On the construction of invariant tori and integrable Hamiltonians

Mikko K.J. Kaasalainen
Merton College

Theoretical Physics
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Trinity 1994

On the construction of invariant tori and integrable Hamiltonians

Mikko K.J. Kaasalainen
Merton College

Theoretical Physics
University of Oxford

Abstract

The main principle of this thesis is to employ the geometric representation of Hamiltonian dynamics: in a broad sense, we study how to construct, in phase space, geometric structures that are related to a dynamical system. More specifically, we study the problem of constructing phase-space tori that are approximate invariant tori of a given Hamiltonian; also, using the constructed tori, we define an integrable Hamiltonian closely approximating the original one. The methods are generally applicable; as examples, we use gravitational potentials that are of interest in stellar dynamics.

First, we construct tori for box and loop orbits in planar, barred potentials, thus demonstrating the applicability of the scheme to potentials that have more than one major orbit family. Also, we show that, in general, the construction scheme needs two types of canonical transformations together: point transformations as well as those expressed by generating functions. To complete the construction scheme, we show how to furnish the tori with consistent coordinate systems, i.e., how to recover the angle variables of a torus labelled by its actions.

Next, the developed methods are employed in creating invariant phase-space tori in non-integrable potentials supporting minor-orbit families. These tori are used to define an integrable Hamiltonian H_0, and a modified form of the standard Hamiltonian perturbation theory is then used to demonstrate that a minor-orbit family can be treated as one made up of orbits trapped by a resonance of H_0.

Finally, we generalize the scheme further by constructing tori in time-reversal asymmetric Hamiltonians (by considering the motion in a rotating frame of reference), and study the transition from locally contained stochasticity to global chaos. Using both near-integrable 'laboratory' Hamiltonians and those for which we construct tori, we investigate the transition in the light of the resonance overlap criterion.

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Trinity 1994

Acknowledgements

First and foremost, I have the pleasure of thanking my supervisor, James Binney, for providing guidance and a constant supply of suggestions and pushes in the right direction. Also, I would like to thank Tom Quinn, Phil Cuddeford and Prasenjit Saha for useful discussions and comments, and Karl Johan Donner for introducing me to the field of galactic dynamics and for suggesting that I go to Oxford.

Ah...then the friends-and-all-that-section. Of course I cannot thank everyone separately, but, to make at least some people leaf through this thesis, I would like to mention a few, for *ipso facto* those few are hereby committed to reading this bally thing, or at least having something nice to say about it: my parents for more things than there is space for, but especially for their encouraging me to go my own way and try and find out the answers myself; Patrick May for innumerable things as well, and for making my vacation visits to Finland always so enjoyable; the same goes with Jukka Aunu; Roger Colinvaux for being a good fellow Mertonian and assisting me in my secondary Oxford project of studying Single Malt Whiskies; Petri Mähönen for keeping up the banner of the Great Finnish Expedition 91-94 to Keble Road with me; and Bodil Holst for poetry and all that. And further thanks go to many people not mentioned here.

Finally, I would like to thank the Osk. Huttunen Foundation, the British Council, and the ORS Awards Scheme for financial support. Also, for those quotation-aficionados: quotations have their place, but in this case, here's the one and only quotation in this thesis – *"Quoting is not unlike playing bits of Menuhin recordings at one's own violin recital."* (M.K.)

Contents

Chapter 1

Introduction

1.1 Phase-space tori and integrability

In an integrable N-dimensional Hamiltonian H, motion is regular and can be described by action-angle variables $(\boldsymbol{J}, \boldsymbol{\theta})$ such that $H = H(\boldsymbol{J})$ and, from Hamilton's equation of motion,

$$\begin{aligned} \dot{\boldsymbol{J}} &= -\frac{\partial H}{\partial \boldsymbol{\theta}} = 0, \\ \dot{\boldsymbol{\theta}} &= \frac{\partial H}{\partial \boldsymbol{J}} \equiv \boldsymbol{\omega}(\boldsymbol{J}), \end{aligned} \tag{1.1}$$

where $\boldsymbol{\omega}$ is the orbital frequency vector. Therefore, the actions $\boldsymbol{J}$ are constant and the angles $\boldsymbol{\theta}$ increase linearly in time:

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\omega}(\boldsymbol{J})t, \tag{1.2}$$

where $\boldsymbol{\theta}_0$ is an initial phase and t the time. Thus, when one knows the relationship between the action-angle variables and a point $\boldsymbol{w} = (\boldsymbol{x}, \boldsymbol{p})$ (coordinates and their canonically conjugate momenta) in phase-space, integrating the motion is trivial. Unfortunately, this relationship cannot usually be given in a closed, analytical form, even if the Hamiltonian is known to be integrable. Fortunately, analytical forms are available for some nontrivial Hamiltonians of different types; therefore, it is natural to seek canonical transformations that could link a system under study to these. Also, if our Hamiltonian is not integrable, one would like to know how it could best be approximated by an integrable one expressable in action-angle form; obviously, this could be advantageous in studying the dynamics of the system. In achieving these, at first glance dynamical, goals, the geometric view of Hamiltonian dynamics is of great help.

In an integrable Hamiltonian, almost all orbits fill densely 'invariant tori'; these are N-dimensional surfaces in phase space that are topologically identical to an N-torus. 'Almost all' means here in the same sense that 'almost all' real numbers are irrational; even though there are infinitely many rational numbers (and they are everywhere dense), they form a set of zero measure. The rational numbers correspond to resonant orbits, i.e. those with commensurable frequencies. Even though such an orbit does not fill a torus densely, it is nevertheless confined to one (and in practice, of course, an orbit with large commensurability integers is a non-resonant one for all computational purposes).

Since the actions for an orbit are constant, the invariant tori can each be labelled by their action variables. The tori are null surfaces: the Poincaré integral $\int_A \mathrm{d}\boldsymbol{p}\cdot\mathrm{d}\boldsymbol{x} = \oint_{\partial A} \boldsymbol{p}\cdot\mathrm{d}\boldsymbol{x}$ vanishes over any region A on the torus whose boundary ∂A can be continuously contracted to a point. Also, the line integral of $\boldsymbol{p}\cdot\mathrm{d}\boldsymbol{x}$ around any two closed paths, related by a continuous 'rubber-band deformation' (à la Binney & Tremaine 1987), will give the same value. This defines the actions:

$$J_i = \frac{1}{2\pi}\oint_{P_i} \boldsymbol{p}\cdot\mathrm{d}\boldsymbol{x}, \tag{1.3}$$

where P_i is a path that cannot be continuously deformed into a point. For an N-torus, there are N such sets of paths; within each set, the paths can be deformed into each other, but the 'rubber band' cannot be smoothly shifted from one set into another. If the Hamilton-Jacobi equation is separable in some coordinates $\boldsymbol{q}$ (see e.g. appendix A), the integrand of (1.3) is just $p_i\,\mathrm{d}q_i$.

