

$$\omega_0 = \frac{\partial H_0}{\partial \mathbf{K}} \quad . \tag{2.11}$$

Let us now take the Fourier transform of (2.10) for $\mathbf{m} \neq 0$. The final term on the right side, $H_1 - H_0$, is independent of angles and makes no contribution for nonzero $\mathbf{m}$. Hence

$$\begin{aligned} i(\omega_0 \cdot \mathbf{m} - n) g_{\mathbf{m}n} = &- \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \, e^{-i(\mathbf{m}\cdot\mathbf{\Phi} - n\theta)} \times \\ &[H_0(\mathbf{K} + G_{\mathbf{\Phi}}) - H_0(\mathbf{K}) - \omega_0 \cdot G_{\mathbf{\Phi}} + V(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta)] \quad , \quad \mathbf{m} \neq 0 \end{aligned} \tag{2.12}$$

where

$$G_{\mathbf{\Phi}} = \sum_{\mathbf{m},n} i\mathbf{m} g_{\mathbf{m}n} e^{i(\mathbf{m}\cdot\mathbf{\Phi} - n\theta)} \quad . \tag{2.13}$$

Equations (2.12) and (2.13) form our object of study. They constitute a closed

system to determine $g_{\mathbf{m}n}$, $\mathbf{m} \neq 0$, and hence $G_{\mathbf{\Phi}}$, with $\mathbf{K}$ as a parameter. Once $G_{\mathbf{\Phi}}$ is known, H_1 is obtained by averaging (2.5) over angles, since G_θ has zero average:

$$H_1(\mathbf{K}) = \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \, H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) \quad . \tag{2.14}$$

Next we can take the Fourier transform of (2.5) for $\mathbf{m} = 0$ but $n \neq 0$ to determine $g_{0\,n}$:

$$ing_{0\,n} = \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \, e^{in\theta} \, H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) \ , \ n \neq 0 \ . \tag{2.15}$$

The one remaining amplitude g_{00} may be put equal to zero, since any other value would merely shift $\mathbf{\Psi}$ by a constant. By taking inverse Fourier transforms we see immediately that a solution of (2.12) and (2.13) leads to a solution of the Hamilton-Jacobi equation through application of (2.14), (2.15), and (2.9). The perturbed frequencies are given by (2.7) and (2.14) as

$$\omega = \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \, H_{\mathbf{K}}(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) \cdot (1 + G_{\mathbf{K\Phi}}) \quad . \tag{2.16}$$

After dividing (2.12) by $i(\omega \cdot \mathbf{m} - n)$, we see that (2.12) and (2.13) together may be summarized succinctly as

$$g = A(g) \quad , \tag{2.17}$$

where $g = [g_{\mathbf{m}n}]$ is a vector made up of the Fourier amplitudes of G, and

$$A_{\mathbf{m}n} = \frac{i}{(\omega_0 \cdot \mathbf{m} - n)} \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \, e^{-i(\mathbf{m}\cdot\mathbf{\Phi} - n\theta)} \\ [H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) - H_0(\mathbf{K}) - \omega_0 \cdot G_{\mathbf{\Phi}}] \ , \ \mathbf{m} \neq 0 \ . \tag{2.18}$$

Solution of (2.17) by iteration is discussed in the following section.

To specialize the equations to the case of a conservative system we merely drop all reference to n and θ : put $n = 0$ where it occurs algebraically, delete the θ integration and its associated factor $(2\pi)^{-1}$, and ignore (2.15).

The Fourier development of G with respect to θ may be inappropriate if the Hamiltonian is not a smooth function of θ. An extreme example is the Hamiltonian of the 'standard map',

$$H(\phi, J, \theta) = \frac{1}{2}J^2 - \frac{k}{2\pi}\cos\phi\,\delta_P(\theta) \tag{2.19}$$

where k is a constant and δ_P is the periodic delta function with period 2π. In this case the Fourier series of H with respect to θ does not converge, and the equations stated above would need special interpretation. Another important example is a model for accelerators in which the θ dependence of magnetic fields producing nonlinear forces is conveniently and accurately represented as a sequence of square steps. The corresponding Hamiltonian has the form

$$H(\mathbf{\Phi}, \mathbf{J}, \theta) = \boldsymbol{\nu}\cdot\mathbf{J} + U(\mathbf{\Phi}, \mathbf{J})f(\theta) \quad , \tag{2.20}$$

where $\boldsymbol{\nu}$ is constant and $f(\theta)$ is zero except over short intervals, centered at nonlinear magnets, on which it is constant. The regions of nonzero f represent a very small fraction of the full domain $[0, 2\pi]$, so that f has almost a delta-function character.

To handle such cases we can avoid Fourier developments in θ, while retaining that for the $\mathbf{\Phi}$ dependence, by using a closed expression for the Green function of the operator

$$i\boldsymbol{\omega}\cdot\mathbf{m} + \frac{\partial}{\partial\theta} \quad . \tag{2.21}$$

This operator appears in (2.10) after a Fourier transform with respect to $\mathbf{\Phi}$ alone:

$$\left(i\boldsymbol{\omega}_0\cdot\mathbf{m} + \frac{\partial}{\partial\theta}\right) g_{\mathbf{m}}(\theta) = -\frac{1}{(2\pi)^d}\int_0^{2\pi}\cdots\int_0^{2\pi} d\mathbf{\Phi}\, e^{-i\mathbf{m}\cdot\mathbf{\Phi}}\times \tag{2.22}$$
$$[H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) - H_0(\mathbf{K}) - \boldsymbol{\omega}_0\cdot G_{\mathbf{\Phi}}] \quad , \quad \mathbf{m} \neq 0 \, ,$$

where

$$G_{\mathbf{\Phi}} = \sum_{\mathbf{m}} i\mathbf{m} g_{\mathbf{m}}(\theta) e^{i\mathbf{m}\cdot\mathbf{\Phi}} \quad . \tag{2.23}$$

The Green function of (2.21) corresponding to periodic boundary conditions is

$$\mathcal{G}(\theta,\theta') = \frac{e^{-i\mu(\theta-\theta')}}{2i\sin(\pi\mu)}[e^{i\pi\mu}h(\theta-\theta') + e^{-i\pi\mu}h(\theta'-\theta)] \quad , \tag{2.24}$$

as is easily verified by direct computation. Here $\mu = \mathbf{m}\cdot\omega_0$ and h is the step function,

$$h(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \quad . \tag{2.25}$$

Application of this Green function in (2.22) gives an integral equation for $g_{\mathbf{m}}(\theta)$:

$$g_{\mathbf{m}}(\theta) = \frac{i}{2\sin(\pi\omega_0\cdot\mathbf{m})} \int_0^{2\pi} d\theta'\, e^{-i\mathbf{m}\cdot\omega_0[\theta-\theta'+\pi\,\mathrm{sgn}(\theta'-\theta)]} \times$$
$$\frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} e^{-i\mathbf{m}\cdot\mathbf{\Phi}} [H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}(\theta'), \theta') - H_0(\mathbf{K}) - \omega_0\cdot G_{\mathbf{\Phi}}(\theta')] \quad . \tag{2.26}$$

Here sgn is the *signum* function,

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases} \quad . \tag{2.27}$$

For the Hamiltonian (2.20) the final factor of the integrand in square brackets reduces to

$$U(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}(\theta'))f(\theta') \quad . \tag{2.28}$$

Consequently, the unknowns of (2.26) are the $g_{\mathbf{m}}(\theta)$ restricted to the support of $f(\theta)$. Once those restricted functions are known, their extensions to all θ are determined automatically by the right hand side of (2.26). Over any interval on which $f(\theta)$ is zero, $g_{\mathbf{m}}(\theta)$ is equal to a constant times $\exp(-i\mathbf{m}\cdot\omega_0\theta)$.

It is easy to check that a solution of (2.26) is necessarily continuous (if the Hamiltonian is piecewise continuous) and periodic in θ with period 2π. Derivatives of a solution are not necessarily continuous; their smoothness will depend on that of the Hamiltonian. A delta function in the Hamiltonian, as in (2.20), will lead to a discontinuity in $g_m(\theta)$.

Although the invariant tori given by (2.3) are usually the objects of interest, one sometimes requires the individual orbits. To compute the latter one must first obtain $G_{\mathbf{K}}$, say by a method given at the end of the next section, and then solve (2.4) for $\mathbf{\Phi}$. One could solve for $\mathbf{\Phi}$ by Newton's method, but there may also be interest in the following method which gives an explicit formula in terms of integrals that can be evaluated numerically. We expand $G_{\mathbf{K}}$ in a Fourier series in $\mathbf{\Psi}$, rather than $\mathbf{\Phi}$:

$$G_{\mathbf{K}}(\mathbf{\Phi},\mathbf{K},\theta) = \sum_{\mathbf{m}} \mathbf{\Gamma}_{\mathbf{m}}(\mathbf{K},\theta) e^{i\mathbf{m}\cdot\mathbf{\Psi}} \quad , \tag{2.29}$$

where

$$\mathbf{\Gamma}_{\mathbf{m}}(\mathbf{K},\theta) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Psi}\, G_{\mathbf{K}}(\mathbf{\Phi},\mathbf{K},\theta)\, e^{-i\mathbf{m}\cdot\mathbf{\Psi}} \quad . \tag{2.30}$$

Since $G_{\mathbf{K}}$ is periodic, (2.4) implies that when ϕ_i changes by 2π, so does ψ_i, while all other components ψ_j are unaffected. Let us suppose that the Jacobian matrix $\partial\mathbf{\Psi}/\partial\mathbf{\Phi} = [\partial\psi_i/\partial\phi_j]$ is nonsingular, so that ϕ_i in turn changes by 2π when ψ_i does so, while all other ϕ_j are constant. We can then change integration variables in (2.30) :

$$\begin{aligned} \mathbf{\Gamma}_{\mathbf{m}} &= \frac{1}{(2\pi)^d} \int_{\phi_1(0)}^{\phi_1(0)+2\pi} \cdots \int_{\phi_d(0)}^{\phi_d(0)+2\pi} d\mathbf{\Phi}\, \det(\partial\mathbf{\Psi}/\partial\mathbf{\Phi})\, G_{\mathbf{K}}\, e^{-i\mathbf{m}\cdot\mathbf{\Psi}} \\ &= \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi}\, \det(\partial\mathbf{\Psi}/\partial\mathbf{\Phi})\, G_{\mathbf{K}}\, e^{-i\mathbf{m}\cdot\mathbf{\Psi}} \quad . \end{aligned} \tag{2.31}$$

The point of this step is that $\mathbf{\Psi}$ is available as an explicit periodic function of

$\mathbf{\Phi}$ through (2.4), and the integral (2.31) can then be evaluated without knowledge of the inverse function $\mathbf{\Phi}(\mathbf{\Psi})$. Thus the desired explicit representation of $\mathbf{\Phi}$ is

$$\mathbf{\Phi} = \mathbf{\Psi} + \sum_{\mathbf{m}} \mathbf{\Gamma}_{\mathbf{m}}(\mathbf{K}, \theta) e^{i\mathbf{m}\cdot\mathbf{\Psi}} \quad , \tag{2.32}$$

where

$$\mathbf{\Gamma}_{\mathbf{m}} = \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} \, \det(1 + G_{\mathbf{\Phi K}}) \, e^{-i\mathbf{m}\cdot(\mathbf{\Phi} + G_{\mathbf{K}})} \, G_{\mathbf{K}} \quad . \tag{2.33}$$

3. ITERATIVE SOLUTION OF THE HAMILTON-JACOBI EQUATION

We now write equations for iterative solution of (2.17). It should be borne in mind that the equations are intended to apply to a reduced version of (2.17) having only a finite number of Fourier modes, even though the truncation will not be acknowledged in the notation of this section. The details of the truncation are covered in Section 4.

The operator has been defined so that (2.17) can be solved by simple iteration,

$$\begin{aligned} g^{(n+1)} &= A(g^{(n)}) \quad , \quad g^{(0)} = 0 \quad , \\ g^{(n)} &\to g \quad , \quad n \to \infty \quad , \end{aligned} \tag{3.1}$$

provided that the Hamiltonian is suitably restricted, and $\omega_0 \cdot \mathbf{m} - n \neq 0$ on the set of modes $(\mathbf{m}, n)$ included. This follows easily from the contraction mapping principle, as is shown in Appendix A. The proof requires weak nonlinearity, a condition which is specified precisely in terms of V, $V_{\mathbf{K}}$, and H_0''. Of course, the bounds on Hamiltonian parameters depend on the values of the divisor $\omega_0 \cdot \mathbf{m} - n$, and thus on $\mathbf{K}$. As is usual in such matters, we find that the restrictions on H of the existence proof are excessively pessimistic. Numerically we find convergence under much weaker conditions.

To see clearly the structure of $A(g)$, let us subtract and add $V(\Phi, \mathbf{K}, \theta)$ in the integrand of (2.18) so as to obtain

$$A_{\mathbf{m}n}(g) = g^{(1)}_{\mathbf{m}n} + \frac{i}{\omega_0 \cdot \mathbf{m} - n} \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\Phi d\theta \, e^{-i(\mathbf{m}\cdot\Phi - n\theta)} \times$$
$$[H_0(\mathbf{K} + G_\Phi) - H_0(\mathbf{K}) - \omega_0(\mathbf{K}) \cdot G_\Phi + V(\Phi, \mathbf{K} + G_\Phi, \theta) - V(\Phi, \mathbf{K}, \theta)] \quad , \tag{3.2}$$

where $g^{(1)}$ is the familiar result for g in lowest order perturbation theory,

$$g^{(1)}_{\mathbf{m}n} = \frac{i}{\omega_0 \cdot \mathbf{m} - n} \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\Phi d\theta \, e^{-i(\mathbf{m}\cdot\Phi - n\theta)} V(\Phi, \mathbf{K}, \theta) \quad . \tag{3.3}$$

Equivalently, $g^{(1)} = A(0)$ is the first iterate of the sequence defined in (3.1). By Taylor's theorem the bracketed expression of (3.2) is of order V^2, since G itself is of order V, according to lowest order perturbation theory. This quadratic dependence on a quantity which can be made small is responsible for the contractive property of A.

Defining

$$F(g)_{\mathbf{m}n} = i(\omega_0 \cdot \mathbf{m} - n)(g - A(g))_{\mathbf{m}n} \quad , \tag{3.4}$$

we also consider the Newton iteration

$$F_g(g^{(p)})(g^{(p+1)} - g^{(p)}) + F(g^{(p)}) = 0 \quad , \quad g^{(0)} = 0 \quad , \tag{3.5}$$

where F_g is the Jacobian matrix (Fréchet derivative) of F:

$$(F_g)_{\mathbf{m}n,\mathbf{m}'n'} = \frac{\partial F_{\mathbf{m}n}}{\partial g_{\mathbf{m}'n'}} \quad . \tag{3.6}$$

At each iteration one solves the linear system (3.5) for $\delta g = g^{(p+1)} - g^{(p)}$; then $g^{(p+1)}$ is computed as $g^{(p+1)} = \delta g + g^{(p)}$.

In the Newton method the first iterate $g^{(1)}$ already involves terms of all orders in V, in contrast to the first iterate (3.3) of plain iteration. Of course one is not compelled to begin either iteration, (3.1) or (3.5), with $g^{(0)} = 0$ as we have indicated, but the choice usually proves to be quite adequate in practice. However, in Section 7 we encounter cases in which a different choice of $g^{(0)}$ is required.