In the above way, an integrable Hamiltonian can be seen as a geometric concept: an everywhere dense set of invariant tori, bound together by the orbital frequencies that define the value of the Hamiltonian as a function of the actions. Locally, a torus for which all the abovementioned geometric properties hold does not care much about whether it has any neighbouring tori or not. This brings us to the concept of near-integrable Hamiltonians. In the much celebrated Kolmogorov-Arnold-Moser (KAM) theorem (see e.g. Lichtenberg & Lieberman 1983 or Arnold 1989) it was formally shown that when an integrable Hamiltonian is perturbed by a small term of the general (sufficiently smooth) form $H_1(\boldsymbol{J},\boldsymbol{\theta})$, most non-resonant invariant tori survive. The KAM theorem does not state how to actually determine the relationship between the measure of invariant tori and the perturbation parameter ϵ. In other words, the proof of the KAM theorem requires a very small value of ϵ to be chosen to ensure that the perturbation is sufficiently small to leave a given torus unbroken; it does not predict which value really destroys an invariant torus. However, from numerical studies, and especially plottings of Poincaré surfaces of section (such as intersections of the orbit with $y = 0$ while $\dot{y} > 0$, forming an $(x, \dot{x})$-plot, a convention used in this treatise), it is known that the KAM picture seems to be generally applicable with larger perturbation strengths as well. The previously simple geometric structure of phase space becomes suddenly exceedingly complex, as already noted by Poincaré (1894). In place of resonant tori, there are sets of 'higher-order' tori surrounding stable resonant orbits, the process repeats for these tori, and so on. Also, in conjunction with the separatrices connecting the unstable (hyperbolical) fixed points (in surfaces of section) that mark the boundaries of the sets of higher-order tori, there are always layers of stochasticity where adjacent orbits diverge exponentially. As one moves further away from integrability, these chaotic areas grow and merge with each other, and more and more invariant tori are destroyed in the process.

In surfaces of section, the manifestations of invariant tori are 'invariant curves': naturally, the resolution of the plot determines which curves can be considered 'invariant'. In theory, the hierarchy of higher-order tori and stochastic layers goes on indefinitely in smaller and smaller scales. Therefore, in this thesis, we follow (to the horror of pure mathematicians) the convention of calling 'computationally invariant' tori invariant. As we have seen, they are quite independent creatures, and thus locally mimic the properties of an integrable

Hamiltonian, never minding all the hullabaloo going on around them.

For clarity (hopefully), we distinguish here between two types of invariant tori and orbit families in the context of near-integrable Hamiltonians (cf. the surface of section plots in this thesis). It is perhaps easiest to imagine an integrable Hamiltonian that is then perturbed in the sense of the KAM theorem: while the original structure of phase space vanishes, invariant 'KAM-tori' still exist after the perturbation. Ones of the first type, called major invariant tori, can be understood as natural continuations of the tori of an unperturbed integrable potential: they are like distortions of the originals. The corresponding orbits in configuration space form major-orbit families, distinguished by the general form of their orbits or tori.

Ones of the second type we call minor invariant tori: these are the higher-order ones created by the perturbation and they do not have counterparts in an integrable Hamiltonian. They are associated with resonances in the integrable Hamiltonian, and in Poincaré surfaces of section they form 'islands' around a closed orbit. The corresponding orbits form minor-orbit families, usually labelled by the rotation number (resonance ratio) of the 'trapping' closed orbit.

One of the goals of this work is to do the above perturbation backwards: start with a given Hamiltonian H that looks like a near-integrable one, but for which no underlying integrable Hamiltonian H_0 is known. Intuitively, one can often clearly see what the major- and minor-orbit families are, and is tempted to speak of resonant trapping and such. Our task is to show that an underlying Hamiltonian H_0, supporting the same major-orbit families as H, can be explicitly created (even if it does not exist in a sort of mathematical Platonic universe, waiting to be found as the H_0 in the direct case would, if somebody had first perturbed it and then forgotten what it was). What is more, the minor-orbit families can then be seen as trapped by the resonant orbits of H_0.

Our problem of fitting an integrable Hamiltonian to any given one is intimately connected with the construction of invariant tori: it would be advantageous if the fitted Hamiltonian preserved accurately the 'surviving' invariant phase-space tori of a near-integrable original

Hamiltonian, so giving an action-angle representation for these tori is a natural starting point. This is an important problem in many fields ranging from plasma physics (Boozer 1982) and semiclassical quantum theory (Martens & Ezra 1987) to stellar dynamics (Binney & Spergel 1984, McGill & Binney 1990 - hereafter McG&B -, Binney & Kumar 1993, Ratcliff, Chang & Schwarzschild 1984) and accelerator physics (Warnock 1991, Warnock & Ruth 1992). Traditionally, invariant tori have been constructed perturbatively: a sequence of transformations is sought that continuously deforms some analytic torus into the required invariant torus. This approach, which culminated in the KAM theorem, is limited by the tendency of tori to change their form abruptly as the underlying dynamical problem is changed: these abrupt changes make the deformation inherently non-analytic and undermine the convergence of perturbative series. Consequently, most problems must be handled non-perturbatively.

Recently (McG&B, Warnock 1991) techniques have been developed for non-perturbatively determining the generating function of a canonical transformation that maps the invariant tori of an integrable Hamiltonian into approximately invariant tori of a given Hamiltonian. Approaches, such as these, that are based on generating functions, can be argued to be inherently superior (McG&B) to ones that directly construct approximate invariant tori. However, these approaches are liable to break down when the given Hamiltonian is far from any known integrable one. In fact, if one can use just a generating function (periodic in angle coordinates) one is lucky: generally, an additional point transformation is necessary to prepare some integrable Hamiltonian to be used for mapping with the original one, as will be discussed in the next chapter.

With the approach that is described in this thesis, one creates invariant tori closely approximating the existing ones of H, and fills the rest of the phase space by creating tori as parts of a global set in which there is a torus corresponding to any given actions $\boldsymbol{J}$, thus defining an integrable Hamiltonian $H_0(\boldsymbol{J})$. The scheme consists of three basic steps: (i) construct the invariant tori, labelled by actions $\boldsymbol{J}$, approximating those of H, (ii) construct consistent angle coordinates $\boldsymbol{\theta}$ for the created tori, and (iii) define how to generalize steps (i) and (ii) for values of $\boldsymbol{J}$ that do not correspond to invariant tori of H. For step (i), we give a method that is often applicable in step (iii) as well. This is very useful when creating tori for a regular grid of $\boldsymbol{J}$-values that may not all correspond to invariant tori of H.

1.2 Integrable Hamiltonians in stellar dynamics

Although the techniques presented in this thesis are generally applicable, we use gravitational potentials of astrophysical interest as case studies. Historically, the developments of Hamiltonian dynamics have always been closely connected with celestial mechanics and the dynamics of stellar and other many-body systems. This tradition still continues, so astrophysical problems usually provide application possibilities for almost any advances in the field of Hamiltonian dynamics.

The 'geometric' torus-fitting technique was originated in McG&B and Binney & Kumar (1993), and the astrophysical examples in these references concerned Hamiltonians governing motion in the meridional plane of axisymmetric potentials. These Hamiltonians are rather special in that, when they are integrable, they have only one orbit family, that of the loop orbits. More general separable potentials admit several major orbit families; de Zeeuw (1985) showed that a general Stäckel potential admits precisely the four major families that had already emerged from Schwarzschild's numerical (1979) study of orbits in a typical non-rotating triaxial potential.