The amplitudes g_{mn} are not all independent, due to the requirement that G be real. In fact,

$$g_{mn} = g^*_{-m,-n} \quad . \tag{3.7}$$

To keep track of independent amplitudes one could use sine and cosine series with real amplitudes, but the bookkeeping proves to be easier if one retains exponential series while removing redundant complex amplitudes. For instance, with $1\frac{1}{2}$ degrees of freedom (2.13) can be rewritten with the help of (3.7) as

$$G_\phi = 2\mathrm{Re} \sum_{n=-\infty}^{\infty} \sum_{m=1}^{\infty} img_{mn}\, e^{i(m\phi - n\theta)} \quad . \tag{3.8}$$

Since the operator $A(g)$ depends on g only through G_ϕ, (2.17) can be considered as a system to determine just the amplitudes that appear in (3.8), namely,

$$g_{mn} \quad , \quad m \geq 1 \quad , \quad -\infty < n < \infty \quad . \tag{3.9}$$

To give an explicit expression for the Jacobian F_g, let us specialize to $1\frac{1}{2}$ degrees of freedom for notational convenience, and make use of (3.8). By formal differentiation, F_g applied to an arbitrary vector h is seen to be as follows:

$$\begin{aligned}[F_g(g)h]_{mn} =& i(m\omega_0(K) - n)h_{mn} + \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta\, e^{-i(m\phi - n\theta)} \times \\ & [H_K(\phi, K + G_\phi, \theta) - \omega_0(K)] \times 2\mathrm{Re} \sum_{n'} \sum_{m' \geq 1} im' h_{m'n'}\, e^{i(m'\phi - n'\theta)} \quad .\end{aligned} \tag{3.10}$$

Because of the 'real part' instruction in the final factor, this is not a linear function

of the complex variable $h_{m'n'}$, although it is linear in the real and imaginary parts of $h_{m'n'}$ separately. Thus to solve (3.5) as a set of linear equations one must regard it as a real system to be solved for the real vectors

$$\begin{bmatrix} \mathrm{Re}(g^{(p+1)} - g^{(p)}) \\ \mathrm{Im}(g^{(p+1)} - g^{(p)}) \end{bmatrix} \quad . \tag{3.11}$$

Assembling the real and imaginary parts of (3.5) to make such a system, and reversing the order of integration and summation in (3.10), we find that (3.5) has the form

$$\begin{bmatrix} \Delta + \mathcal{L}_{rr} & \mathcal{L}_{ri} \\ \mathcal{L}_{ir} & \Delta + \mathcal{L}_{ii} \end{bmatrix} \begin{bmatrix} \mathrm{Re}(X) \\ \mathrm{Im}(X) \end{bmatrix} = \begin{bmatrix} -\mathrm{Im}(F) \\ \mathrm{Re}(F) \end{bmatrix} \quad , \tag{3.12}$$

where $X_{mn} = m(g_{mn}^{(p+1)} - g_{mn}^{(p)})$ and the sums on matrix indices are understood to include only the modes indicated in (3.9). The matrix elements are given by

$$\begin{aligned} \Delta(mn|m'n') &= (\omega_0 - n/m)\delta_{mm'}\delta_{nn'} \quad , \\ \mathcal{L}_{rr}(mn|m'n') &= \mathrm{Re}[L(m-m', n-n') - L(m+m', n+n')] \quad , \\ \mathcal{L}_{ri}(mn|m'n') &= -\mathrm{Im}[L(m-m', n-n') + L(m+m', n+n')] \quad , \\ \mathcal{L}_{ir}(mn|m'n') &= \mathrm{Im}[L(m-m', n-n') - L(m+m', n+n')] \quad , \\ \mathcal{L}_{ii}(mn|m'n') &= \mathrm{Re}[L(m-m', n-n') + L(m+m', n+n')] \quad , \end{aligned} \tag{3.13}$$

where

$$L(\mu, \lambda) = \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta \, e^{-i(\mu\phi - \lambda\theta)} [H_K(\phi, K + G_\phi, \theta) - \omega_0(K)] \quad . \tag{3.14}$$

Thus a Newton iteration requires only one more Fourier transform than a plain iteration, namely (3.14) which determines the coefficient matrix of the linear

equations (3.12). Furthermore, there is the computational advantage that the coefficient matrix is symmetric; symmetry follows from the property

$$L(\mu,\lambda) = L(-\mu,-\lambda)^* \quad . \tag{3.15}$$

A notable feature of the Newton iteration is that individual small divisors $\omega_0 - n/m$ do not occur. This is in sharp contrast to perturbation theory (including KAM theory) and the simple iteration (3.1). It does not mean that the famous small-divisor problem has gone away completely, however, since the Jacobian $F_g(g)$ might have singular points as a function of K. The latter correspond to zeros of the determinant D of system (3.12),

$$D(g;K) = \det(\Delta + \mathcal{L}) \quad ,$$

$$M = \begin{bmatrix} \mathcal{L}_{rr} & \mathcal{L}_{ri} \\ \mathcal{L}_{ir} & \mathcal{L}_{ii} \end{bmatrix} \quad . \tag{3.16}$$

The determinant $D(g^{(p)};K)$ appears as a divisor in iterate $g^{(p+1)}$.

With any finite number of modes it should be possible to choose K so as to avoid zeros of D, but in principle it might happen that K would have to be changed at each iteration. In the examples we have treated numerically it was easy to choose a fixed K so that no zero of D was encountered in the course of iteration. We had only to pick K so that the unperturbed frequency $\omega_0(K)$ was not too close to any n/m on the moderately large set of (m,n) permitted. Moreover, it is easy to prove that for sufficiently weak nonlinearity such a choice always leads to a non-vanishing D (see Appendix A).

If a solution of (2.17) is reached without encountering a zero of D, then the frequencies of the corresponding quasi-periodic motion on a torus are given by (2.16). Although ω is typically close to $\omega_0(\mathbf{K})$, one does not know the exact value of ω until the calculation is finished. For some purposes it would be better to

know the final value of ω at the outset. In particular, we should like to fix ω at an appropriate 'non-resonant' value when we study the breakup of invariant tori as perturbation strengths are increased. The condition that ω be non-resonant, which must be satisfied on a KAM surface, is

$$\left|\frac{1}{\omega\cdot\mathbf{m}-n}\right| \leq \beta\left[\sum_i |m_i| + |n|\right]^{\gamma} \quad , \quad \gamma > d \; , \tag{3.17}$$

for all $\mathbf{m}$, n with $\mathbf{m} \neq 0$, for some fixed β and γ. According to (2.16), ω is a function of g, $g_{\mathbf{K}}$, and $\mathbf{K}$. By arranging an iteration in which all three of these arguments vary, we devise a method such that the final frequency is specified at the start.

We continue in $1\frac{1}{2}$ degrees of freedom and restrict the unperturbed Hamiltonian to be quadratic in J:

$$H_0(J) = \nu J + \alpha J^2/2 \; , \quad \alpha \neq 0 \; . \tag{3.18}$$

The generalization to an arbitrary H_0 in $d + 1/2$ degrees of freedom is only a matter of elaborating notation. By analogy to $\alpha \neq 0$, in the generalization one must impose the KAM condition

$$\det\left(\frac{\partial^2 H_0}{\partial J_i \partial J_j}\right) \neq 0 \; , \tag{3.19}$$

in the region of phase space considered. This is a condition for the frequency map $(\mathbf{K} \mapsto \omega_0(\mathbf{K}))$ to be invertible; it must be possible to set the unperturbed frequencies by setting the actions.

Let us define K_0 so that $\omega_0(K_0) = \omega$, where ω is the desired final frequency:

$$\omega = \omega_0(K_0) = \nu + \alpha K_0 \; . \tag{3.20}$$

Since frequency shifts are typically small, it is reasonable to begin an iteration with $K = K_0$. Using (3.18) and (3.20) one can rewrite (2.16), the expression for

ω, as

$$K = K_0 - \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} d\phi d\theta \left[G_\phi G_{K\phi} + \frac{1}{\alpha} V_K(\phi, K + G_\phi, \theta)(1 + G_{K\phi}) \right] \quad . \tag{3.21}$$

To verify (3.21), observe that G_ϕ and $G_{\phi K}$ have zero average over (ϕ, θ). Since $V_K(\phi, K, \theta)$ also has zero average, it is easy to check that $\alpha(K - K_0)$ is second order in the perturbation strength. That makes it reasonable to incorporate Eq. (3.21) into an iterative scheme.

In addition to the vector equation (2.12) for g and K, and the scalar equation (3.21) for g, g_K, and K, we invoke the derivative of (2.12) with respect to K, in order to close the system with a vector equation for g, g_K, and K. The resulting set of equations is the following:

$$F_{1\,mn} = i[m\omega_0(K) - n]g_{mn} + \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta\, e^{-i(m\phi - n\theta)} \\ [\alpha G_\phi^2/2 + V(\phi, K + G_\phi, \theta)] = 0 \quad , \quad m \neq 0 \, , \tag{3.22}$$

$$F_{2\,mn} = i[m\omega_0(K) - n]g_{K,mn} + \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta\, e^{-i(m\phi - n\theta)} \\ [\alpha G_\phi + V_K(\phi, K + G_\phi, \theta)][1 + G_{\phi K}] = 0 \quad , \quad m \neq 0 \, , \tag{3.23}$$

$$F_3 = \alpha(K - K_0) + \frac{1}{(2\pi)^2} \int_0^{2\pi} d\phi \int_0^{2\pi} d\theta \\ [\alpha G_\phi + V_K(\phi, K + G_\phi, \theta)][1 + G_{\phi K}] = 0 \quad . \tag{3.24}$$

Under suitable conditions these equations may be solved by simple iteration. Sufficient conditions for convergence, a bit stronger than those for (2.12) alone, will be presented in a later publication. A solution by Newton iteration is also feasible. Although we expect that such an iteration will afford the best convergence, in the present work we use a method that is somewhat easier to program.

The method may be described as a hybrid of Newton and plain iteration. Starting with $K = K_0$, we solve (3.25) for g by Newton iteration. We substitute g in (3.26), which then becomes a linear equation for g_K. The coefficient matrix of this equation is F_{1g}, the Jacobian already computed in the Newton iteration. After solving (3.27)for g_K, we substitute g, g_K, and K into the right-hand side of (3.21), and call the result a new value for K. We next use g_K to extrapolate g to the new value of K:

$$g + g_K(K - K_0) \ . \tag{3.28}$$

To repeat the cycle we return for another Newton iteration of (3.29) at the new K, with (3.28) as the zeroth iterate. In practice, we find that (3.30) does not have to be solved very accurately before extrapolation to the next K. In fact, we usually do only one or two Newton steps with (3.31) before finding a new g_K and extrapolating.

4. SUCCESSIVE CANONICAL TRANSFORMS AND EXPANSION OF MODE SET IN A K.A.M. FRAMEWORK

The iterative methods of the previous section apply to the Hamilton- Jacobi equation restricted to a finite set B of Fourier modes. For any B, the iterations are guaranteed to converge for sufficiently weak nonlinearities, the latter being defined in Appendix A. Unfortunately, the conditions of Appendix A require that the strength of nonlinearities decrease at a rapid rate as the mode set B is expanded. Although the conditions in question are crude sufficient conditions for convergence, undoubtedly far from being necessary, the situation motivates one to seek a scheme in which an indefinite expansion of B with a fixed Hamiltonian could be justified theoretically.

On the other hand, in our numerical experiments with Newton's method we do not encounter a limitation on the mode set, as long as we begin an iteration with a solution obtained on a smaller set and stay away from the difficult parameter region close to the break-up of the KAM curve. Rigorous analysis of a

procedure to expand B seems to be a hard problem, however, which we have not yet seriously studied.

Lacking a precise knowledge of how the Newton method behaves under expansion of the mode set, we propose to expand the mode set within a KAM framework. Successive canonical transformations give smaller and smaller residual perturbations, which can be handled by successive truncated Hamilton- Jacobi equations on larger and larger mode sets.

If G is the generating function obtained from the first truncated Hamilton-Jacobi equation, then the original Hamiltonian H is transformed by G into

$$H_1(\mathbf{\Psi}, \mathbf{K}, \theta) = H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) + G_\theta. \tag{4.1}$$

Because of the truncation, H_1 is not purely a function of $\mathbf{K}$ as in the ideal situation of Eq. (2.5). The new perturbation V_1 is H_1 minus its average:

$$H_1(\mathbf{\Psi}, \mathbf{K}, \theta) = H_{10}(\mathbf{K}) + V_1(\mathbf{\Psi}, \mathbf{K}, \theta), \tag{4.2}$$

$$H_{10}(\mathbf{K}) = \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Psi} d\theta H_1(\mathbf{\Psi}, \mathbf{K}, \theta). \tag{4.3}$$

A new truncated Hamilton-Jacobi equation, based on (4.2) and a bigger mode set, is used to construct a new generating function $G_1(\mathbf{\Psi}, \mathbf{K}, \theta)$, and so on. To define notation, we write the second canonical transformation and Hamilton-Jacobi equation as

$$\mathbf{K} = \mathbf{K}_1 + G_{1\mathbf{\Psi}} \ , \tag{4.4}$$

$$\mathbf{\Psi}_1 = \mathbf{\Psi} + G_{1\mathbf{K}_1} \ , \tag{4.5}$$

$$H_1(\mathbf{\Psi}, \mathbf{K}_1 + G_{1\mathbf{\Psi}}, \theta) + G_{1\theta} = H_2(\mathbf{K}_1) \ . \tag{4.6}$$

To express H_1 as a function of $\mathbf{\Psi}$ and $\mathbf{K}$, rather then ϕ and $\mathbf{K}$, we use the Fourier method described at the end of Section 2. With

$$H_1(\mathbf{\Psi}, \mathbf{K}, \theta) = \sum_{\mathbf{m},n} H_{1\,\mathbf{m}n}(\mathbf{K}) e^{i(m\cdot\mathbf{\Psi} - n\theta)} \tag{4.7}$$

we have

$$\begin{aligned} H_{1\,\mathbf{m}n}(\mathbf{K}) = \frac{1}{(2\pi)^{d+1}} \int_0^{2\pi} \cdots \int_0^{2\pi} d\mathbf{\Phi} d\theta \det(1 + G_{\phi\mathbf{K}}) e^{-i\mathbf{m}\cdot(\mathbf{\Phi} + G_{\mathbf{K}}) + in\theta} \\ \times [H(\mathbf{\Phi}, \mathbf{K} + G_{\mathbf{\Phi}}, \theta) + G_\theta] \ . \end{aligned} \tag{4.8}$$

In order to define and solve the new Hamilton-Jacobi equation (4.6), the $\mathbf{K}$ dependence of H_1 will be needed, but generally in a much smaller region than was required for the J dependence of H. Numerically, one might solve the old equation for a few values of $\mathbf{K}$, and then use a finite-element basis, such as a spline basis, to interpolate those values. The values of $\mathbf{K}$ should be centered at that $\mathbf{K} = \mathbf{K}_*$ which gives the desired frequency ω by one of the algorithms of the previous section. Comparing (2.16) and (4.3) we see that the frequency function used in those algorithms is actually $\partial H_{10}(\mathbf{K})/\partial \mathbf{K} = \omega_{10}(\mathbf{K})$; (In Section 3 we pretended that this was the true expression for the frequency, but because of the

truncation it was only an approximation!). Just as we started the first Hamilton-Jacobi iteration with $\mathbf{K}$ such that $\omega(\mathbf{K}) = \omega$, it is appropriate to start the second one at a new action $\mathbf{K}_1$ such that $\omega_{10}(\mathbf{K}_1) = \omega$, which is to say at $\mathbf{K}_1 = \mathbf{K}_*$. If the second equation (4.6) is solved only in lowest order perturbation theory, as in our calculation reported in Section 7, then the only value of $\mathbf{K}_1$ needed is $\mathbf{K}_*$.

5. NUMERICAL ALGORITHMS AND SELECTION OF MODES

Successful application of the methods of Sections 3 and 4 requires attention to a few issues of numerical analysis and programming, which we now review. The first requirement for numerical modeling of (2.17) is an integration rule for the Φ, θ integrations of (2.18). We discretize integrals so as to make the Fast Fourier Transform (FFT) applicable. For a function $f(\phi)$ of period 2π we take

$$\begin{aligned} \hat{f} &= \frac{1}{2\pi} \int_0^{2\pi} e^{-im\phi} f(\phi) d\phi \\ &\simeq \frac{1}{J} \sum_{j=0}^{J-1} e^{-im(2\pi j/J)} f(2\pi j/J) \equiv \tilde{f}(m) \ . \end{aligned} \tag{5.1}$$

This apparently naive integration rule, based on equally spaced mesh points with equal weights, is suited to functions f which are well-approximated by trigonometric polynomials.[22] In fact, $\hat{f}(m) = \tilde{f}(m)$ *exactly* for $|m| < J - p$ if $f(\phi)$ is a trigonometric polynomial of degree p; i.e., if

$$f(\phi) = \sum_{n=-p}^{p} \hat{f}(n) e^{in\phi}. \tag{5.2}$$

This statement will be recognized as an elaboration of Nyquist's criterion. To prove this we introduce (5.2) in the right-hand side of (5.1) and find

$$\tilde{f}(m) = \sum_{n=-p}^{p} K(m,n)\hat{f}(n), \tag{5.3}$$

$$\mathbf{K}(m,n) = \frac{1}{J}\sum_{j=0}^{J-1} e^{i(n-m)(2\pi j/J)} \tag{5.4}$$

Evaluation of the geometric sum (5.4) shows that $\mathbf{K}(m,m) = 1$, and $\mathbf{K}(m,n) = 0$ for $m \neq n$ provided that $(n-m)/J$ is not an integer. But the latter ratio will never be a non-zero integer if $|m| < J-p$, since $|n| \leq p$. For an arbitrary $f(\phi)$ one can appraise accuracy experimentally by checking validity of the approximation

$$f(\phi) \simeq \sum_{m \epsilon C} \tilde{f}(m) e^{im\phi}, \tag{5.5}$$

where C is some appropriate subset of the values m less in magnitude than J; for instance, all m with $|m| \leq [J/2]$.

Our finite dimensional model of (2.17) is obtained by approximating the Φ, θ integrals as in (5.1), and by truncating the Fourier series for G_{Φ}. In 1-1/2 degrees of freedom

$$G_{\Phi} = 2Re \sum_{(m,n)\epsilon B} img_{mn} e^{i(m\phi - n\theta)}, \tag{5.6}$$

where B is some finite set with $m \geq 1$. Having chosen B such that its elements satisfy $m \leq M-1$, how should one choose the number of points in the discretization of integrals? Let $\Psi(\phi,\theta)$ denote the bracketed expression in the integrand of (2.18). The assumption that (2.17) can be modeled on the set B is equivalent to saying that Ψ (evaluated near a solution of the truncated equation) is dominated by modes within B. Thus Ψ is a trigonometric polynomial of order $M-1$ in

ϕ and order $N-1$ in θ, plus a remainder, assumed to be small, which contains in general modes of arbitrarily high order. If we take $2M$ and $2N$ mesh points for the ϕ and θ integrals respectively, then the Fourier transform of the aforementioned trigonometric polynomial will be represented exactly for $|m| \leq M$, $|n| \leq N$, thus for all modes in B and more. It therefore seems reasonable to start with $2M$, $2N$ mesh points. We do so, and then assess the amount of change as additional mesh points are included. In our calculations to date, the inclusion of additional mesh points has been necessary in a few cases, but we have never required more than $4M$, $4N$.