A two-dimensional barred Stäckel potential admits two orbit families – those of the box and loop orbits – and these orbits occupy most of phase space in many non-rotating two-dimensional barred potentials. In this thesis, our examples are concerned with this case.

Innumerable problems in stellar dynamics could be more readily tackled if it were routinely possible to fit an integrable Hamiltonian to the Hamiltonian for stellar motion in an arbitrarily specified galactic potential; the development of such a technique would make available to galactic dynamics the powerful machinery of Hamiltonian perturbation theory, which was developed more than a century ago to handle the dynamics of the Solar system.

Hitherto the applicability of perturbation theory to galactic dynamics has been confined to nearly spherical (or planar and nearly axisymmetric) systems, since only for such nearly symmetrical systems has it been possible to approximate the Hamiltonian by an integrable one. None the less, the majority of what we know about the dynamics of galaxies has been learnt either from such studies or from N-body simulations. In particular the whole of the theory of spiral structure and most of the theory of the stability of spherical systems rely upon Hamiltonian perturbation theory.

A great many important problems in galactic dynamics involve large departures from sphericity, and are inaccessible to standard methods. The classical example is the dynamics of halo stars in the solar neighbourhood, which move far from the Galactic plane and thus respond significantly to the flattening of the Galactic potential. Other examples are provided by the motions of stars in strongly barred potentials. Awareness of the importance of this class of motion has been increasing in recent years as we become more aware of the frequency and dynamical importance of bars at the centres of galaxies such as our own (e.g. Blitz et al. 1993; Kormendy 1993), and of the triaxiality of the heavy haloes in which galaxies are believed to be embedded (Binney 1978; Kuijken & Tremaine 1993).

In addition to perturbative modeling of phenomena associated with resonances, the methods for torus construction can be used in creating models for steady-state stellar systems, which remains a fundamental problem of dynamical astronomy. In accordance with the Jeans theorem (Binney & Tremaine 1987, section 4), one can describe stellar systems using a phase-space distribution function $f(\boldsymbol{J})$ that depends only on the actions $\boldsymbol{J}$ (N-dimensional action space is a faithful representation of $2N$-dimensional phase space). Thus, if the locations of the tori $\boldsymbol{J}$ in phase space are known, one can recover the real-space densities $\rho(\boldsymbol{J};\boldsymbol{x})$ with the aid of the fact that, by the time averages theorem (Arnold 1989, Binney & Tremaine 1987), the probability of finding a star on a torus is uniform in the angle variables.

In this way, one can iteratively construct self-consistent potential-density-pairs corresponding to given distribution functions. Having such semianalytical models at our disposal would naturally be preferable to just watching N-body models dance on a computer screen (after all, important though they are, such simulations are not much more than speeded-up films when it comes to understanding what exactly is going on physically).

Finally, we mention another important dynamical consequence of the possibility of expressing the dynamics of a system in action-angle variables: under a suitably slow change of Hamiltonian, the actions remain constant, i.e., they are adiabatic invariants. This allows slow evolution of a system to be studied analytically. Historically, the concept of adiabatic invariance has played a very important role in the development of modern dynamics, starting with Einstein (Solvay conference 1911) and leading to several important principles, such as the 'WKB' and 'EBK' quantization conditions, secular perturbation theory and the 'KBM' method of averaging (e.g. Lichtenberg & Lieberman 1983).

Outline of the thesis

In view of the general applicability of the methods presented here, this thesis is meant to be quite self-contained and suitable for all readers possessing some knowledge of general Hamiltonian dynamics. The main emphasis is on astrophysical applications, but technically the astrophysical content is rather implicit; even if you do not have the faintest idea of what a galaxy looks like, you should be completely able to follow the discussion.

The thesis is organized as follows. In chapter 2 we explain how approximate invariant tori are generated for a given orbit family, and we employ the techniques in constructing tori for various points in the action space of a non-axisymmetric potential. In chapter 3 we show how the angle variables and frequencies are obtained for a given torus. We also describe the use of a torus construction method originated by Warnock (1991); this scheme is more limited than our geometric one, but it can obtain highly accurate approximations to the existing invariant tori of a given Hamiltonian. Also, we briefly discuss the computational efficiency of the methods in three degrees of freedom.

In chapter 4, we show how constructed tori can be employed to define an integrable Hamiltonian H_0 closely approximating the original one, and Hamiltonian perturbation theory is then used to demonstrate that a minor-orbit family can be treated as one made up of orbits trapped by a resonance of H_0. In chapter 5, we extend our methods to time-reversal asymmetric Hamiltonians (motion in a rotating frame of reference). In this connection, we study how tori can be constructed in the regions of phase-space where motion in the given Hamiltonian is chaotic. Also, we examine the transition from locally contained stochasticity to globally spread one in the light of the resonance overlap criterion. Finally, in chapter 6 we draw the conclusions, discuss the results and speculate a bit on future work. In four appendices, we present mainly mathematical material that is essential to the methods, but not necessarily needed in the main text.

Most of the material in chapters 2 to 4 is covered in the references Kaasalainen & Binney (1994a,b,c) and Kaasalainen (1994).

Chapter 2

Canonical transformations and torus construction

In this chapter, we view tori as geometric concepts. The basic idea is to map invariant tori in the phase space of a known integrable 'toy' Hamiltonian H_{T} to the phase space of a given 'target' Hamiltonian H with the aid of generating functions and point transformations; we employ the convention that primed variables relate to the tori of H, while the unprimed ones correspond to those of H_{T}. For H_{T}, the transformation between the ordinary phase-space coordinates $(\boldsymbol{x}, \boldsymbol{p})$ and the actions and angles $(\boldsymbol{J}, \boldsymbol{\theta})$ is known. Since the mappings are not obtained perturbatively, H_{T} need not closely resemble H. At first we study the case where we map the toy tori into existing invariant tori of H, i.e., H is either integrable or near-integrable, possessing 'KAM-tori' on which quasiperiodic motion is preserved. These tori can be labelled by action variables $\boldsymbol{J}'$, and ideally the torus-construction scheme 'finds' these target tori. Later (chapters 4 and 5) we shall show how to fill the rest of the phase space by creating tori as parts of a global set in which there is a torus corresponding to any given actions $\boldsymbol{J}'$, thus defining an integrable Hamiltonian $H_0(\boldsymbol{J}')$.

With the geometric approach, an invariant torus is constructed using purely its geometric characteristics:

(i) The torus is a null surface, i.e. Poincaré's integral $\int_A \mathrm{d}\boldsymbol{p}\cdot \mathrm{d}\boldsymbol{x}$ vanishes over any region A on the torus. This is automatically ensured because the transformation is canonical.

(ii) The torus lies entirely within a surface of constant H: any null torus on which H is constant is an orbital torus of H (McG&B). This fact can be employed in creating a practical procedure.