In connection with Newton's method, there is a point concerning discretization that we wish to emphasize. One should first specify a discretization of the nonlinear system, and stay with that same discretization when the Jacobian is computed for the Newton linearization. One might be tempted instead to derive the exact Jacobian F_g, as in (3.10), and then discretize it independently, say with a set of mesh points different from that used for F itself. To understand that temptation, note that if the integral $L(\mu, \lambda)$ of (3.14) is approximated as a sum by an FFT with $2M$ and $2N$ points for ϕ and θ, respectively, then only the output of the FFT for $|\mu| \leq N-1$, $|\lambda| \leq N-1$ approximates the integral, according to the view developed above. On the other hand, evaluation of the Jacobian matrix (3.13) requires

$$0 \leq \mu \leq 2(M-1) \quad , \quad -2(N-1) \leq \lambda \leq 2(N-1) \; . \tag{5.7}$$

To approximate $L(\mu, \lambda)$ on the full intervals (5.7) would require about twice as many mesh points. In fact, it would be an error to use twice as many points for L as for F! If we take $2M$, $2N$ points for ϕ, θ in both L and F, and accept the full range (5.7) of μ and λ as provided by the FFT for L, then we in fact get the exact Jacobian of the discretized nonlinear system, even though many of the FFT values for L are not good approximations of the integral (3.10). We are aware of other studies in which inconsistent discretizations of a function and its Jacobian led to poor convergence of the Newton iteration.

In some cases the equations (3.12) are nearly singular and must be approached with care. That is the case near the transition to chaos, as discussed in Section 7. We have solved the equations without difficulty using the program LEQ2S of the IMSL program library,[23] but have had some trouble with less sophisticated routines. LEQ2S uses Bunch's method,[24] which is efficient for real symmetric indefinite matrices. It improves the solution by iteration, if necessary, returning an error message if iteration fails to produce a solution to the working precision. The error message was never received in the examples of Section 7, even though the matrices had normalized determinants as small as 10^{-28}. Also, there was little change in results in passing from single to double precision.

For Newton iteration the selection of the mode set B is a crucial matter. For solution of (2.17) by plain iteration it is not a bad choice to let B consist of all modes up to limits, say $1 \leq m \leq M-1$, $|n| \leq N-1$. Let us call this set B_0. It is a poor choice for Newton's iteration, because on the one hand most of the modes in B_0 have negligible amplitudes, while on the other hand the corresponding Jacobian matrix is typically very large and sparse. A bit of numerical experimentation shows that a mode is negligible if it is not both fairly close to resonance ($\omega_0 m - n$ small) and driven by the perturbation, either directly or through harmonics. For this reason it is appropriate to "let the equation decide" which modes are important. We employ a simple auto-adaptive algorithm as follows. We take at least $2M$, $2N$ mesh points for ϕ, θ integrations throughout the Newton sequence, so that we are able to evaluate $A(g)_{mn}$ of (2.18) for all $(m,n) \in B_0$. After computing a Newton iterate $g^{(p)}$ we evaluate $|mA(g^{(p)})_{mn}/K|$ for all $(m,n) \in B_0$, and identify as "appreciable" those modes for which this quantity exceeds in magnitude some preassigned a. For the next iteration, B consists of just the appreciable modes, in the sense that in (3.5) all vector indices are restricted to those modes. After the $g_{mn}^{(p+1)}$ are obtained from (3.5) for $(m,n) \in B$ they are summed to form a new G_Φ, and hence $A(g^{(p+1)})$. After a few iterations, typically 3 or 4 in the examples we have treated, B converges to a definite set (perhaps with a slight ambiguity concerning the highest mode num-

bers). Further iterations with fixed B serve to refine the solution. After a little experience one can choose the lower bound a reasonably, taking into account limitations due to the choice of M, N and rounding error. To give an example of the smallness of B compared to B_0, we mention runs described in Section 7 with $M = 128$, $N = 64$. In this case B_0 had 16129 elements, while B typically had fewer than 120. Consequently, the computation time for solution of (3.12) was not great; much more time was required for the inevitable Fourier transforms.

6. TESTING THE METHOD IN A SOLUBLE EXAMPLE

For a first view of the convergence properties of our method we examine a simple case in which the exact answer is available in analytic form; namely, the 4^{th}-order single-resonance model in $1\frac{1}{2}$ degrees of freedom. The Hamiltonian is

$$H(\phi, K, \theta) = \nu J + \frac{1}{2}\alpha J^2 + \epsilon J^2 \cos(4\phi - \theta) \quad , \tag{6.1}$$

where ν, α, and ϵ are constants. In this example the Hamilton-Jacobi equation (2.5) can be solved by an Ansatz to the effect that $G(\phi, K, \theta)$ depends on ϕ, θ only in the combination $4\phi - \theta$. Then $G_\theta = -\frac{1}{4}G_\phi$, and the Hamilton-Jacobi equation is

$$(\nu - \frac{1}{4})(K + G_\phi) + \frac{1}{2}\alpha(K + G_\phi)^2 + \epsilon(K + G_\phi)^2 \cos(4\phi - \theta) = H_1(K) - \frac{1}{4}K \quad . \tag{6.2}$$

Hence

$$J = K + G_\phi = \frac{\frac{1}{4} - \nu \pm \left[\left(\frac{1}{4} - \nu\right)^2 + 4E\left(\frac{\alpha}{2} + \epsilon\cos 4\psi\right)\right]^{1/2}}{2\left(\frac{\alpha}{2} + \epsilon\cos 4\psi\right)} \quad , \tag{6.3}$$

where

$$\begin{aligned} E &= H_1(K) - \frac{1}{4}K = \text{constant} \quad , \\ \psi &= 4\phi - \theta \quad . \end{aligned} \tag{6.4}$$

Equation (6.3) gives the family of invariant surfaces $J(\phi, \theta; E)$ parametrized by E, whereas our method based on solving (2.17) will give $J(\phi, \theta)$ parametrized instead

by K. The relation between the two parameters is obtained by integrating (6.3) with respect to ϕ, noting that the integral of G_ϕ is zero by periodicity. Then

$$K(E) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \, J(\phi, \theta; E) \tag{6.5}$$

provided that the integral exists.

We wish to explore completely the dependence of J on parameters ν, α, ϵ, and E, and see to what extent the iterative solution of (2.17) can reproduce the exact J in various regions of parameter space. Let us first introduce parameters more convenient for the purpose, as follows:

$$J = \left[\frac{\frac{1}{4} - \nu}{2\epsilon}\right] \frac{1 \pm [1 + \gamma(a + \cos 4\psi)]^{1/2}}{a + \cos 4\psi} \quad , \tag{6.6}$$

$$a = \frac{\alpha}{2\epsilon} \quad , \tag{6.7}$$

$$\gamma = \frac{4E\epsilon}{(\frac{1}{4} - \nu)^2} \quad . \tag{6.8}$$

Clearly the parameter $|(\frac{1}{4} - \nu)/2\epsilon|$ just determines the scale of J; we shall henceforth take it to be 1. To allow for both signs of $(\frac{1}{4} - \nu)/2\epsilon$, notice that a change in sign of ϵ is equivalent to rotating the figure by $\pi/4$: $J(\psi) \to J(\psi - \pi/4)$. We shall therefore take $(\frac{1}{4} - \nu)/2\epsilon$ to be $+1$, since that choice gives us all curves, modulo rotation and scale change. The special case $\nu = 1/4$ will be treated separately.

In Appendix B we show how the various types of curves evolve as parameters vary. Here we review the results in a series of graphs. We plot the points $(x, y) = (J^{1/2} \cos\psi, J^{1/2} \sin\psi)$; this yields a surface of section at $\theta = 0$ in the ϕ coordinate. In each figure we show separatrices (if any), and a few other typical curves corresponding to various values of γ. Figure 1 is for the case $a = 0$ in which the nonlinear term of H_0 is absent. In this case there is unbounded motion on

invariant curves which extend to infinity, a feature which persists for $0 \leq |a| < 1$. As a varies from 0 to +1, the inner separatrix evolves into a square, and the outer separatrices into straight lines continuing the sides of the square. As a varies from 0 to -1, the four segments of the inner separatrix become more concave inward, while the corners (hyperbolic fixed points) move out to infinity. Figure 2 shows the situation at $a = -0.9$. As a increases through +1, the separatrices 'connect at infinity' and then move in to enclose islands as shown in Figures 3 and 4 for $a = 1.05$ and $a = 20$, respectively. Unbounded curves that were symmetrical about $y = \pm x$ turn into islands, while those symmetrical about the x and y axes turn into bounded curves with 90° rotational invariance. Indeed, within a certain distance of the origin Figs. 1 and 3 are similar in appearance. As a decreases through -1, the separatrices move to infinity, leaving only curves with 90° rotational invariance, as shown in Fig. 5 for $a = -1.01$. As $a \to -\infty$, the curves become concentric circles. For the degenerate case $\nu = 1/4$ see Appendix B.

In Figures 6, 7 and 9 we show results obtained by Newton's method applied to (2.17), beginning iteration with $g = 0$. Each graph is a separatrix obtained by numerical solution of (2.17) using the value of K given by the formulas of Appendix B, (B9) and (B10). We present results for separatrices only, since in this example they are the most difficult cases to compute. For a fixed mode set, convergence and accuracy are much better on curves far from separatrices.

We use the method of mode selection described in the previous section, which leads as expected to a set B including only harmonics of the perturbation: $(m,n) = \lambda(4,1)$, $\lambda = 1, 2, \cdots, 31$. The set B_0 from which this set was selected consists of all (m,n) such that $1 \leq m \leq 127$, $|n| \leq 31$. The numbers of integration points for ϕ, θ integrals were 512, 128 respectively. Calculations on the VAX 8600 were done first in single precision then repeated in double precision (roughly 7 then 16 decimal digits accuracy).

We measure convergence by the normalized residual

$$r = \frac{\|g - A(g)\|}{\|g\|} , \tag{6.9}$$

where double bars indicate the Hermitian norm

$$\|g\| = \left(\sum_{m,n \in B} |g_{mn}|^2 \right)^{1/2} . \tag{6.10}$$

We monitor the condition number of the Jacobian matrix $\Delta + \mathcal{L}$ by computing the normalized determinant D_N.[25] This is defined as the determinant D divided by the product of the Euclidean row lengths of the matrix. We find D_N to be a useful although crude 'condition number' of the Jacobian. A D_N becoming much smaller than 1 in the course of iteration is a possible warning of an impending singular Jacobian.

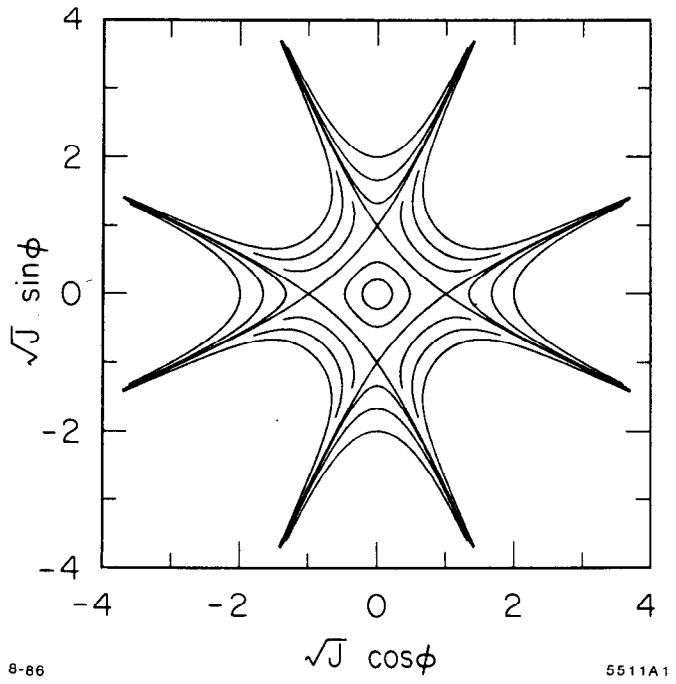
In Figures 6 and 7 we show the cases $a = 0.98$ and $a = -0.9$, respectively; the latter is to be compared with Figure 2. In both cases the residual r was about 10^{-7} (in single precision after 7 iterations) or 10^{-16} (in double precision after 9 iterations). Away from the separatrix similar residuals are found after 3 or 4 iterations. Graphs indistinguishable from Figures 6 and 7 can be made in 5 iterations in single precision, and with about half as many integration points in each variable. Figure 9 shows the case of $a = 20$, to be compared to Fig. 4, which has inner and outer separatrices corresponding to two different values of K; compare (B9). This is an easier case requiring fewer modes and about half as many iterations.

Except near the vertices of the separatrices, where there is a little rounding due to truncation of the Fourier series, the agreement with the exact values is good. For instance at $a = 20$ on the outer separatrix there is agreement to 8 significant figures at $\phi = \pi/4$, 5 figures at $\phi = \pi/8$, and 2 figures at $\phi = 0$. Curves away from separatrices, having continuous derivatives, are more accurate globally.

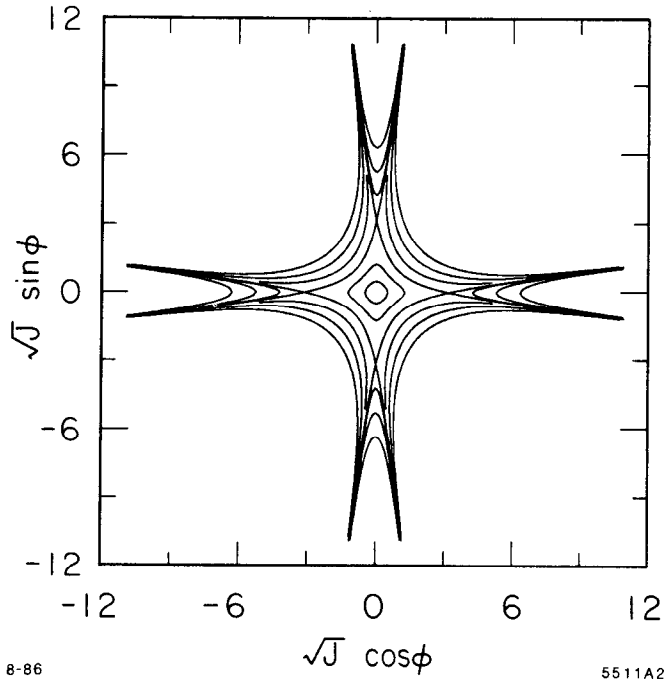
A difficult case is $a = 1.05$ on the outer separatrix (not the inner). Here Newton's method fails to converge if $g = 0$ is the starting point. In fact for $1 < a \leq 1.2$, sequences beginning at $g = 0$ do not converge. One can try stepwise continuation from larger a, but we have found an easier method which provides convergence from $g = 0$. The procedure is to perturb the diagonal of the Jacobian matrix $\Delta + \mathcal{L}$ by a small positive constant δ which is gradually decreased as the iteration proceeds. The initial δ is somewhat smaller than $\omega_0(K) - 1/4$. For $a = 1.05$ we start with $\delta = 0.015$, whereas $\omega_0(K) - 1/4 = 0.052$. Keeping δ at this value for the first three iterations, and then decreasing it gradually to zero, we produce in 9 iterations the curve of Fig. 8, to be compared to Fig. 3. In single precision the residual r is 1.7×10^{-6}.

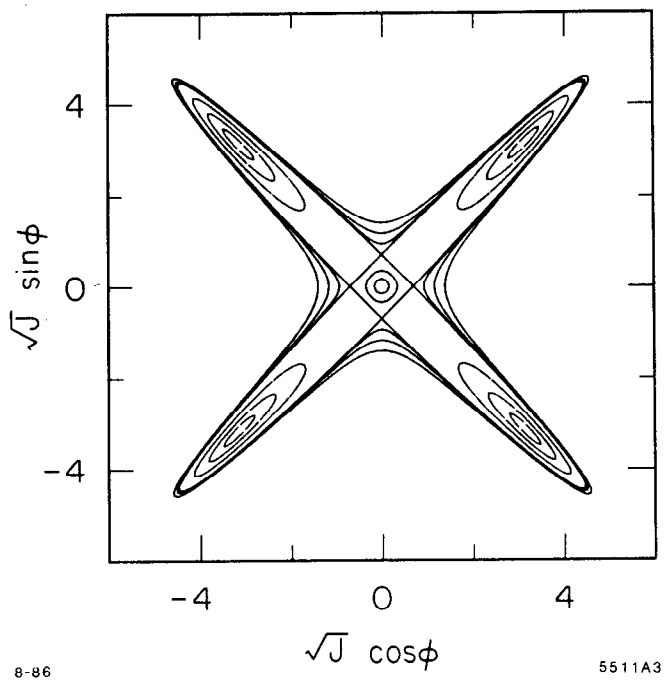
Curves such as those of Fig. 9, which may be regarded as typical in many applications, can be produced by simple iteration of (2.17), in very small computation time. The more difficult cases require Newton's method.

The relative difficulty of computing separatrices seems not to be associated with a very ill-conditioned Jacobian. For instance, D_N is of order 10^{-3} in the calculation of Fig. 9. In comparison to values encountered near the transition to chaos this is not small; see Section 7.

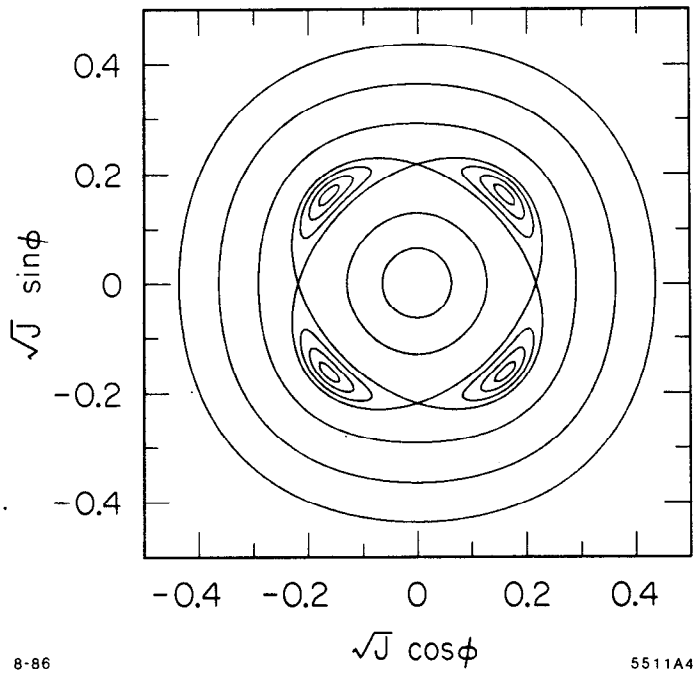

4
2
0
-2
-4
-4
-2
0
2
4
√J sinϕ
√J cosϕ
8-86
5511A1

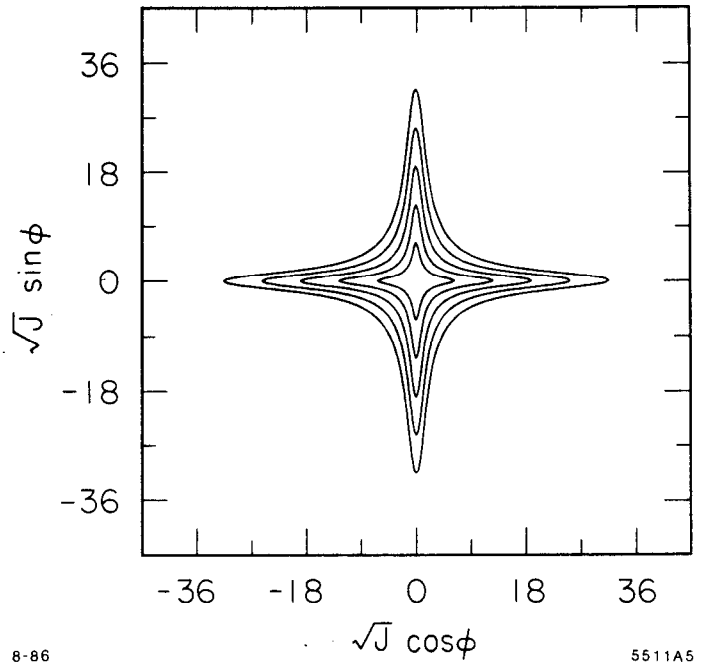
Figures 1-2


12
6
0
-6
-12
-12
-6
0
6
12
√J sinϕ
√J cosϕ
8-86
5511A2

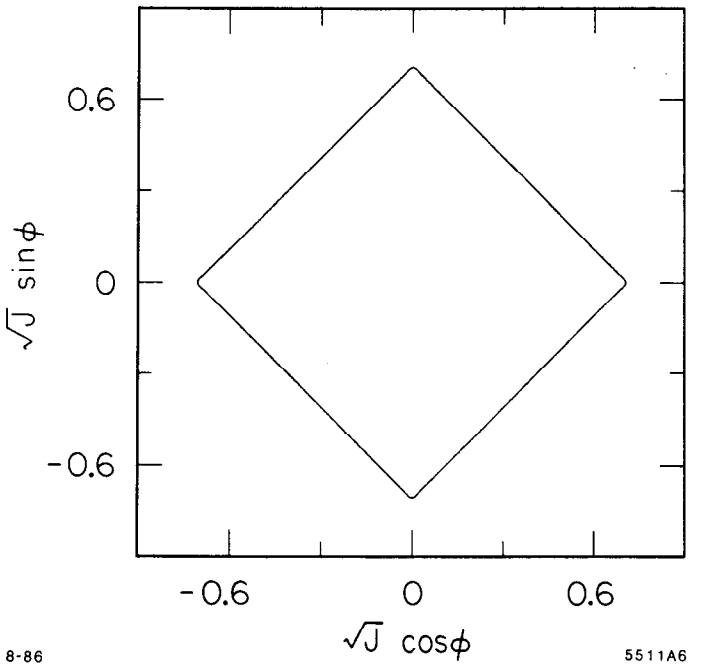

4
0
-4
$\sqrt{J}\sin\phi$
-4
0
4
$\sqrt{J}\cos\phi$
8-86
5511A3

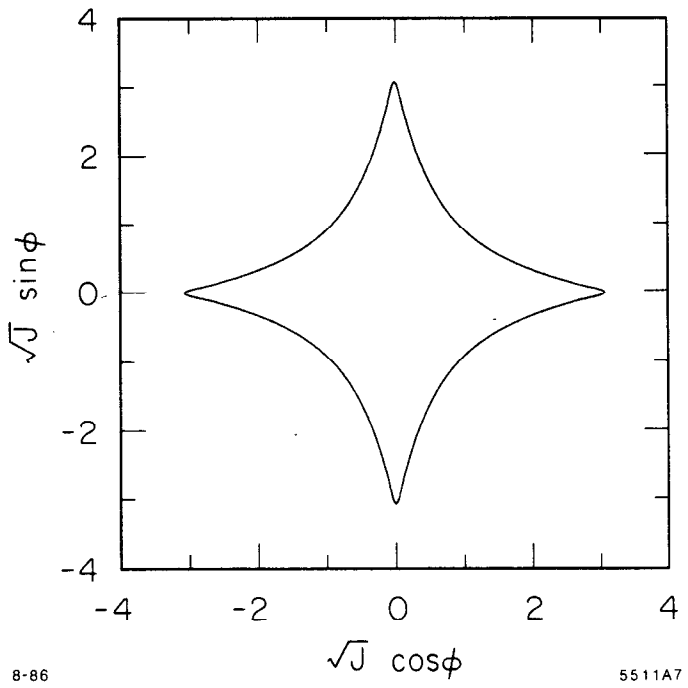
Figures 3-4


0.4
0.2
0
-0.2
-0.4
$\sqrt{J}\sin\phi$
-0.4
-0.2
0
0.2
0.4
$\sqrt{J}\cos\phi$
8-86
5511A4

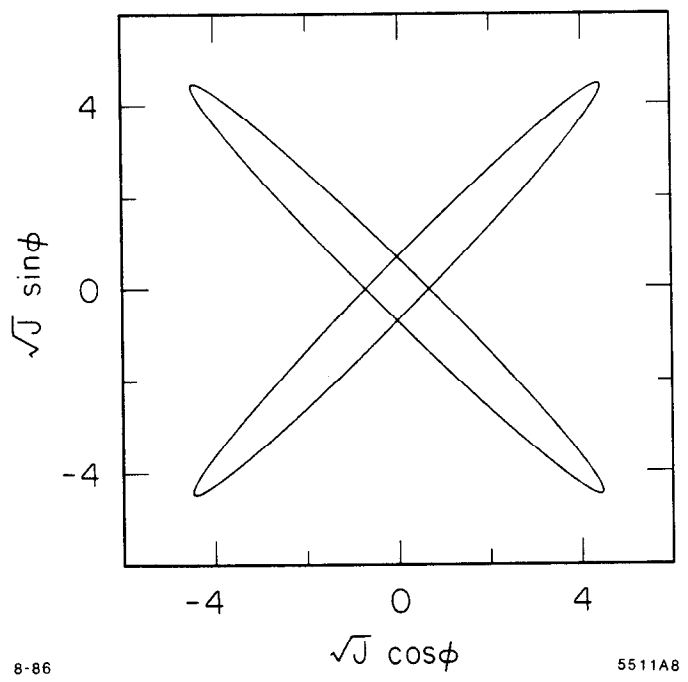

36
18
0
-18
-36
√J sinφ
-36
-18
0
18
36
√J cosφ
8-86
5511A5

Figures 5-6


0.6
0
-0.6
√J sinφ
-0.6
0
0.6
√J cosφ
8-86
5511A6


4
2
0
-2
-4
-4
-2
0
2
4
$\sqrt{J}\sin\phi$
$\sqrt{J}\cos\phi$
8-86
5511A7

Figures 7-8


4
0
-4
-4
0
4
$\sqrt{J}\sin\phi$
$\sqrt{J}\cos\phi$
8-86
5511A8

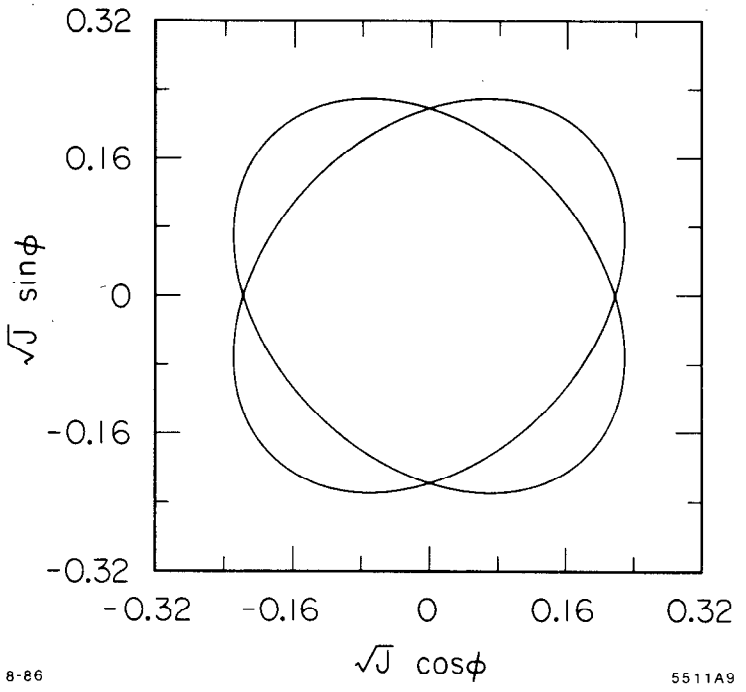


Figure 9

Figures 1-5: 4^{th} order single-resonance model. Typical sections of invariant surfaces at $\theta = 0$. $(\sqrt{J}\cos\phi, \sqrt{J}\sin\phi)$ is plotted from the formula (6.6). The parameter a of (6.7) has the values $a = 0,\ -0.9,\ 1.05,\ 20,\ -1.01$ for Figures 1-5 respectively.

Figures 6-9: 4^{th} order single-resonance model. Separatrices at $\theta = 0$, obtained by solving the Hamilton-Jacobi system (2.17) by Newton's method. The parameter a of (6.6) has the values $a = 0.98,\ -0.9,\ 1.05,\ 20$ for Figures 6-9, respectively. For $a = 1.05$ the Newton method was modified as explained in Section 6.

7. TWO NEIGHBORING RESONANCES AND CHAOTIC BEHAVIOR; A NON-INTEGRABLE EXAMPLE

We now pass from the trivial case of the preceding section to a non-integrable case in which chaotic motion can arise; namely, a two resonance model in 1 1/2 degrees of freedom. The Hamiltonian is

$$H = \nu J + \frac{1}{2}\alpha J^2 + \epsilon_1 J^{5/2}\cos(5\phi - 3\theta) + \epsilon_2 J^2 \cos(8\phi - 5\theta) \ . \qquad (7.1)$$

At small ϵ_1, ϵ_2 we find a KAM curve at a perturbed frequency near

$$\omega_\star = \frac{5^{1/2} - 1}{2} = (\text{golden mean}) - 1 = 0.6180339\cdots \qquad (7.2)$$

and try to investigate its breakup as the ϵ's are increased. The frequency $\omega_\star$ lies between the two resonances, 3/5=.6 and 5/8=.625. For convenience we choose $\nu = 0.5$, $\alpha = 0.1$, so that the corresponding K is near 1. The unperturbed resonances are said to be at $J = J_r$, where $\nu + \alpha J_r = n/m = 3/5,\ 5/8$; thus $J_{r1} = 1,\ J_{r2} = 1.25$.

We use the Newton method to solve (2.17) for the KAM curves, incorporating variations of K during iteration so as to hold the frequency close to the desired value $\omega_\star$. The method described at the end of Section 3 was used to vary K, with a change of K at every iteration.

The half widths of the resonance islands are estimated by

$$\Delta J = 2\left[\frac{\epsilon J_r^{m/2}}{\alpha}\right]^{1/2} \qquad (7.3)$$

for the term $\epsilon J^{m/2}\cos(m\phi - n\theta)$. We show results for a sequence of cases with ϵ's and ΔJ's as follows:

i) $\epsilon_1 = 2\epsilon_2 = 6 \times 10^{-5}$

$\Delta J_1 = 0.049 \quad \Delta J_2 = 0.054$

ii) $\epsilon_1 = 2\epsilon_2 = 8 \times 10^{-5}$

$\Delta J_1 = 0.057 \quad \Delta J_2 = 0.062$

iii) $\epsilon_1 = 2\epsilon_2 = 10^{-4}$

$\Delta J_1 = 0.063 \quad \Delta J_2 = 0.070$

iv) $\epsilon_1 = 2\epsilon_2 = 1.2 \times 10^{-4}$

$\Delta J_1 = 0.069 \quad \Delta J_2 = 0.076$

v) $\epsilon_1 = 2\epsilon_2 = 1.25 \times 10^{-4}$

$\Delta J_1 = 0.070 \quad \Delta J_2 = 0.078$

By the resonance overlap criterion[26] and the magnitudes of the ΔJ's, one can expect these cases to be stochastic or nearly so, for some regions of J in the interval $[J_{r1}, J_{r2}]$, since $J_{r2} - J_{r1} = 0.25$ is comparable to $\Delta J_1 + \Delta J_2$.

We immediately face the question of how to recognize the arrival of stochasticity in the Hamilton-Jacobi formalism. One possibility is that the canonical transform equation (2.4) develops a singularity so that one can no longer solve for $\phi = \phi(\psi, K, \theta)$, even if G exists and is known. That can happen if the Jacobian of (2.4),

$$\frac{\partial \psi}{\partial \phi} = 1 + G_{\phi K}(\phi, K, \theta) \tag{7.4}$$

develops a zero for some (ϕ, θ). Another possibility is that the Jacobian $\Delta + \mathcal{L}$ of the Hamilton-Jacobi system (2.17) becomes singular. We try to monitor the calculation for both of these possibilities, by plotting $\partial\psi/\partial\phi$ and by computing the normalized determinant D_N of the Jacobian as defined in Section 6. Here we actually plot $\partial\psi/\partial\phi$ at $\theta = 0$. Since the function's minimum with respect to ϕ does not vary much with θ, this gives a good indication of the presence or absence of zeros.

To appreciate the meaning of $\partial\psi/\partial\phi$ it is worthwhile to note that

$$\partial\psi/\partial\phi = \partial J/\partial K \ , \tag{7.5}$$

where ψ and J are both regarded as functions of ϕ, K, and θ. Thus the heuristic

picture of $\partial\psi/\partial\phi = 0$ is that two curves that differ infinitesimally in their K values make contact.

In Figures 10 to 13 we show $J(\phi,\theta)$ plotted against $\phi/2\pi$ at $\theta = 0$, for cases (i) to (iv). Figures 14 to 17 show the corresponding graphs of $\partial\psi/\partial\phi$. The buildup of high modes is quicker and more pronounced in $\partial\psi/\partial\phi$ than in J. The anticipated zeros of $\partial\psi/\partial\phi$ seem on the verge of appearance in case (iv), Fig. 17. We find, however, that as the ϵ's are increased the behavior of $\partial\psi/\partial\phi$ becomes increasingly sensitive to the number of modes included in the truncated Hamilton-Jacobi equation. We must therefore treat the question of mode selection before trying to estimate the critical ϵ's for appearance of zeros.

The computations for Figures 10 to 17 were based on mode sets selected by the method of Section 5, with the parameter a taken to be 10^{-10}. Recall that a is a lower bound for allowed values of $|mA(g)_{mn}/K|$; typically the maximum value of the latter is around 5×10^{-3} in these examples. In each case the calculation was done first with 8 Newton iterations on a relatively small mode set B selected from an initial set B_0 consisting of all (m,n) with $1 \leq m \leq 63$, $|n| \leq 31$. The starting point of the Newton sequence was $g = 0$. The result of this iteration was then taken as the starting point for 8 additional Newton iterations on a larger set B selected from a B_0 made up of all (m,n) with $1 \leq m \leq 127$, $|n| \leq 63$. For the ϕ and θ integrations there were 256 and 128 mesh points, respectively, for the first 8 iterations, and 512 and 256 for the second. The calculations were done in double precision, but single precision usually produces graphs that are visually indistinguishable from those shown. Double precision is required to compute the residual perturbation after one canonical transformation and to rule out the possibility of severe rounding error. The results were not materially affected by varying a over the range 10^{-8} to 10^{-12}, but at the smaller values the mode set is needlessly large. In single precision a must not be smaller than 10^{-8}, to avoid inclusion of modes at the level of round-off noise.

In Tables 1 and 2 we give data on the computations for cases (i) and (iv),

respectively; cases (ii) and (iii) have intermediate behavior. For each iterate $g^{(p)}$ we tabulate the number n_B of modes in B, the normalized residual r defined in (6.9), the normalized determinant D_N of the Jacobian matrix defined following (6.10), and $\Delta\omega/\omega$, which is the fractional deviation of the frequency from the desired value ω_*. The frequency was computed by (2.16), with the integral approximated on the same mesh used in the main calculation. At the end of each stage of the calculation we give a final value $(\Delta\omega/\omega)_f$ of the frequency deviation, and the ratio v_1/v, where v is the absolute value of the original Hamiltonian perturbation V averaged over (ϕ, θ), and v_1 is a similar average of the residual perturbation V_1 as defined in (4.2).

We see from Table 1 that the method works very well in case (i). The convergence is rapid, the residual perturbation is smaller than the original perturbation by a factor 6×10^{-8}, and the final frequency has the desired value to machine accuracy. Expansion of the mode set from stage one (40 modes) to stage two (77 modes) served to decrease the residual perturbation substantially. There is, however, no change at the level of visual inspection of the graphs of J and $\partial\psi/\partial\phi$ between stage one and stage two.