2.1 Generating functions

The first goal of our scheme is to seek a generating function $S(\boldsymbol{\theta},\boldsymbol{J}')$ that maps the tori of H_{T} into those of H. The generating function is given by

$$S(\boldsymbol{\theta},\boldsymbol{J}') = \boldsymbol{\theta}\cdot\boldsymbol{J}' - \mathrm{i}\sum_{\mathbf{n}\neq 0} S_{\mathbf{n}}(\boldsymbol{J}')\, \exp(\mathrm{i}\boldsymbol{n}\cdot\boldsymbol{\theta}). \tag{2.1}$$

When the real-valuedness of $\boldsymbol{J}$ is combined with time-reversibility of H, this simplifies to

$$S(\boldsymbol{\theta},\boldsymbol{J}') = \boldsymbol{\theta}\cdot\boldsymbol{J}' + 2\sum_{\mathbf{n}>0} S_{\mathbf{n}}(\boldsymbol{J}')\, \sin(\boldsymbol{n}\cdot\boldsymbol{\theta}), \tag{2.2}$$

where only half of $\boldsymbol{n}$-space is included in the sum because $S_{\mathbf{n}}$ is real and $S_{-\mathbf{n}} = -S_{\mathbf{n}}$. When the target Hamiltonian respects additional symmetries, (2.1) simplifies further. (Later we will drop the symmetry of time-reversibility; this will only make the series contain both sines and cosines.) The toy actions $\boldsymbol{J}$ are related to the target actions $\boldsymbol{J}'$ by

$$\boldsymbol{J} = \frac{\partial S(\boldsymbol{\theta},\boldsymbol{J}')}{\partial\boldsymbol{\theta}} = \boldsymbol{J}' + 2\sum_{\mathbf{n}>0} \boldsymbol{n}\, S_{\mathbf{n}}(\boldsymbol{J}')\, \cos(\boldsymbol{n}\cdot\boldsymbol{\theta}). \tag{2.3}$$

The coefficients $S_{\mathbf{n}}$ are found by nonlinear least-squares iteration (Levenberg-Marquardt algorithm, e.g. Press et al. 1986), where the objective is to minimize the variation of H, measured for a large set of points $\boldsymbol{\theta}$, along the trial torus labelled by $\boldsymbol{J}'$ (McG&B). If that torus is an existing invariant torus of H, the value of H on it will not vary after the iteration (except for computational noise). In addition to $S_{\mathbf{n}}$, the parameters defining the toy Hamiltonians play an important role, and they must be optimized together with the coefficients.

The main question is, of course: what is a suitable toy Hamiltonian for a given target one? In principle, the main criterion is that the orbits of H_{T} must be topologically similar to those of H. More precisely, (2.3) implies that each surface of constant $\boldsymbol{\theta}$ (and variable $\boldsymbol{J}$) cuts a given torus $\boldsymbol{J}'$ exactly once, otherwise the form $\boldsymbol{J} = \boldsymbol{J}(\boldsymbol{J}'; \boldsymbol{\theta})$ of (2.3) cannot be used. Thus the toy and target tori must be of similar types.

McG&B and Binney & Kumar (1993) studied cases where only one family of topologically similar orbits, and consequently only one toy Hamiltonian H_{T}, is needed. They used axisymmetric potentials; in these, the motion occurs at a constant z-axis angular momentum $L_z \equiv J_z'$ and can be followed from its projection onto the (r, z)-plane, the meridional plane. The major invariant tori are all of the same type, corresponding to one major-orbit family. A suitable toy Hamiltonian with similar tori is the spherically-symmetric isochrone Hamiltonian (see appendix A)

$$H_{\mathrm{I}} = \frac{1}{2}p^2 - \frac{k}{b + \sqrt{b^2 + (r - r_0)^2}}, \tag{2.4}$$

where k, b and r_0 are parameters.

In general, however, there are several orbit families, each needing toy tori that reproduce the characteristic topological properties of the family. Such Hamiltonians are produced e.g. by nonaxisymmetric planar potentials.

Two orbit families and H_{T}:s

We illustrate this case by using the potential

$$\Phi = \frac{1}{2} \ln \left(x^2 + \frac{y^2}{q^2} + R_{\mathrm{c}}^2 \right), \tag{2.5}$$

with $R_{\mathrm{c}} = 0.14$. Motion in this potential, representing that of an elongated galaxy, has been studied by Richstone (1984), Binney & Spergel (1984), Levison & Richstone (1987), Miralda & Schwarzschild (1989) and others.

A typical $(x, \dot{x})$ surface of section is shown in Fig. 2.1 ($q = 0.8$, $H = -0.2$). Two major types of orbits are dominant: those that have a constant sense of circulation around the origin ('loops') and those that have not ('boxes'); their invariant curves circulate around $(x, \dot{x}) \approx (0.4, 0)$ and $(0, 0)$, respectively. Similar orbits are supported by Stäckel potentials (appendix B, Binney & Tremaine 1987, de Zeeuw 1985), whose Hamilton-Jacobi equations separate. In addition to these, a minor family of orbits appears as a pair of islands. These orbits may be considered to be trapped by a 3:2 resonance in a perturbed Stäckel potential. As q is decreased from unity and the potential becomes more elongated, more and more of phase space is occupied by similar families. When talking about a $r : s$ resonance, we shall follow the convention that $T_1/T_2 = \omega_2/\omega_1 = r/s$, where $\boldsymbol{T}$, $\boldsymbol{\omega}$ are, respectively, the fundamental periods and frequencies of the orbit in the underlying integrable Hamiltonian. In the surface of section (θ_1', J_1'), θ_2' =const., this resonance produces r islands; in (θ_2', J_2'), θ_1' =const., s islands. For example, the $(x, \dot{x})$ and $(y, \dot{y})$ surfaces of section of a box orbit have, respectively, r and s islands for one integrated orbit. The islands of Fig. 2.1 correspond to a doublet of 3:2 orbits.

As can be seen from e.g. Fig. 2.1, boxes and loops occupy fundamentally different parts of phase space. Consequently, one cannot map loop-type tori into box-type ones or vice versa, so we must have a way of generating two types of toy tori. By mapping the tori of a Stäckel potential, one could, in principle, recover the invariant tori of both major orbit families from a single toy Hamiltonian H_{T}. However, since analytic expressions are not available for the angle variables of Stäckel tori, it is more convenient to obtain the box and loop tori as images of the tori of two different toy Hamiltonians, whose orbits topologically resemble boxes and loops; we stress the fact that quantitatively they need not resemble H closely.