In passing through the cases from (i) to (iv) the convergence gradually becomes slower. The situation at case (iv) is shown in Table 2. Here the expansion of the mode set from stage one (40 modes) to stage two (117 modes) gives relatively little decrease in the residual perturbation. It is gratifying that v_1/v is still small compared to 1, even though much bigger than in case (i).

Table 1				
Case *(i)* $\epsilon_1 = 2\epsilon_2 = 6 \times 10^{-5}$				
p	n_B	r	D_N	$\Delta\omega/\omega$
1	2	8.4×10^{-6}	1.	–
2	16	4.5×10^{-5}	.51	1.2×10^{-4}
3	36	6.0×10^{-8}	0.027	-3.9×10^{-6}
4	39	4.0×10^{-12}	0.027	9.9×10^{-8}
5	40	6.9×10^{-14}	0.027	-3.5×10^{-9}
6	40	1.7×10^{-15}	0.027	8.9×10^{-11}
7	40	2.4×10^{-16}	0.027	-2.2×10^{-12}
8	40	2.7×10^{-16}	0.027	5.4×10^{-14}
$v_1/v = 1.1 \times 10^{-5}$, $(\Delta\omega/\omega)_f = -6.7 \times 10^{-16}$				
1	61	2.6×10^{-11}	3.8×10^{-9}	4.1×10^{-11}
2	68	1.5×10^{-10}	3.0×10^{-6}	6.4×10^{-11}
3	73	2.1×10^{-10}	4.6×10^{-10}	-1.4×10^{-12}
4	75	4.0×10^{-13}	6.1×10^{-11}	9.7×10^{-14}
5	76	4.2×10^{-14}	6.0×10^{-11}	-7.7×10^{-14}
6	77	1.7×10^{-14}	5.2×10^{-11}	-5.0×10^{-15}
7	77	$< 3 \times 10^{-39}$	5.2×10^{-11}	-1.9×10^{-15}
8	77	$< 3 \times 10^{-39}$	5.2×10^{-11}	4.5×10^{-17}
$v_1/v = 6.4 \times 10^{-8}$, $(\Delta\omega/\omega)_f < 3 \times 10^{-39}$				

Table 2				
Case *(iv)* $\epsilon_1 = 2\epsilon_2 = 1.2 \times 10^{-4}$				
p	n_B	r	D_N	$\Delta\omega/\omega$
1	2	3.3×10^{-5}	1.0	–
2	17	4.4×10^{-4}	.13	5.1×10^{-4}
3	37	1.4×10^{-5}	1.6×10^{-5}	-4.5×10^{-5}
4	42	1.1×10^{-7}	2.6×10^{-5}	-4.4×10^{-6}
5	43	2.2×10^{-9}	2.7×10^{-5}	1.9×10^{-6}
6	43	1.8×10^{-10}	2.7×10^{-5}	-5.3×10^{-7}
7	43	1.1×10^{-11}	2.7×10^{-5}	1.5×10^{-7}
8	43	2.6×10^{-12}	2.7×10^{-5}	-4.4×10^{-8}
$v_1/v = 7.3 \times 10^{-4}$, $(\Delta\omega/\omega)_f = -4.0 \times 10^{-9}$				
1	98	7.1×10^{-7}	3.2×10^{-20}	1.6×10^{-6}
2	113	1.4×10^{-8}	1.4×10^{-20}	4.5×10^{-6}
3	117	1.4×10^{-9}	1.1×10^{-20}	-1.5×10^{-6}
4	117	1.2×10^{-10}	1.1×10^{-20}	4.7×10^{-7}
5	117	2.1×10^{-11}	1.1×10^{-20}	-1.5×10^{-7}
6	117	7.3×10^{-13}	1.1×10^{-20}	5.3×10^{-8}
7	117	1.3×10^{-12}	1.1×10^{-20}	-2.1×10^{-8}
8	117	1.2×10^{-13}	1.1×10^{-20}	1.5×10^{-8}
$v_1/v = 1.5 \times 10^{-4}$, $(\Delta\omega/\omega)_f = -1.6 \times 10^{-9}$				

In case (iv) the function $\partial\psi/\partial\phi$ changes appreciably between stage one and stage two; on the other hand $J(\phi)$ changes little. In particular, the minimum value of $\partial\psi/\partial\phi$, our object of interest, changes from 0.45 to 0.35 when the mode set is expanded. This leads us to consider a further expansion of the mode set. It turns out in case (iv) that the set of Table 2, 117 modes, is already close to the

biggest that can be used, at least without extreme delicacy in the computation. We were able to add a few more modes by gradually expanding the set B_0 from which B is selected, at constant a, taking at each stage the previous solution to start a sequence of 5 iterations. Finally with B_0 consisting of all (m,n) with $1 \leq m \leq 147$, $|n| \leq 73$, we obtained a solution on a set B containing 146 modes, with $r = 8.6 \times 10^{-11}$, $D_N = 3.0 \times 10^{-26}$, and $v_1/v = 2.8 \times 10^{-4}$. Beyond 146 modes it was difficult to achieve convergence. In passing from 117 to 146, the measure of residual perturbation v_1/v no longer decreased; in fact it nearly doubled. The minimum of $\partial\psi/\partial\phi$ went through values as small as 0.22, and ended at 0.29.

A 4% increase in perturbation strengths takes us from case (iv) to case (v) with $\epsilon_1 = 2\epsilon_2 = 1.25 \times 10^{-4}$. This small step produces a large change in behavior. In case (v) a solution could not be generated starting at $g = 0$ as was done in the previous cases. Starting instead with the solution of case (iv), Table 2, and using the same B_0 we could generate in 5 iterations a solution with B containing 121 modes, and $r = 9.8 \times 10^{-10}$, $D_N = 4.0 \times 10^{-22}$, $v_1/v = 2.8 \times 10^{-4}$. Expanding the mode set as far as possible we arrived at a solution with 156 modes, $r = 2.7 \times 10^{-9}$, $D_N = 4.4 \times 10^{-28}$, and $v_1/v = 3.3 \times 10^{-4}$. The graph of $\partial\psi/\partial\phi$ for the solution is shown in Fig. 18; its minimum value is close to zero, namely 0.077. Thus, our best guess for $\epsilon_1 = 2\epsilon_2$ at the first appearance of a zero of $\partial\psi/\partial\phi$ is 1.25×10^{-4}, the value of case (v). The curve $J(\phi)$ in case (v) looks much the same as in case (iv).

For an independent check of accuracy and to locate empirically the transition to chaos we have performed numerical integrations of Hamilton's ordinary differential equations, taking initial conditions on our alleged invariant tori at the surface of section $\theta = 0$. The integration program used was not ideal for our purposes, but it allowed us to maintain sufficient accuracy for a few thousand turns (a turn meaning one intersection of an orbit with the surface of section). To control accuracy we did "backtracking": after N turns forward in θ we reversed steps to do N turns backward, and demanded that the initial and final values of (J, ϕ) agree within an error considerably smaller than the error we would tolerate

after $2N$ forward turns.

In Fig. 19 we show points generated for case (i) in a run with 4000 (forward) turns, plotted together with an enlargement of a small segment of the curve of Fig. 10. A further enlargement would be necessary to see any clear discrepancy between the points and the curve; the agreement appears to be better than one part in 10^6. In Fig. 20 we show similar results for case (iii), on a somewhat larger scale, from a run with 4000 turns. The agreement is still good, and there is no sign of stochastic behavior. Points for case (iv) from a run with 3000 turns are shown in Fig. 21. Here there is a noticeable scatter of points about the curve, the latter being from the run of Table 2 entailing 117 modes. It is hard to say whether the scatter represents chaotic behavior or merely a high-order island chain corresponding to high modes not included in the Hamilton-Jacobi equation. Finally, in Fig. 22 we show points for case (v) from a run with only 1500 turns. Here the appearance of chaotic behavior is quite definite; case (v) seems to be a little beyond transition. On the basis of backtracking experiments we think if unlikely that the scatter in Figures 21 and 22 is due to numerical error. As expected, the number of turns allowed by the backtracking criterion decreases sharply as the transition to chaos is approached.

In summary, the hypothesis that the transition to chaos corresponds to the first appearance of a zero of $\partial\psi/\partial\phi$ seems consistent with our experience in integrating Hamilton's equations. The Hamilton-Jacobi method, in the form based on a single canonical transformation, appears to over-estimate slightly the critical perturbation strengths; we cannot say by how much until more careful integrations have been performed. Weighing the evidence, we provisionally put the transition at $\epsilon_1 = 2\epsilon_2 = (1.2 \pm 0.05) \times 10^{-4}$. It should be possible to refine the estimate based on the Hamilton- Jacobi method, but it will be necessary to make at least one additional canonical transformation, in order to allow sufficiently many modes. Of course, the complementary work on Hamilton's equations should be improved, perhaps with help of methods guaranteed to produce a symplectic map.[27] Notice that it would be difficult to study the breakup of a prescribed

KAM surface by integration of Hamilton's equations alone. Finding a point on the surface to take as an initial condition would be a formidable task in itself. By combining the Hamilton-Jacobi method and Hamilton's equations, one commands an approach which is much more powerful than either method alone.

As is seen in Tables 1 and 2, the normalized determined D_N is impressively small compared to 1 near transition. For fixed ϵ's it is strongly dependent on the number of modes, decreasing sharply as the number is increased. At fixed B_0 it also decreases rapidly as the ϵ's are increased, indicating that a singularity of the Jacobian is being approached. For the B_0 of stage one ($1 \leq m \leq 63$, $|n| \leq 31$) we found $D_N = 2.7 \times 10^{-2}$, 4.0×10^{-3}, 4.6×10^{-4}, 2.7×10^{-5}, for cases (i)-(iv) respectively, while for stage two ($1 \leq m \leq 127$, $|n| \leq 63$) we found $D_N = 5.2 \times 10^{-11}$, 3.4×10^{-13}, 2.3×10^{-16}, 1.1×10^{-20}. Finally in case (v) on the largest mode set employed, $D_N = 4.4 \times 10^{-28}$. Notice that the decrease of D_N as a function of $\epsilon_1 = 2\epsilon_2$ is very much steeper in stage two than in stage one,as is reasonable if stage two represents a better model of the exact Hamilton-Jacobi system.

In all experiments we have found that $\partial\psi/\partial\phi$ acquires a zero before D_N vanishes, although D_N is always very small when $\partial\psi/\partial\phi$ has a zero. For ϵ's slightly larger than in case (v), say $\epsilon_1 = 2\epsilon_2 = 1.3 \times 10^{-4}$, we find solutions with negative minima of $\partial\psi/\partial\phi$ but with D_N still positive. These may have to do with "cantori", (tori with gaps, that exist beyond transition) and are interesting objects of further research. Perhaps $D_N = 0$ has to do with a final disappearance of cantori. In any event, smallness of D_N should be useful as a quick indicator of near-chaotic regions when one wishes to avoid the relatively costly calculation of $\partial\psi/\partial\phi$.

As was noted above, the residual perturbation after one canonical transformation turned out to be quite small, even when the KAM surface is close to breakup. That is encouraging for study of the critical region by means of further transformations. As a first step in such a program, we have computed the sec-

ond canonical transformation in lowest order, for cases (i) - (iv). We first found $V_1(\psi, K, \theta)$ by means of (4.2), (4.3), (4.7), and (4.8), and then calculated the generating function G_1 of (4.4), (4.5) from its Fourier coefficients,

$$g_{1mn} = \frac{i}{\omega m - n} \frac{1}{(2\pi)^2} \int_0^{2\pi} d\psi \int_0^{2\pi} d\theta e^{-i(m\psi - n\theta)} V_1(\psi, K, \theta) \ , \tag{7.6}$$

where ω and K are the final values of frequency and action obtained in the previous calculation of G. We used 512 and 256 mesh points for (ψ, ϕ) and θ integrations, respectively, and retained values of m and $|n|$ up to 255 and 127, respectively, in (7.6). This computation must be done in double precision, because V_1 owes its smallness to close cancellations.

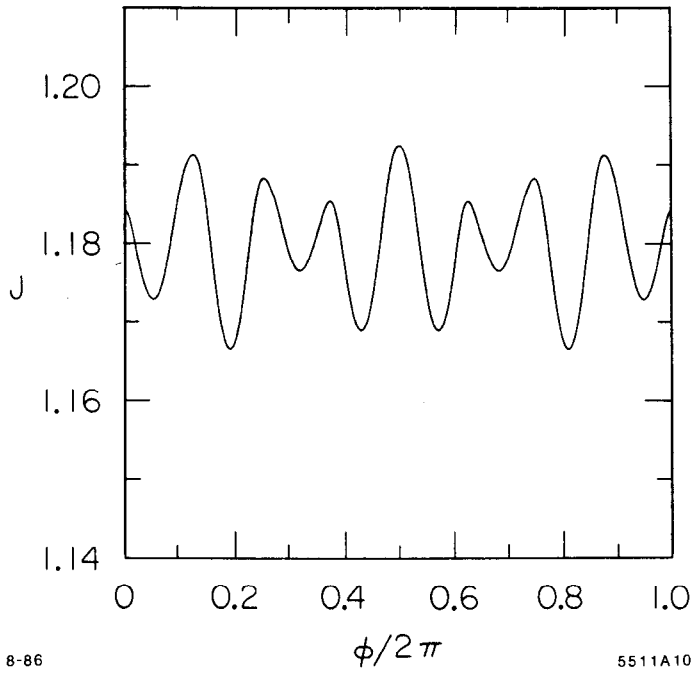
We present results in terms of the residual torus distortion, $K - K_1 = G_{1\psi}$; the notation is that of Section 4. The main torus distortion, associated with the first canonical transformation, is $J - K = G_\phi$. In Table 3 we give the ratio of the averaged absolute values of these quantities,

$$\frac{(\Delta J)_1}{\Delta J} = \frac{\langle |K - K_1| \rangle}{\langle |J - K| \rangle} = \frac{\langle |G_{1\psi}| \rangle}{\langle |G_\phi| \rangle} \ , \tag{7.7}$$

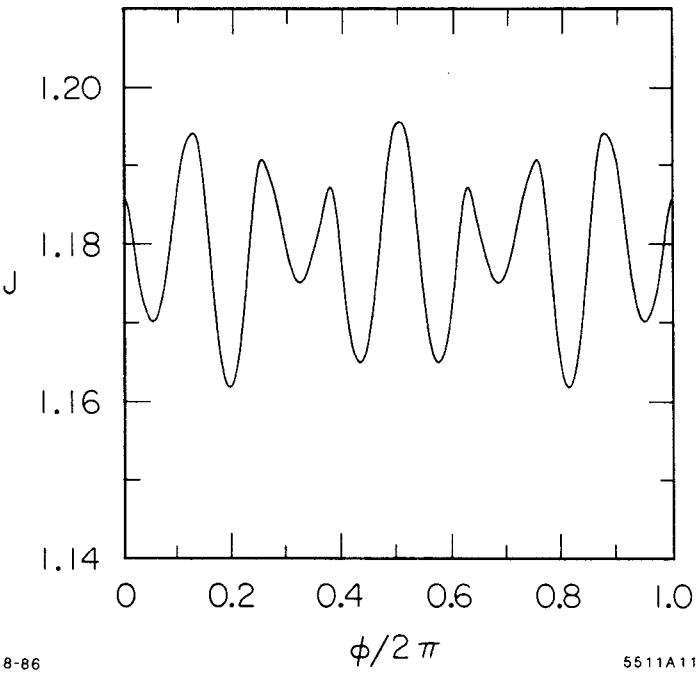
the averages being over ψ or ϕ, and θ. This ratio is small compared to 1 in cases (i)-(iv), but because of small divisors not as small as v_1/v. Except for fine details on a scale defined by Table 3, the tori obtained by the first canonical transformation would seem to be good representations of the actual KAM surfaces, even close to breakup.