For loops, we have used the isochrone Hamiltonian H_{I}, and for boxes the two-dimensional harmonic oscillator

$$H_{\mathrm{H}}(\boldsymbol{x}, \boldsymbol{p}) = \frac{1}{2} \sum_{i=1}^{2} (p_i^2 + \omega_i^2 x_i^2), \tag{2.6}$$

where the ω_i are suitable frequencies. In terms of action-angle coordinates $(\boldsymbol{J}, \boldsymbol{\theta})$, we have

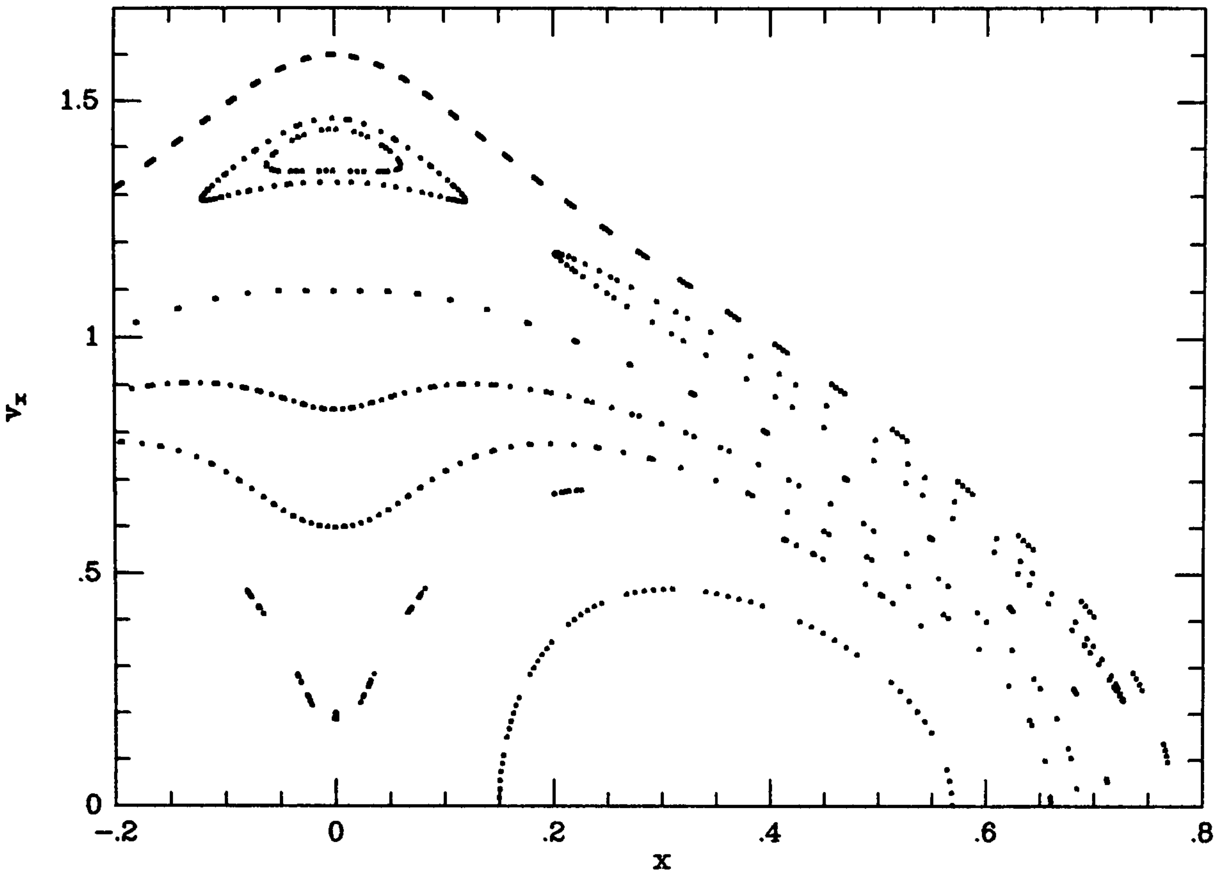

Figure 2.1: Invariant curves of box- and loop-orbits and 3:2 islands.

(e.g. Lichtenberg & Lieberman 1983)

$$\begin{aligned} H_{\mathrm{H}} &= \sum \omega_i J_i, \\ x_i = \sqrt{\frac{2J_i}{\omega_i}}\sin\theta_i, &\qquad p_i = \sqrt{2J_i\omega_i}\cos\theta_i. \end{aligned} \tag{2.7}$$

In the target potential Φ, the equation of motion, $\dot{J}_i = -(\partial H/\partial\theta_i)$, of a box-type toy action variable reads

$$\begin{aligned} \dot{J}_i &= -\Big(\boldsymbol{p}\cdot\frac{\partial \boldsymbol{p}}{\partial\theta_i} + \frac{\partial\Phi}{\partial\boldsymbol{x}}\cdot\frac{\partial\boldsymbol{x}}{\partial\theta_i}\Big) \\ &= p_i\,\big(\omega_i x_i - \frac{1}{\omega_i}\frac{\partial\Phi}{\partial x_i}\big). \end{aligned} \tag{2.8}$$

Hence $\dot{J}_i$ vanishes when either $p_i = 0$ or $\omega_i^2 x_i = \partial\Phi/\partial x_i$. If the potential is symmetric in x_i, the latter condition holds at $x_i = 0$. Thus $\dot{J}_i = 0$ at $\theta_i = m\pi/2$. By the argument given in appendix B of McG&B, it now follows that all $S_{\mathbf{n}}$ with n_i an odd integer vanish. Because of the symmetries, only a part of the complete $2\pi \times 2\pi$-space of toy angles need be explored:

if necessary, the rest of the $\boldsymbol{\theta}$-space can be filled using that part. For the time-symmetric box-orbit case with potential symmetric in both coordinates, the patch $[0, \pi/2] \times [0, \pi]$ is sufficient. This, of course, is an important factor in reducing computational effort.

2.2 Point transformations

In this section, we discuss some situations where point transformations are necessary. As mentioned in Introduction, point transformations are, in fact, generally essential and the first thing to be considered in practice. The discussion was started with generating functions mainly because it seems to be more convenient to start with them and then add point transformations (which are, of course, also derivable from generating functions) in hindsight. First we consider the box orbits.

Although the tori of box orbits can in principle be recovered by the straightforward scheme of McG&B, the first few experiments along these lines showed that this procedure is inefficient: many box orbits are physically very different from the rectangular boxes formed by the toy orbits in configuration space, having narrow waists. To form these waists by the scheme of McG&B, a great many terms $S_{\mathbf{n}}$ must be included in the generating function, and this makes the calculation costly and cumbersome. Moreover, when the cosine-dependent sum in (2.3) becomes large, meaningless negative values of J_i are likely to be reported during the iteration. Steps in which such negative values are encountered have to be abandoned, and this further slows the calculation.

An efficient resolution of this difficulty is to compound the transformation (2.3) with a point transformation. The latter is motivated by the fact that box orbits are nearly (and with Stäckel potentials exactly) bounded by constant values of confocal elliptic coordinates, in the same way that harmonic-oscillator orbits are confined by constant values of Cartesian coordinates (see appendix B).

In a point transformation $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$ between two vectors expressed in arbitrary coordinate systems, the mapping of the momenta is obtained using $p_i \equiv \partial\mathcal{L}/\partial\dot{x}_i$ ($\mathcal{L}$ is the Lagrangian) and the chain rule. Since $(\partial\dot{x}_j/\partial\dot{x}'_i) = (\partial x_j/\partial x'_i)$, we have that, for any point transformation,

$$p'_i = \frac{\partial \boldsymbol{x}}{\partial x'_i} \cdot \boldsymbol{p}. \tag{2.9}$$

Let (u, v) be confocal elliptic coordinates for the (x', y') plane. In terms of these coordinates the target Hamiltonian is

$$H(u, v, p_u, p_v) = \frac{1}{2G\Delta^2}(p_u^2 + p_v^2) + \Phi(x'(u, v), y'(u, v)), \tag{2.10}$$

where

$$G \equiv \sinh^2 u + \cos^2 v. \tag{2.11}$$

When a particle moves on a box orbit in a Stäckel potential $\Phi(x', y')$ that has the same focal distance Δ as the coordinate system, the particle explores a rectangle in the (u, v) plane. Conversely, if we cause the particle to move over a rectangle in the (u, v) plane by immersing it in a potential quadratic in

$$\begin{aligned} x &\equiv f(u), \\ y &\equiv g(v), \end{aligned} \tag{2.12}$$

where f and g are odd functions, then the image of its motion in the (x', y') plane will look like a box orbit in a Stäckel potential. This suggests that a useful point transformation is of the form (2.12). Using (2.9), we have

$$\begin{aligned} p_u(x, p_x) &= \left.\frac{\mathrm{d}f}{\mathrm{d}u}\right|_{u(x)} p_x, \\ p_v(y, p_y) &= \left.\frac{\mathrm{d}g}{\mathrm{d}v}\right|_{v(x)} p_y. \end{aligned} \tag{2.13}$$

The torus-finding procedure is now the following. Given a value for $\boldsymbol{J}'$, we use equation (2.3) to evaluate $\boldsymbol{J}$ for the first point on our grid of $\boldsymbol{\theta}$ values. Inserting $(\boldsymbol{J}, \boldsymbol{\theta})$ into equations (2.7), we evaluate $(\boldsymbol{x}, \boldsymbol{p})$, and inserting these values into (2.12) and (2.13) we obtain

(u, v, p_u, p_v). Finally, equation (2.10) yields $H(\boldsymbol{J}', \boldsymbol{\theta})$. This procedure is then repeated for each value of $\boldsymbol{\theta}$ on our grid, and in this way we evaluate the variance of H on the trial torus. Then the $S_{\mathbf{n}}$ and the functions f and g are adjusted so as to minimize this variance, and $H(\boldsymbol{J}', \boldsymbol{\theta})$ is re-evaluated around the torus.

The initial value of Δ in the iteration can be chosen by integrating a few orbits in the target potential and finding the elliptic coordinates that best fit these orbits. In addition to describing properly the geometry of the system, (2.12) optimizes the structure of the generating function S: for Stäckel potentials, all the coefficients $S_{\mathbf{n}}$ with both $n_i \neq 0$ vanish because $p_u = p_u(u)$ and $p_v = p_v(v)$. Thus the greatest weight in S in Stäckel-like potentials lies on the coefficients S_{i0} and S_{0j}.

We have found the most useful choices for the functions f and g to be $f(u) = X \sinh au$ and $g(v) = Yv$ or $Y \sin v$, where X and Y are varied with the $S_{\mathbf{n}}$, and a is fixed. The parameter a (typically between 1 and 1.5) mainly affects the shape of the torus obtained with $S \equiv 0$, and the distribution of the sample points on the torus used in the minimization procedure. When the actions lie close to the box-loop frontier, a often has to be larger than when they are further away from the frontier.

Since $|v| \leq \pi/2$ or $|\sin v| \leq 1$, the minimization procedure must correct for steps that lead outside the allowed range; usually this happens only in the beginning of the iteration, if at all. The choices for g given above seem to be more efficient than functions with infinite ranges. Varying X and Y is equivalent to varying the frequencies ω_i, so ω_i should be kept fixed during iteration when optimizing X and Y.

Equation (2.3) sometimes yields negative values for J_i even when the basic scheme of McG&B is compounded with the point transformation just described – this problem arises when seeking the tori of orbits that lie near or beyond the box-loop transition. However, when a point transformation is employed, the absolute value of any negative J_i tends to be small, and the iterations can usually be successfully continued by employing $|J_i|$ in place of J_i.

Box orbits in the toy potential actually consist of motion in two independent one-dimensional potentials; since systems with one degree of freedom are always integrable, the choice of the two potentials is not limited to the harmonic oscillator. For any one-dimensional potential the relation between the action and angle variables and a phase-space point can always be found, at least numerically. We have used the harmonic oscillator because then the relation is simple and analytic, and because it has worked well in the test cases.

To emphasize the usefulness of point transformations, it is perhaps worth noting that for Stäckel potentials the torus fitting for both boxes and loops can, in fact, be done employing only the transformation (2.12) without the generating function S (for loops, the polar coordinates r, φ are used instead of x, y). Using the Stäckel integrals E, I_2 corresponding to a pair of actions, one can compute $p_u(u)$ and $p_v(v)$ and thus integrate the functions $f(u)$ and $g(v)$ numerically from (2.12) and (2.13), with the toy actions being the same as the target actions. Thus there the generating function S is superfluous, and the target angles coincide with the toy angles. This fact can act as a guide to the optimization of the functions f and g.

Loop orbits with small J_r'

For the majority of loop orbits, the coefficients $S_{\mathbf{n}}$ can be obtained by mapping invariant tori of the isochrone Hamiltonian H_{I}, very much as described in McG&B. Using polar coordinates for the plane, a general phase-space point is $\boldsymbol{w} \equiv (r, \varphi, p_r, p_\varphi)$, and the action-angle coordinates are $(J_r, J_\varphi, \theta_r, \theta_\varphi)$. For completeness, the transformations $\boldsymbol{w} \leftrightarrow (\boldsymbol{J}, \boldsymbol{\theta})$ for the full three-dimensional isochrone potential are given in appendix A (a generalization of Saha 1989 and Goldstein 1980). As explained in appendix B of McG&B, all $S_{\mathbf{n}}$ with n_φ an odd number vanish when the target potential is symmetric in each quadrant.

The standard scheme of McG&B works for the majority of loop orbits, but it necessarily fails for orbits near any closed loop orbit $J_r' = 0$. The problem is that, under the transformation (2.3), $J_r' = 0 \Rightarrow J_r = 0$: the term periodic in $\boldsymbol{\theta}$ has zero mean, and, since $J_r \geq 0$, must vanish whenever $J_r' = 0$. But $J_r = 0$ in turn implies that $p_r = 0$ and thus that r is constant. Hence any transformation of the type (2.3) maps all closed loop orbits $J_r' = 0$ into circles rather than into the ellipses formed by closed orbits in the target potential. A similar problem plagues the recovery of 'shell orbits' in the flattened axisymmetric potentials of McG&B and Binney & Kumar (1993), and is discussed in Kaasalainen & Binney (1994b).

Our resolution of this difficulty is to compound the transformation (2.3) with a point transformation which maps a circular orbit in $H_{\rm I}$ into the closed loop orbit in H that has the same value of J_φ. Let (r, φ) be the polar coordinates in the plane of the toy potential and (u, v) the confocal elliptic coordinates for the plane of the target potential, with a suitable focal distance Δ. (In the following, one can also substitute $u \longrightarrow r'$ and $v \longrightarrow \varphi'$ in coordinates and momenta, (r', φ') being the polar coordinates in the target plane. Usually elliptic coordinates best reflect the geometry of the system.)