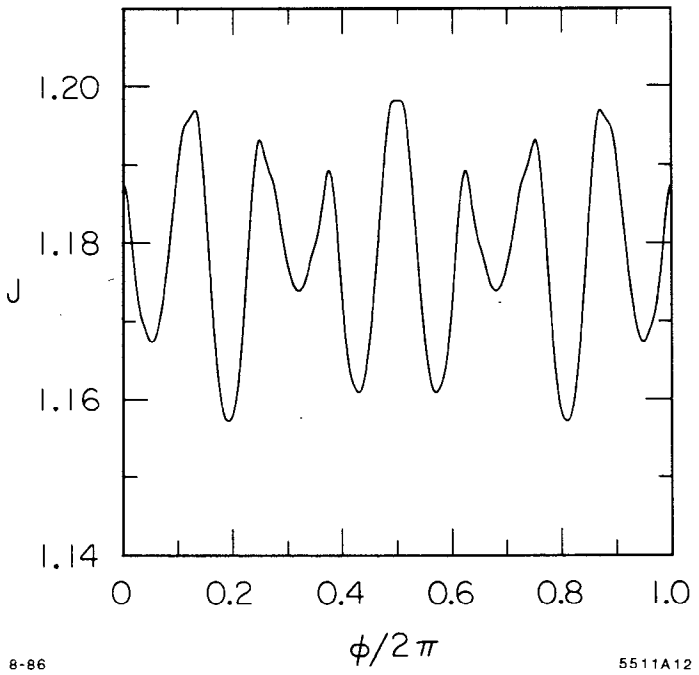
Table 3

case	$(\Delta J)_1/\Delta J$
(i)	2.8×10^{-6}
(ii)	5.0×10^{-5}
(iii)	6.5×10^{-4}
(iv)	4.1×10^{-3}

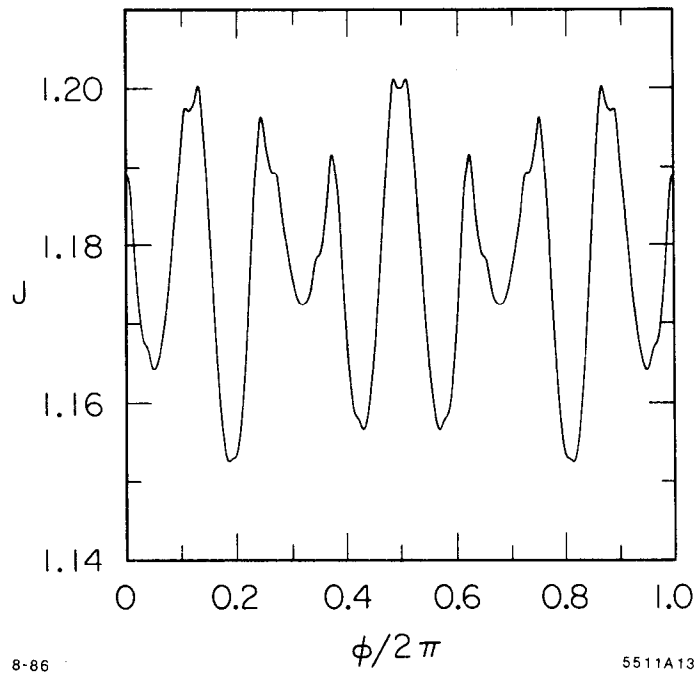


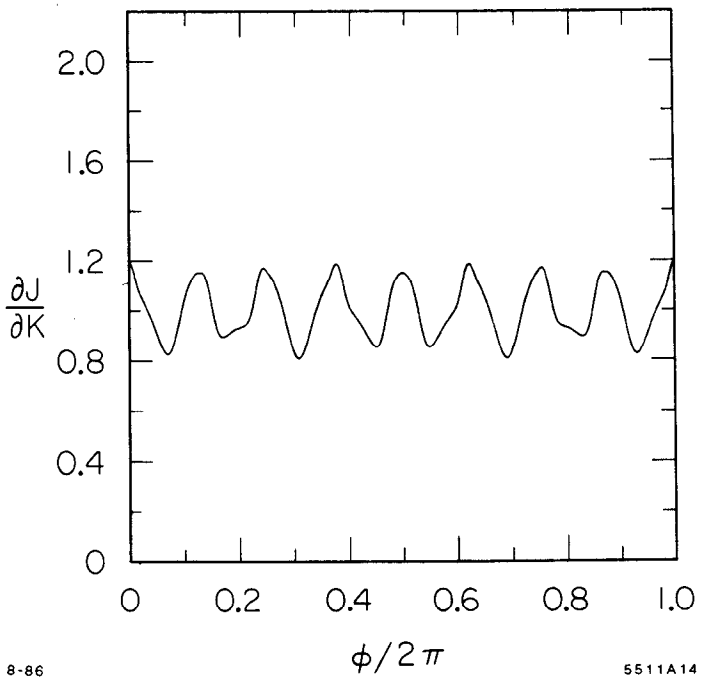
Figures 10-11



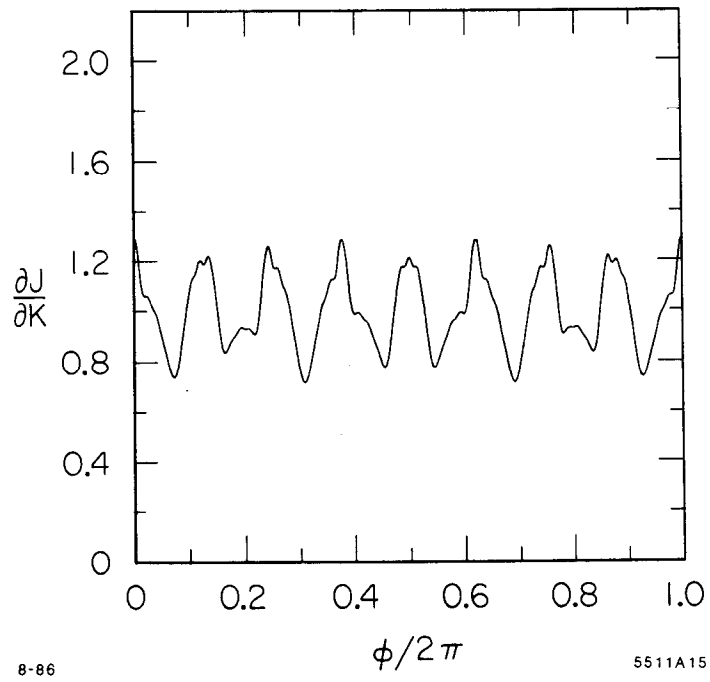


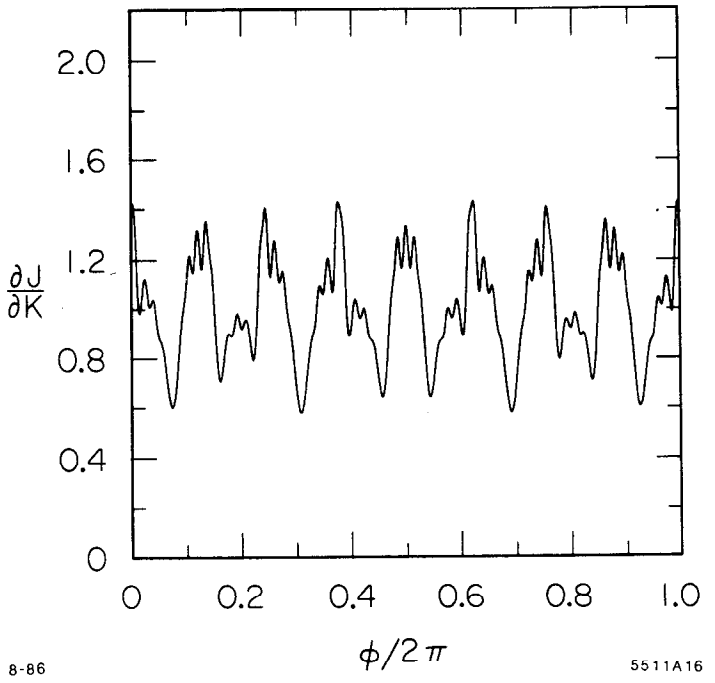
Figures 12-13



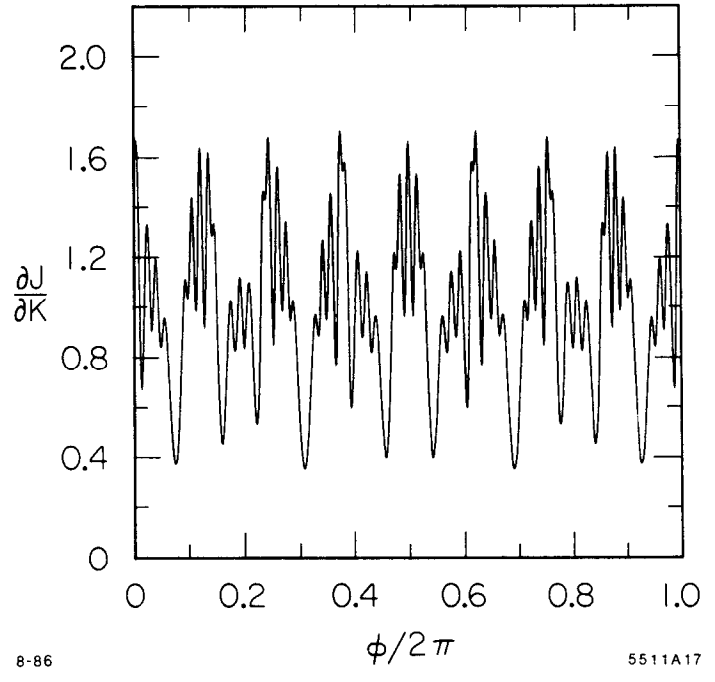


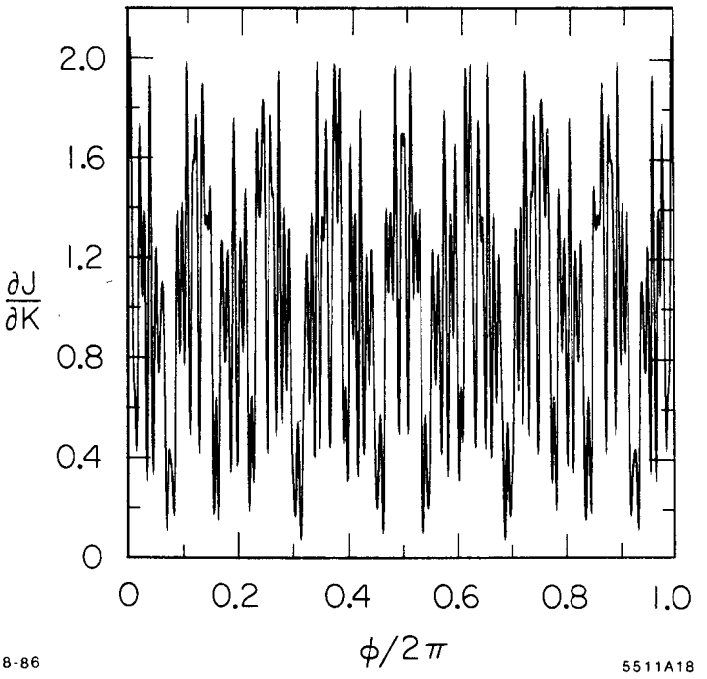
Figures 14-15



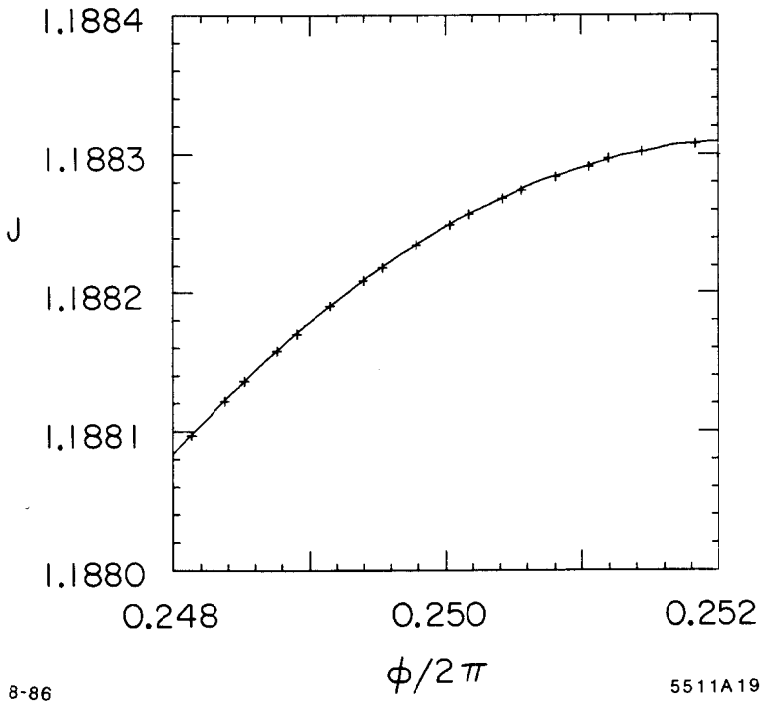


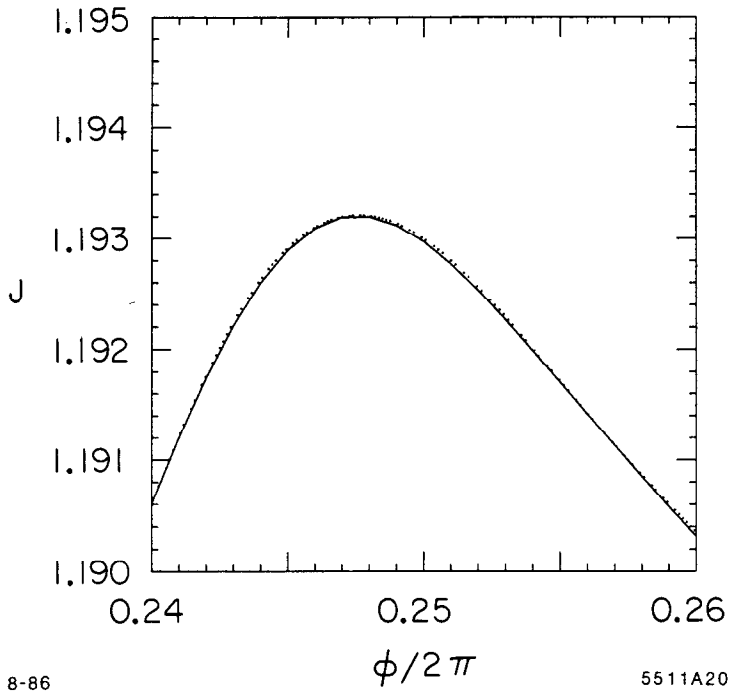
Figures 16-17



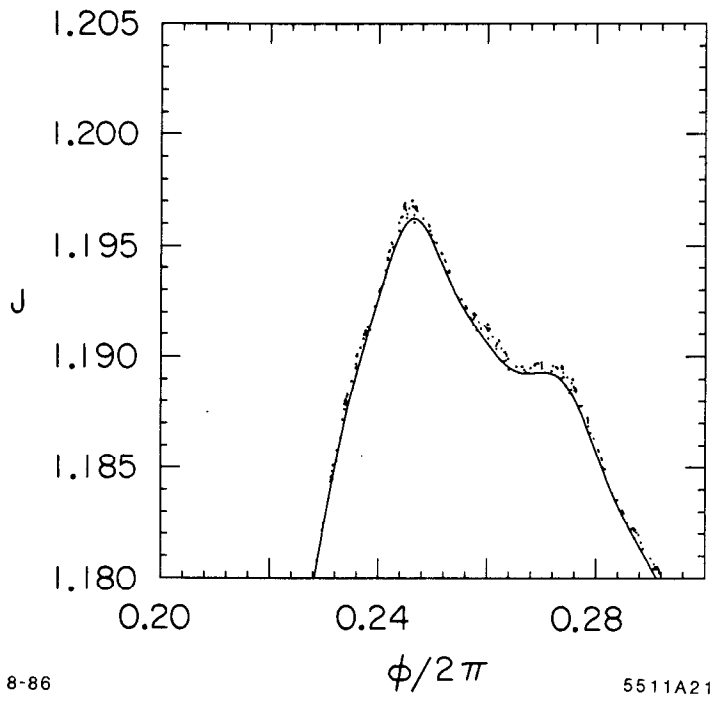


Figures 18-19





Figures 20-21



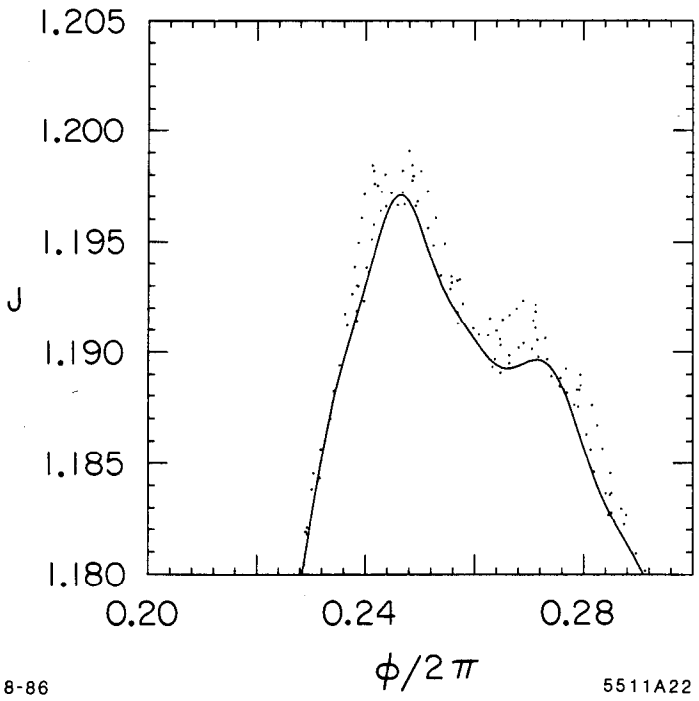


Figure 22

Figures 10-13: Two resonance model. Section of invariant surface at $\theta = 0$: J plotted *vs.* $\phi/2\pi$. Figures 10-13 correspond to cases (i)-(iv), respectively.

Figures 14-17: Two-resonance model. $\partial J/\partial K = 1 + G_{\phi K}$ at $\theta = 0$ plotted *vs.* $\phi/2\pi$. Figures 14-17 correspond to cases (i)-(iv) of Section 7, respectively.

Figure 18: Two resonance model. $\partial J/\partial K = 1 + G_{\phi K}$ at $\theta = 0$, plotted *vs.* $\phi/2\pi$. Case (v), the best candidate for the transition to chaos on the basis of Hamilton-Jacobi solutions alone.

Figure 19: A small segment of the curve of Figure 10, case (i), plus points from numerical integration through 4000 turns.

Figure 20: A small segment of the curve of Fig. 12, case (iii), plus points from numerical integration through 4000 turns.

Figure 21: A small segment of the curve of Fig. 13, case (iv), plus points from numerical integration through 3000 turns.

Figure 22: A small segment of the curve of Fig. 18, case (v), plus points from numerical integration through 1500 turns.

8. THE RESIDUE CRITERION

In this section we would like to make the connection between John Greene's residue criterion[20,21] and the associated Hamilton-Jacobi equation. To do this we need to solve the H-J equation over a finite time interval, locate an appropriate fixed point of the resulting map, and linearize about that point to calculate the residue.