While there are many possible forms for the transformation, we have found practical the form

$$r = x(v)u \ , \ \tan\varphi = \sin v \, \frac{y(u)}{z(v)}, \tag{2.14}$$

where x, y, z are functions to be determined. z is a cosine-like function; especially, $z(\pm\pi/2) = 0$. If the target potential is a Stäckel one, x and y are constants, and only z needs to be solved for. Using (2.9), we obtain

$$\begin{aligned} p_u &= x p_r + \frac{\mathrm{d}y}{\mathrm{d}u} \frac{z}{f(u,v)} \sin v \, p_\varphi, \\ p_v &= \frac{\mathrm{d}x}{\mathrm{d}v} u p_r + \frac{y}{f(u,v)} \Big(z \cos v - \frac{\mathrm{d}z}{\mathrm{d}v} \sin v \Big) p_\varphi, \end{aligned} \tag{2.15}$$

where

$$f(u, v) \equiv z^2 + y^2 \sin^2 v. \tag{2.16}$$

The image $[u_0(v), \, p_u(v), \, p_v(v)]$ of the closed orbit $[r = a, \, p_r = 0, p_\varphi = J_\varphi' \equiv (2\pi)^{-1} \oint \boldsymbol{p}' \cdot \mathrm{d}\boldsymbol{q}']$ in the toy potential is found by direct integration of the equations of motion. From (2.14) we now obtain x explicitly:

$$x(v) = a/u_0(v). \tag{2.17}$$

The equations (2.15) constitute a pair of differential equations for y and z. We define $\tilde{y}(v) \equiv y(u_0(v))$, so that there is only one variable v, and numerically integrate the pair

$$\begin{aligned} \frac{\mathrm{d}\tilde{y}}{\mathrm{d}v} &= \frac{p_u}{p_\varphi}\frac{\mathrm{d}u_0}{\mathrm{d}v}\Big(\frac{z}{\sin v} + \frac{\tilde{y}^2}{z}\sin v\Big), \\ \frac{\mathrm{d}z}{\mathrm{d}v} &= -\frac{p_v}{p_\varphi}\Big(\frac{z^2}{\tilde{y}\sin v} + \tilde{y}\sin v\Big) + z\cot v\ . \end{aligned} \tag{2.18}$$

(Both derivatives vanish at $v = 0$.) From (2.15) we have the initial condition $\tilde{y}(0)/z(0) = p_v(0)/p_\varphi$; the other initial condition is only a scaling convention and can be taken to be $\tilde{y}(0) = 1$. Note that y and z are completely determined by the target potential only, and the toy potential has an effect only on a in $x = a/u_0$; thus $y(u)$ and $z(v)$ need to be determined only once, and x is easily changed when u_0 is stored. When integrating (2.18), the result is consistent ($z(\pm\pi/2) = 0$) only when p_φ is correct, i.e., it equals J'_φ, the azimuthal action of the target torus.

The radius a of the toy circular orbit of angular momentum p_φ in the isochrone potential

$$\Phi(r) = -\frac{k}{b + \sqrt{b^2 + (r - r_0)^2}} \tag{2.19}$$

is given by

$$(a - r_0)^2 = \frac{(p_\varphi + \sqrt{4bk + p_\varphi^2})^4}{16k^2} - b^2, \tag{2.20}$$

where the potential parameters k, b, r_0 can be found by seeking the best fit to the target potential in a suitable region, or they can be some values obtained with the usual torus-fitting procedure close to the region in which the additional transformation becomes necessary. Also, it is convenient for a to have approximately the average value of u_0.

When $J'_r > 0$, $y(u)$ must be extrapolated beyond the range in u explored by the closed orbit. A practical way is to attach suitable polynomials to both ends of $y(u)$; they must be of at least second order in u, because continuous second derivatives are needed in the minimization procedure. These polynomials must be such that unique solutions are obtained when solving for u and v from (2.14) with given r and φ.

We have also experimented with a general point transformation between polar and elliptic coordinates (analogous to that between Cartesian and elliptic ones used for the box orbits), although there is no immediate geometrical need for it. For the loop orbits, such a transformation does not seem to be particularly advantageous, especially because it introduces more parameters.

Examples

As in McG&B, we tested the apparatus on Stäckel potentials. For these potentials one can calculate the $S_{\mathbf{n}}$ independently of the torus-fitting machinery, enabling the latter to be rigorously tested. The relations $p_u(u)$ and $p_v(v)$ (u, v being the elliptic coordinates; see appendix B) establish a phase-space torus in terms of u and v; the momenta p_u, p_v and the possible values for u, v are determined by the values of the integrals E and I_2. This torus can then be expressed in the action-angle variables of a toy Hamiltonian, i. e. one obtains relations $[\boldsymbol{J}(u,v), \boldsymbol{\theta}(u,v)]$. Thus, using e. g. a Newton-Raphson procedure, $\boldsymbol{J}$ can be determined on an equally-spaced $\boldsymbol{\theta}$-grid. Two discrete Fourier transforms give now the values of $\boldsymbol{n}\, S_{\mathbf{n}}(\boldsymbol{J}')$ from (2.3).

In our test cases, the structure of S was quite simple in the sense that the magnitudes of the coefficients $S_{\mathbf{n}}$ usually decrease in a regular manner with increasing $|\,\boldsymbol{n}\,|$. Thus it is easy to construct a starting S (containing usually about 60 terms and orders to 20), and from the results of successive trial iterations to deduce whether to include coefficients of higher order in either direction. As mentioned in McG&B, it is essential to 'oversample' the trial torus: one often needs at least 35 points in one quadrant of one angle to cover the torus properly. Also, after the convergence has saturated with one sampling grid, it is sometimes useful to continue with a denser grid. Loop-orbit tori generally converge within about ten steps, while box-orbit tori can require tens of steps for convergence.

The introduction of the above described point transformation for boxes has the effect of introducing local minima in χ^2; the latter were surprisingly little in evidence in McG&B and are slightly more troublesome now. Usually a few exploratory iterations are required to find the most useful $S_{\mathbf{n}}$ to include in the generating function, and to test different starting values of the other parameters (k, b, r_0 in the case of loop orbits and X, Y, Δ in the case of box orbits). The best results are obtained when the values of these parameters do not change much as the generating function converges: if a parameter changes by, say, a factor 2, a better final torus can usually be obtained by starting again with a better value of that parameter.