To solve the H-J equation over a finite time interval it is necessary to respecify the problem and convenient to change notation slightly. We consider a canonical transformation $(\phi, J) \mapsto (\phi_i, J_i)$ defined implicitly by

$$\begin{aligned} J &= J_i + \mathcal{G}_\phi(\phi, J_i, \theta, \theta_i) \ , \\ \phi_i &= \phi + \mathcal{G}_{J_i}(\phi, J_i, \theta, \theta_i) \quad , \end{aligned} \tag{8.1}$$

where θ_i is the initial time. The H-J equation which is appropriate for the finite-time map consists of the requirement that the new Hamiltonian be identically zero

$$H(\phi, J_i + \mathcal{G}_\phi, \theta) + \mathcal{G}_\theta = 0 \quad . \tag{8.2}$$

In this case the new coordinates are the initial conditions provided that we also impose the boundary condition

$$\mathcal{G}(\phi, J_i, \theta_i, \theta_i) = 0 \quad . \tag{8.3}$$

In this case $\mathcal{G}$ is not a periodic function of θ; however, it does satisfy

$$\mathcal{G}(\phi, J_i, \theta + 2\pi, \theta_i + 2\pi) = \mathcal{G}(\phi, J_i, \theta, \theta_i) \quad , \tag{8.4}$$

since the original Hamiltonian is periodic in θ.

To study the neighborhood of a periodic orbit with period $2\pi q$, we note that such a periodic orbit is a fixed point of the map in (8.1) at (ϕ_0, J_0) provided that

$$\begin{aligned} \mathcal{G}_\phi(\phi_0, J_0, \theta_i + 2\pi q, \theta_i) &= 0 \ , \\ \mathcal{G}_{J_i}(\phi_0, J_0, \theta_i + 2\pi q, \theta_i) &= 0 \quad . \end{aligned} \tag{8.5}$$

To calculate the residue of that fixed point we linearize for small deviations about it by setting

$$\begin{aligned} \phi &= \phi_0 + \delta\phi \ , \quad \phi_i = \phi_0 + \delta\phi_i \ , \\ J &= J_0 + \delta J \ , \quad J_i = J_0 + \delta J_i \ . \end{aligned} \tag{8.6}$$

From (8.1) if we now keep terms linear in the deviation from the fixed point we obtain the linear map

$$\begin{pmatrix} \delta J \\ \delta\phi \end{pmatrix} = \frac{1}{1+\mathcal{G}_{\phi J}} \begin{pmatrix} (1+\mathcal{G}_{\phi J})^2 - \mathcal{G}_{\phi\phi}\mathcal{G}_{JJ} & \mathcal{G}_{\phi\phi} \\ -\mathcal{G}_{JJ} & 1 \end{pmatrix} \begin{pmatrix} \delta J_i \\ \delta\phi_i \end{pmatrix} \quad , \tag{8.7}$$

where all partial derivatives of $\mathcal{G}$ are evaluated at $(\phi_0, J_0, \theta_i + 2\pi q, \theta_i)$. Denoting the matrix above as M_q, the frequency or tune ν_q of the oscillation about the fixed point is given by

$$\text{Trace}(M_q) = 2\cos 2\pi\nu_q = \frac{1 + (1+\mathcal{G}_{\phi J})^2 - \mathcal{G}_{\phi\phi}\mathcal{G}_{JJ}}{1+\mathcal{G}_{\phi J}} \quad . \tag{8.8}$$

Therefore, the residue is given by[20]

$$R = \frac{1}{4}[2 - \text{Trace}(M_q)] = \frac{1}{4}\left(\frac{\mathcal{G}_{\phi\phi}\mathcal{G}_{JJ} - \mathcal{G}_{\phi J}{}^2}{1+\mathcal{G}_{\phi J}}\right) \quad . \tag{8.9}$$

In the case of an integrable system with Hamiltonian $H_0(J)$ we find

$$\mathcal{G} = -H_0(J)\theta \tag{8.10}$$

which yields $R = 0$.

In the case of a nonintegrable system we test the existence of some KAM curve with irrational frequency ω by considering the residues R_n of a sequence of periodic orbits of increasing period $2\pi q_n$ as $n \to \infty$. The elements of the sequence correspond to frequencies in the continued fraction representation of ω, say p_1/q_1, p_2/q_2, $\cdots$. According to Greene there are three distinct cases:

1. $R \to 0$, there is a KAM curve with frequency ω.
2. $R \to \pm\infty$, there is *no* KAM curve with frequency ω.
3. $R \to R_0$, the transition case.

If attention is restricted to solutions $\mathcal{G}$ with bounded second derivatives, then case (2) can arise only if

$$1 + \mathcal{G}_{\phi J}{}^{(n)} \to 0, \quad n \to \infty \, . \tag{8.11}$$

This recalls our condition that $1 + G_{\phi K}$ should first acquire a zero at transition; *cf.* Section 7. The latter condition refers to the G which generates an orbit covering an invariant surface, which is a different object from the $\mathcal{G}^{(n)}$ of (8.11). Nevertheless, for large n the orbit generated by $\mathcal{G}^{(n)}$ lies close to the surface generated by G. The failure of either condition, $1 + \mathcal{G}_{\phi J}{}^{(n)} \neq 0$ or $1 + G_{\phi K} \neq 0$, means that the corresponding canonical transformation, (8.1) or (2.4), is no longer well defined. It seems reasonable that the two conditions (the former taken in the limit $n \to \infty$) should fail simultaneously as parameters approach critical values. Further exploration of the meaning of cases (1) - (3) in the Hamilton-Jacobi formalism might be rewarding.

9. OUTLOOK AND REMARKS ON RELATED WORK

Our method and results were described briefly in Section 1, and the most important numerical results were described in detail in Section 7. Here we suggest lines for further investigation and mention current work by other authors.

We have seen that iterative solution of the truncated Hamilton-Jacobi equation provides an effective method to approximate invariant tori, even under extreme conditions, for the case of certain simple Hamiltonians in $1\frac{1}{2}$ degrees of freedom. We believe that these Hamiltonians entail most of the generic difficulties, and therefore we expect equally good convergence in more complicated systems. Computations to test this expectation, especially for models of accelerators, are underway.

It will also be interesting to test the method in problems of higher dimension. We have verified that solutions by plain iteration are feasible in $2\frac{1}{2}$ degrees of freedom,[28] but we have not yet explored Newton's method and near-chaotic regions in that case. We expect that the generalization of $\partial\psi/\partial\phi = 0$, namely $\det(1 + G_{\Phi,\mathbf{K}}) = 0$, will be a correct criterion for the breakup of KAM tori in higher dimensions.

A feature of our method which may be favorable for its application to complicated systems is that it reduces most of the computation to performing the Fast Fourier Transform. This suggests the use of enhanced FFT algorithms, for computers using vector or parallel processing, which are now being developed.[29]

In the example of Section 7 we were able to estimate the critical perturbation strengths for breakup of the golden mean KAM surface to an accuracy that seems to be about 5%. More careful work will required to improve the estimate. It should be emphasized that the example of Section 7 is much more difficult than examples studied more accurately in the literature, for instance area-preserving maps of the plane such as the standard map and the quadratic map.[20,21] In our scheme such maps can be treated with one-dimensional Fourier transforms

rather than the two-dimensional transforms required in the present examples. Therefore we should be able to treat simple maps with considerable refinement, using many more Fourier modes than we used the present study. In particular it should be possible to study self-similar behavior on progressively smaller scales. A stringent test of our method will be to calculate the critical parameter value k_c of the standard map, and compare the result with the accepted value determined by Green's residue criterion.

In Section 7 we mentioned that it would apparently be necessary to perform at least two canonical transforms to improve the computations very close to transition. Another possibility, perhaps more inspiring, would be to modify Newton's method in some way, so that more modes could be included in the first canonical transformation. In the literature there are several proposals for improving the behavior of Newton's method near a singularity of the Jacobian.[30–33]

To relate our approach to other work, we note first that it may be regarded as an instance of Galerkin's general method for solving an operator equation by finite-dimensional projections.[15,34,35] More specifically, it resembles the 'harmonic balance' technique for solving nonlinear ordinary differential equations as applied in electrical engineering by Kundert and Sangiovanni-Vincentelli.[36]

In an interesting paper in accelerator theory, Guignard and Hagel[37] have treated the equations of motion (in second order Lagrangian form) by successive linearizations. Although this recalls Newton's method, and may share its good convergence properties, issues concerning periodicity of coefficients are treated in such a way that the resulting scheme is quite different from our Newton method. Frequencies are approximated by rationals, so that each linear equation in a sequence can be treated by Floquet theory. Another scheme depending on approximation by rationals is that of Eminhizer, Helleman, and Montroll.[38] These authors avoid small denominators by imposing the requirement of rational frequencies.

Forest[39] has applied the Lie-algebraic approach to obtain explicit formulas

for approximate invariants as functions of the original canonical variables. This gives an implicit representation of the invariant torus, rather than the explicit parametrization of Eq. (2.3).

Zanetti and Turchetti[40] studied a circle map close to transition, using Fourier series and perturbation theory for direct computation of the invariant torus. This work is somewhat akin in spirit to ours, and anticipated our finding that modified classical techniques can be used near the transition to chaos.

Working with analytic perturbations (the standard map and its relatives), Greene and Percival[41] studied convergence of the Fourier series for the invariant torus with near-critical perturbation strengths. Their technique was much different from ours, but a similar study of convergence could be done in our formulation by moving the contours of angular integrations into the complex plane

ACKNOWLEDGEMENTS

We would like to thank Robert Helleman, Giorgio Turchetti, and Eugene Wayne for helpful comments and references, and Tor Raubenheimer for aid in programming. The computations were done at Lawrence Berkeley Laboratory. R. Warnock owes many thanks to the Theoretical Physics Group at LBL and the Accelerator Department at SLAC for kind hospitality.

REFERENCES

1. F. Ch. Iselin, in 'Nonlinear Dynamics Aspects of Particle Accelerators', Lecture Notes in Physics, Vol. 247, Springer, Berlin, 1986.

2. A. Wrulich, *ibid.*

3. R. V. Servranckx , IEEE Trans. Nuc. Sci. **NS-32**, No. 5 (1985) 2186.

4. V. I. Arnold, 'Mathematical Methods of Classical Mechanics', Springer, Berlin, 1978.

5. G. Gallavotti, "The Elements of Mechanics", Springer, Berlin, 1983.

6. A. J. Lichtenberg and M. A. Lieberman, "Regular and Stochastic Motion", Springer, Berlin, 1983.

7. G. Hori, Publ. Astr. Soc. Japan **18** (1966) 287.

8. A. Deprit, Cel. Mech. **1** (1969) 12; A. Deprit, J. Henrard, and A. Rom, Science, **168** (1970) 1569.

9. A. N. Kaufman, in "Topics in Nonlinear Dynamics", A.I.P. Conference Proceedings No. 46, Amer. Inst. Phys., New York, 1978.

10. R. Littlejohn, "A Pedestrian's Guide to Lie Transforms: A New Approach to Perturbation Theory in Classical Mechanics", Lawrence Berkeley Laboratory, unpublished report, 1978; J. Math. Phys. **20** (1979) 2445, *ibid.* **23** (1982) 742.

11. J. R. Cary, Physics Reports **79** (1981) 129, J. R. Cary and R. Littlejohn, Ann. Phys. **151** (1983) 1.

12. B. McNamara, J. Math. Phys. **19** (1978) 31.

13. L. Michelotti, Particle Accelerators **16** (1985) 233; Report FERMILAB-Conf-86/30, Fermi National Accelerator Laboratory, Batavia, Illinois, 1986, to be published in "1984 U.S. Summer School on High-Energy Accelerators", AIP Conference Proceedings.

14. A. J. Dragt and J. M. Finn, J. Math. Phys. **20** (1979) 2649, **17** (1976) 2215; A. J. Dragt, in "Physics of High Energy Particle Accelerators", AIP Conference Proceedings No. 87, Amer. Inst. Phys., New York, 1982.

15. M. A. Krasnosel'skii, G.M. Vainikko, P.P. Zabreiko, Ya. B. Rutitskii, and V. Ya. Stetsenko, 'Approximate Solution of Operator Equations', Wolters-Noordhoff, Groningen, 1972.

16. L. V. Kantorovich and G. P. Akilov, 'Functional Analysis in Normed Spaces', Pergamon, Oxford, 1964.

17. J. Pöschel, Comm. Pure Appl. Math. **35** (1982) 653.

18. L. Chierchia and G. Gallavotti, Nuovo Cimento **B67** (1982) 277; G. Gallavotti, in 'Scaling and Self-Similarity in Physics', J. Fröhlich, Ed., Birkhäuser, Basel, 1983.

19. D. F. Escande and F Doveil, J. Stat. Phys. **26** (1981) 257; Phys. Lett. **83A** (1981) 307.

20. J. M. Greene, J. Math. Phys. **20** (1979) 1183; contribution to 'Nonlinear Dynamics and the Beam-Beam Interaction', AIP Conference Proceedings, No. 57, Amer. Inst. Phys., New York, 1979.

21. R. S. MacKay, in 'Nonlinear Dynamics Aspects of Particle Accelerators', Lecture Notes in Physics, Vol. 247, Springer-Verlag, Berlin, 1986; 'Renormalization in Area Preserving Maps', Ph.D. Thesis, Princeton Univ., 1982.

22. V. I. Krylov and L. G. Kruglikova, 'Handbook of Numerical Harmonic Analysis', Israel Program for Scientific Translations, Jerusalem, 1969.

23. IMSL, 2500 City West Blvd., Houston, Texas.

24. J. R. Bunch, SIAM J. Numer. Anal. **8** (1971) 656.

25. S. D. Conte, 'Elementary Numerical Analysis', McGraw-Hill, New York, 1965

26. B. V. Chirikov, Phys. Reports **52**, No. 5 (1979) 264.

27. R. D. Ruth, IEEE Trans. on Nucl. Sci. **NS-30**, No. 4 (1983) 2669.

28. R. D. Ruth, T. Raubenheimer, and R. L. Warnock, IEEE Trans. Nucl. Sci. **NS-32** No. 5 (1985) 2206.

29. P. N. Swartztrauber, Parallel Computing **1** (1984) 45.

30. A. Griewank, SIAM Review **27** (1985) 537.

31. H. B. Keller in 'Applications of Bifurcation Theory', P. H. Rabinovitz, Ed., Academic Press, New York, 1977.

32. F. H. Branin Jr., IBM J. Res. Devel. **16** (1972) 504.

33. P. M. Anselone and R. H. Moore, J. Math. Anal. Appl. **13** (1966) 476.

34. M. Urabe, Arch. Rat. Mech. Anal. **20** (1965) 120.

35. N. S. Kurpel', 'Projection-Iterative Methods for Solution of Operator Equations', Translations of Math. Monographs, No. 46, Amer. Math. Soc., Providence, R.I., 1976.

36. K. S. Kundert and A. Sangiovanni-Vincentelli, 'Simulation of Nonlinear Circuits in the Frequency Domain', preprint, Dept. of Elec. Eng. and Comp. Sci., Univ. of California, Berkeley, 1986.

37. G. Guignard and J. Hagel, Particle Accelerators **18** (1986) 129.

38. C. R. Eminhizer, R. H. G. Helleman, and E. W. Montroll, J. Math. Phys. **17** (1976) 121.

39. E. Forest, SSC Central Design Group Reports SSC-29 and SSC-30, Lawrence Berkeley Lab.

40. G. Zanetti and G. Turchetti, Lett. Nuovo Cimento **41** (1984) 90.

41. J. M. Greene and I. C. Percival, Physica **3D** (1981) 530.

APPENDIX A

EXISTENCE OF SOLUTIONS OF THE TRUNCATED HAMILTON-JACOBI EQUATION

When the action parameter K and the Hamiltonian H are suitably restricted, the Hamilton-Jacobi equation for the truncated periodic generating function, Eq. (2.17), has a unique small solution which can be computed by iteration. Furthermore, that solution is locally a smooth function of K. The region of smoothness in K shrinks as the number of Fourier modes is increased. Under somewhat stronger restrictions on the Hamiltonian, similar statements apply if the frequency ω rather than the action K is specified. We now make these statements precise by proving some theorems. The proofs are elementary, but they give a worthwhile first view of the problem.

For the analysis at fixed K the equation in question is (2.17),

$$g = A(g, K), \tag{A1}$$

where the vector function A is defined by (2.18), with G_ϕ given by the truncated Fourier series

$$G_\phi = 2\mathrm{Re}\left(\sum_{(m,n)\in B} img_{mn} e^{i(m\phi - n\theta)} \right) \tag{A2}$$

The set B is any finite set of pairs of integers (m, n) such that $m \geq 1$. For notational simplicity we take $1\frac{1}{2}$ degrees of freedom; the generalization to higher dimensional phase space will be obvious. In (A1) the argument K refers to the explicit occurrence of K in $H(\phi, K + G_\phi, \theta)$, $H_0(K)$, and $\omega_0(K) = H_0'(K)$.

We seek a solution at some $K = K_0$ such that

$$\omega_0(K)m - n \neq 0 \ , \quad (m,n) \in B. \tag{A3}$$

We suppose that $H_0''(J), V(\phi, J, \theta)$, and $V_J(\phi, J, \theta)$ are real and continuous for

$$|J - K_0| < R \ , \qquad \phi, \theta \in [0, 2\pi], \tag{A4}$$

for some positive R independent of ϕ, θ. We regard (A1) at $K = K_0$ as a fixed point problem for the vector $x = \{x_{mn}\}$, where

$$x_{mn} = m\left(g_{mn} - g_{mn}^{(1)}\right) \ , \quad (m,n) \in B \ . \tag{A5}$$

The $g_{mn}^{(1)}$ are the coefficients of g to lowest order in perturbation theory, as defined in (3.3). The quantity x_{mn}, being generally smaller than mg_{mn}, is an appropriate object for an iterative analysis. In terms of x the problem is

$$x = B(x, K) \tag{A6}$$

$$B(x,K)_{mn} = \frac{im}{m\omega_0(K) - n} \frac{1}{(2\pi)^2} \int\int d\phi d\theta e^{-i(m\phi - n\theta)} \tag{A7}$$
$$\left[H_0(K + G_\phi) - H_0(K) - \omega(K)G_\phi + V(\phi, K + G_\phi, \theta) - V(\phi, K, \theta)\right] \ .$$

We consider (A6) on a complete metric space S_γ with distance $d(x,y) = ||x - y||$, where

$$||x|| = \max_{(m,n)\in B} |x_{mn}| \ . \tag{A8}$$

The space S_γ is the ball with radius γ,

$$S_\gamma = \left\{x \mid \ ||x|| \leq \gamma\right\} . \tag{A9}$$

We look for conditions on the Hamiltonian such that for some γ,

(*a*) $x \in S_\gamma$ implies $B(x, K) \in S_\gamma$.

(*b*) There is a $\beta < 1$ such that

$$||B(x_1) - B(x_2)|| \leq \beta ||x_1 - x_2|| \quad \text{for all } x_1,\ x_2 \in S_\gamma\ . \tag{A10}$$

If (a) and (b) hold, the contraction mapping principle implies that there is a unique solution $g \in S_\gamma$. Further, that solution is obtained by iteration, starting with any element of S_γ.

If $x \in S_\gamma$, then from (A2) we see that

$$|G_\phi| \leq |G_\phi^{(1)}| + |G_\phi - G_\phi^{(1)}| \leq \Gamma + d||x|| \leq \Gamma + d\gamma\ , \tag{A11}$$

where

$$\Gamma = \sup_{\phi,\theta} \left| G_\phi^{(1)}(\phi, \theta) \right|\ , \tag{A12}$$

$$d = 2 \sum_{m,n \in B} 1\ . \tag{A13}$$

Here $G_\phi^{(1)}$ is the part of G_ϕ due to $g^{(1)}$, and d ('diameter of B') is twice the number of elements in B. To find conditions for (A10), we start by restricting γ and the perturbation V so that

$$\Gamma + d\gamma < R\ . \tag{A14}$$

Then by (A4) and (A11) the Hamiltonian functions evaluated at $J = K_0 + G_\phi$ are within their domains of continuity when $x \in S_\gamma$. Let us define

$$C = \max_{(m,n)\in B} \left| \frac{m}{\omega_0(K_0)m - n} \right|, \tag{A15}$$

$$\epsilon = \sup_{\phi,\theta} \sup_{|J-K_0|<R} |V(\phi, J, \theta)|, \tag{A16}$$

$$\epsilon' = \sup_{\phi,\theta} \sup_{|J-K_0|<R} |V_J(\phi, J, \theta)|, \tag{A17}$$

$$\alpha = \sup_{|J-K_0|<R} |H_0''(J)|. \tag{A18}$$

A crude upper bound for Γ is

$$\Gamma \leq Cd\epsilon, \tag{A19}$$

but since one can often get better bounds in specific cases we shall write results in terms of Γ itself.

Now if $x \in S_\gamma$, the mean value theorem applied to (A17) gives

$$||B(x)|| \leq C \left[\frac{\alpha}{2} (\Gamma + d\gamma)^2 + \epsilon' (\Gamma + d\gamma) \right] ; \tag{A20}$$

thus $B(x) \in S_\gamma$ and condition (a) holds if

$$C(\Gamma + d\gamma) \left[\frac{\alpha}{2} (\Gamma + d\gamma) + \epsilon' \right] \leq \gamma \tag{A21}$$

For the contraction condition (A10) consider G_{ϕ_1}, G_{ϕ_2} corresponding to x_1, x_2 respectively; clearly

$$|G_{\phi_1} - G_{\phi_2}| \leq d||x_1 - x_2|| \quad . \tag{A22}$$

Note this lemma:

If $a(u)$ has a continuous second derivative for $|u| < \xi$ and $b(u) = a(u) - a'(0)u$, then

$$|b(u_1) - b(u_2)| \leq \frac{|u_1^2 - u_2^2|}{2} \sup_{|u|<\xi} |a''(u)|, \tag{A23}$$

Applying the lemma to $a(u) = H_0(K_0 + u)$ we have

$$\left|H_0(K + G_{\phi_1}) - \omega_0(K)G_{\phi_1} - H_0(K + G_{\phi_2}) + \omega_0(K)G_{\phi_2}\right| \leq \alpha d(\Gamma + d\gamma)||x_1 - x_2|| \tag{A24}$$

Hence (A10) holds provided that

$$C\left[\alpha d(\Gamma + d\gamma) + \epsilon' d\right| < 1. \tag{A25}$$

The sufficient conditions for a solution, (A21)and (A25), may be put into a more appealing form. Inequality (A21) is

$$P(\gamma) = a\gamma^2 + b\gamma + c = a(\gamma - \gamma_+)(\gamma - \gamma_-) \leq 0$$

$$a = Cd^2\frac{\alpha}{2} \quad , \quad b = Cd(\alpha\Gamma + \epsilon') - 1 \ , \tag{A26}$$

$$c = C\Gamma\left(\frac{\alpha}{2}\Gamma + \epsilon'\right)$$

The roots $\gamma_\pm$ of $P(\gamma)$ must be real, and $\gamma_- \leq \gamma \leq \gamma_+$, for (A21) to hold. The discriminant is

$$D = b^2 - 4ac = (1 - \epsilon' Cd)^2 - 2\alpha\Gamma Cd\,. \tag{A27}$$

Regarded as a polynomial in ϵ', D has roots

$$\epsilon'_{\pm} = \frac{1 \pm (2\alpha\Gamma Cd)^{1/2}}{Cd}. \tag{A28}$$

Reality of $\gamma_{\pm}$ requires $\epsilon' \leq \epsilon_-$ or $\epsilon' \geq \epsilon_+$, but the contraction condition (A25) requires $\epsilon' < 1/Cd$, so that only $\epsilon' \leq \epsilon_-$ is allowed. Inequality (A25) is more restrictive than $\gamma \leq \gamma_+$; it requires

$$\gamma < \widehat{\gamma} = \frac{1 - Cd(\alpha\Gamma + \epsilon')}{\alpha Cd^2} \tag{A29}$$

Thus (A21) and (A25) require

$$\gamma_- \leq \gamma < \widehat{\gamma}, \tag{A30}$$

$$\epsilon' \leq \epsilon_-. \tag{A31}$$

Conversely, (A30), (A31) imply (A21), (A25).

We must verify that (A30) allows $\gamma > 0$, as is necessary by the definition of γ. Since $c = a\gamma_+\gamma_-$ and $b = -a(\gamma_+ + \gamma_-)$, the roots $\gamma_{\pm}$ are positive if $b < 0$. But it is easy to show that $b < 0$ is implied by (A31).

Accounting for (A30), (A31), and (A14), we can now summarize the results in

> Theorem 1: Suppose that $\omega_0(K)m - n \neq 0$, $(m,n) \in B$, and that $H_0''(J)$, $V(\phi, J, \theta)$ and $V(\phi, J, \theta)$ are real and continuous for $|J - K_0| < R$ and $\phi, \theta \in [0, 2\pi]$. Then Eq. (A1), truncated to the finite mode set B, has a unique solution in the ball S_γ of (A9) with

$$\frac{1 - Cd(\alpha\Gamma + \epsilon') - \left[(1 - \epsilon' Cd)^2 - 2\alpha\Gamma Cd\right]^{1/2}}{\alpha C d^2} \leq \gamma$$
$$< \min\left[\frac{1 - Cd(\alpha\Gamma + \epsilon')}{\alpha C d^2}, \frac{R - \Gamma}{d}\right] \tag{A32}$$

provided that

$$\epsilon' + \left(\frac{2\alpha\Gamma}{Cd}\right)^{1/2} < \frac{1}{Cd}\,. \tag{A33}$$

For any finite mode number set B and any R, it is possible to choose the Hamiltonian parameters $\Gamma, \epsilon', \alpha$ so that (A33) holds and the set of γ satisfying (A32) is not empty. Provided that $(R - \Gamma)/d > \gamma_-$, the only further restriction on the Hamiltonian parameters is (A33). Sufficient conditions for (A33) to hold are $\epsilon' < 1/4\mathrm{Cd}$, $\alpha\Gamma < 1/4Cd$, or, in view of (A19), $\epsilon' + (2\alpha\epsilon)^{1/2} < 1/Cd$.

Notice that the solution referred to is the same for all γ in the interval (A32). Thus, the solution is actually in $\mathcal{S}_\gamma \supset \mathcal{S}_{\gamma_-}$, for any γ obeying (A32). The solution can be obtained by iteration of (A6), beginning with any $x_0 \in \mathcal{S}_\gamma$, say $x_0 = 0$. Iteration of (A1) beginning with $g = 0$ leads to precisely the same sequence so that in a numerical calculation there is no point in using (A6).

To show that the solution of Theorem 1 is locally a smooth function of K we apply the implicit function theorem to

$$\mathcal{F}(x, K) = x - B(x, K) = 0. \tag{A34}$$

The idea of the argument is that if (A3) holds at $K = K_0$, it also holds in some neighborhood of K_0, by continuity. To ensure that we stay in the domain (A4) of the Hamiltonian functions, we replace the requirement (A14) by

$$\Gamma + d\gamma < R' < R, \tag{A35}$$

where the difference between R' and R allows for a small variation of K. We ask whether the unique solution $x = x_0$ in S_γ at $K = K_0$ belongs to a family $x(K)$, $|K - K_0| < \delta$, $x(K_0) = x_0$. The function $\mathcal{F}(x, K)$ is continuously differentiable in both arguments, x and K, in some neighborhood of (x_0, K_0). If the Jacobian $\mathcal{F}_x$ is nonsingular at (x_0, K_0), then the implicit function theorem guarantees that the desired solution family $x(K)$ exists, and in fact has a continuous K derivative, in some neighborhood of K_0. A sufficient condition for regularity of the Jacobian is

$$||B_x(x_0, K_0)|| = \sup_h \frac{||B_x(x_0, K_0)h||}{||h||} < 1 \tag{A36}$$

Calculating the Jacobian as in Section 3 and using mean-value arguments as above, we see that (A36) holds under our previous contraction condition (A25).

Theorem 2: Replace R in (A32) by some $R' < R$. Then under the conditions of Theorem 1 so modified, the unique solution $g_\star$ of (A1) in S_γ for $K = K_0$ is a member of a family of solutions $g(K)$, defined and continuously differentiable with respect to K in some neighborhood of K_0, such that $g(K_0) = g_\star$.

APPENDIX B

FORMULAS FOR INVARIANT CURVES IN THE 4TH ORDER RESONANCE MODEL

We wish to explore the function $J(\phi)$ of Eq. (6.6) for various values of a and γ, with the factor $(1/4 - \nu)/2\epsilon$ put equal to 1.

There is a qualitative difference between the cases $|a| < 1$ and $|a| > 1$; we discuss them separately. Let us define the square root symbol in (6.6) to be

always nonnegative, and patiently treat in order the various cases as follows. On each branch we quote inequalities that follow directly from the requirement $J > 0$.

(A) $|a| < 1$

(A.1) The branch $J = \frac{1-[1+\gamma(a+\cos 4\psi)]^{1/2}}{a+\cos 4\psi}$

$$\gamma < 0 \quad .$$

(A.1.a) $J > 0$ for all ψ if and only if

$$-\frac{1}{1+a} \leq \gamma < 0$$

(A.1.b) For $\gamma < -\frac{1}{1+a}$,

$$a + \cos 4\psi \leq \frac{1}{|\gamma|}$$

(A.2) The branch $J = \frac{1+[1+\gamma(a+\cos 4\psi)]^{1/2}}{a+\cos 4\psi}$

$$a + \cos 4\psi > 0 \qquad (\dagger)$$

(A.2.a) $J > 0$ for all ψ satisfying $(\dagger)$ if and only if

$$-\frac{1}{1+a} < \gamma$$

(A.2.b) For $\gamma < -\frac{1}{1+a}$,

$$0 < a + \cos 4\psi \leq \frac{1}{|\gamma|}$$

In Figures 1 and 2 we show phase space plots of $(x, y) = (J^{1/2} \cos\psi, J^{1/2} \sin\psi)$ for case (A). These may be regarded as surfaces of section at $\theta = 0$, where $\psi = \phi$. Each plot is for fixed a, and shows a few curves from each family that exists at that a; the curves of a family correspond to various γ.

Figure 1 is for the case of no nonlinear term in the unperturbed Hamiltonian, $a = 0$, whereas Fig. 2 is for $a = -0.9$. The bounded closed curves near the origin are from (A.1.a) with

$$\frac{-1}{1+a} < \gamma < 0 \quad . \tag{B1}$$

The separatrix corresponds to the lower limit of γ. The unbounded curves symmetrical about the x and y axes are from case (A.2.a) with

$$\frac{-1}{1+a} < \gamma < \infty \quad . \tag{B2}$$

The unbounded curves symmetrical about the lines $y = \pm x$ are made up of cases (A.1.b) and (A.2.b) together; for the same γ the two cases give fragments which fit together to make continuous curves, for

$$-\infty < \gamma \leq \frac{-1}{1+a} \quad . \tag{B3}$$

(B) $|a| > 1$

(B.1) The branch $J = \frac{1-[1+\gamma(a+\cos 4\psi)]^{1/2}}{a+\cos 4\psi}$

(B.1.a) $a < 0$. $J > 0$ for all ψ if and only if

$$-\infty < \gamma < 0$$

(B.1.b) $a > 0$. $J > 0$ for all ψ if and only if

$$-\frac{1}{1+a} \leq \gamma < 0$$

(B.1.c) $a > 0$. For $\gamma < -\frac{1}{1+a}$,

$$a + \cos 4\psi \leq \frac{1}{|\gamma|} \quad , \quad \text{hence} \quad -\frac{1}{a-1} < \gamma$$

(B.2) The branch $J = \frac{1+[1+\gamma(a+\cos 4\psi)]^{1/2}}{a+\cos 4\psi}$

$$a > 0 \quad \text{only}$$

(B.2.a) $J > 0$ for all ψ if and only if

$$-\frac{1}{1+a} \leq \gamma$$

(B.2.b) For $\gamma < -\frac{1}{1+a}$,

$$a + \cos 4\psi \leq \frac{1}{|\gamma|} \quad , \quad \text{hence} \quad -\frac{1}{a-1} < \gamma$$

Figures 3 shows the case $a = 1.2$. The curves near the origin correspond to case (B.1.b) with

$$\frac{-1}{1+a} \leq \gamma < 0 \quad , \tag{B4}$$

the lower limit of γ giving the separatrix. The outside curves, invariant under rotation by 90°, correspond to (B.2.a) for

$$\frac{-1}{1+a} \leq \gamma < \infty \quad . \tag{B5}$$

The islands, having no rotational symmetry, correspond to cases (B.1.c) and

(B.2.b) together with

$$\frac{-1}{a-1} \leq \gamma \leq \frac{-1}{a+1} \quad . \tag{B6}$$

Each island has an outer piece given by (B.2.b) and an inner piece given by (B.1.c); the pieces join at the turning points where $dJ/d\psi$ goes to infinity.

It is easy to see how Figure 1 evolves into Figure 3 as a crosses 1. The ends of the infinite separatrices join at ∞ on the lines $y = \pm x$ when $a = 1$, and then move inward. Unbounded curves that were symmetrical about $y = \pm x$ turn into islands, while those symmetrical about the x and y axes turn into bounded curves with 90° rotational invariance. Indeed, within a certain distance of the origin Figures 1 and 3 are similar in appearance.

The remaining case $a < -1$ is shown in Figure 4 for $a = -1.01$. This corresponds to (B.1.a) with

$$-\infty < \gamma < 0 \quad . \tag{B7}$$

Here all curves have symmetry under 90° rotations, and they approach concentric circles under further decrease of a. Comparing to Figure 3, we see that separatrices have gone to infinity, so that no unbounded curves can turn into islands as they did for positive a.

In the special case $\nu = 1/4$,

$$J = \left[\frac{E/\epsilon}{a + \cos 4\psi}\right]^{1/2} \quad , \tag{B8}$$

which consists entirely of curves with 90° rotation symmetry, all unbounded for $|a| < 1$, all bounded for $|a| > 1$. For $|a| < 1$ there are no stable trajectories that stay close to the origin in phase space.

The action integral (6.5) giving the invariant K makes sense only for the bounded curves having symmetry under 90° rotations. For unbounded curves it is divergent, while for islands J is defined as a real function only between turning

points, not on the whole interval $[0, 2\pi]$. Correspondingly, our Hamilton-Jacobi formalism applies only to the rotationally invariant curves, which are the KAM surfaces that we aspire to compute.

For E corresponding to bounded separatrices it is possible to evaluate the integral (6.5) analytically. For $|a| < 1$ on the separatrix,

$$K = \frac{1}{\pi}\left|\frac{\nu - \frac{1}{4}}{\epsilon}\right|(1-a^2)^{-1/2}\ln\left[\frac{2^{1/2}+(1-a)^{1/2}}{(1+a)^{1/2}}\right] \quad . \tag{B9}$$

For $a > 1$ on the separatrices

$$K = \frac{2}{\pi}\left|\frac{\nu - \frac{1}{4}}{\epsilon}\right|(a^2-1)^{-1/2}\tan^{-1}\left[\frac{(1+a)^{1/2} \pm 2^{1/2}}{(1+a)^{1/2} \mp 2^{1/2}}\right]^{1/2} \quad , \tag{B10}$$

where the upper (lower) sign corresponds to the separatrix outside (inside) the islands.