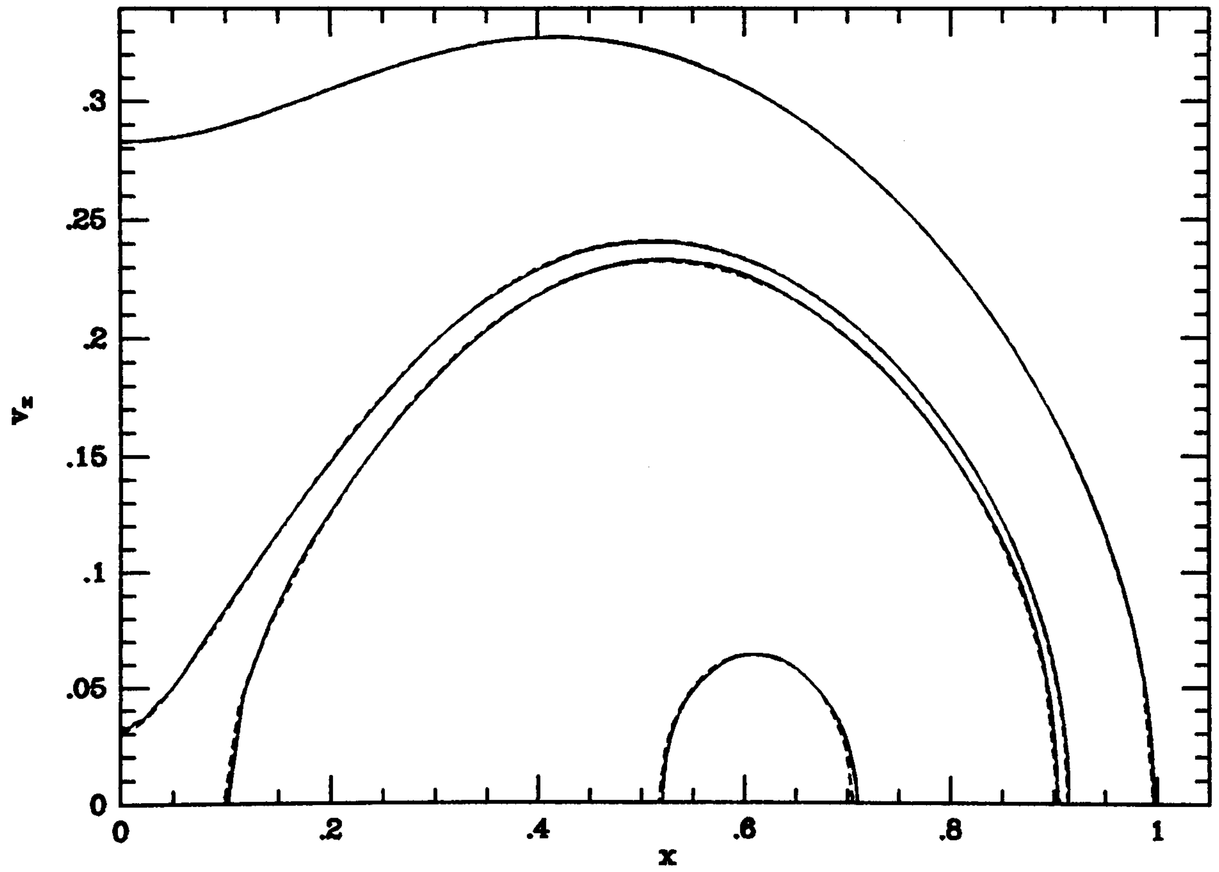

Figure 2.2: $(x, \dot{x})$ surface of section for a Stäckel potential at $H = -0.486$. The full curves show the sections of numerically fitted tori, while the dashed lines show analytically obtained sections.

In the case of Stäckel potentials, the fitting procedure typically gave the correct energy to an accuracy of one part in 10^4 or better, and the rms variation of the Hamiltonian was usually 10^{-4} or smaller. What is more, the focal distance Δ, used as a parameter in the point transformation for boxes, converged to the correct value to an accuracy of at least one part in 10^3.

To demonstrate that the algorithm works well in all parts of action space, we show two surfaces of section in two different potentials. First, consider the Stäckel potential of section 3.5 of Binney & Tremaine (1987). In Fig. 2.2, results for four orbits are shown in the $(x, \dot{x})$ surface of section at $H = -0.486$. In each case the full curves are obtained by torus fitting, while the dashed lines are sections obtained analytically. The uppermost box orbit is evaluated at $\boldsymbol{J}' = (0.1732, 0.1556)$, corresponding to the value -0.04 of the Stäckel integral I_2; the dashed line exactly overlies the full curve.

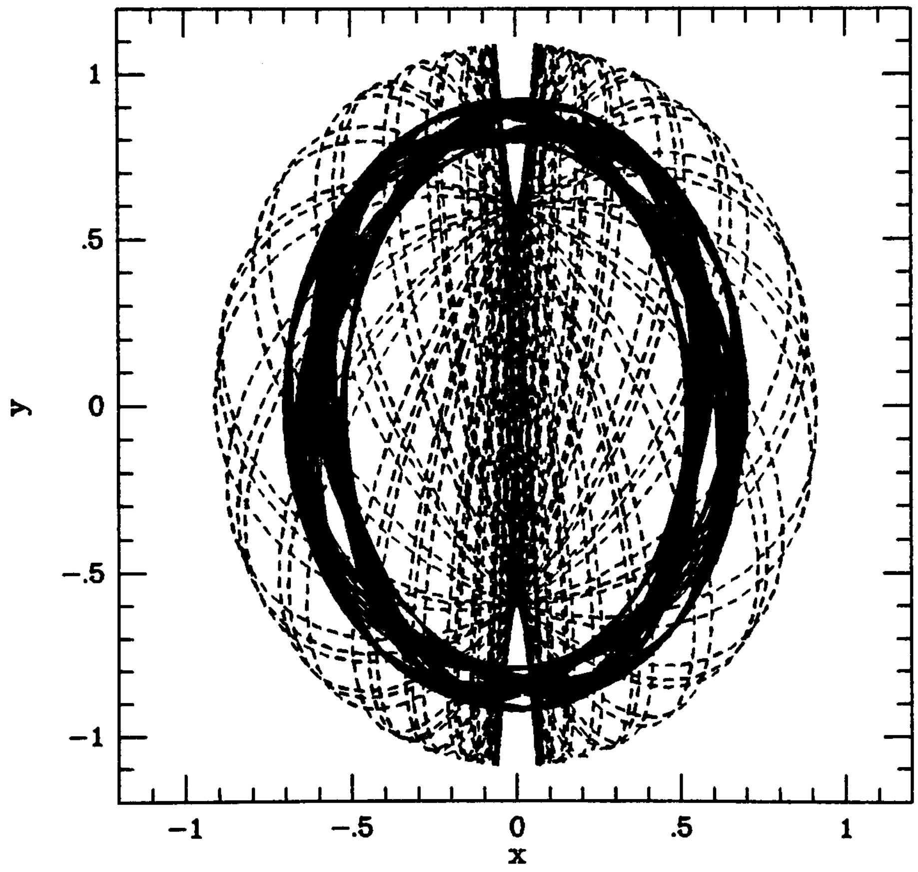

Figure 2.3: The appearances in real space of the lowest loop orbit and the middle box orbit of Fig. 2.2.

The next three orbits lie in the parts of the action space that offer the most difficult challenges: the transition zone between the two orbit families, and loops with small radial actions. Here, the box and the loop orbits in the transition zone correspond to $\boldsymbol{J}' = (0.0998, 0.2217)$ and $\boldsymbol{J}' = (0.0891, 0.2317)$ ($I_2 = -0.0005$ and $I_2 = 0.003$), respectively. The lowest loop orbit at $\boldsymbol{J}' = (0.006, 0.3061)$ ($I_2 = 0.05$) was constructed with the aid of the point transformation for loops with small J_r'. In Fig. 2.3, this loop orbit is shown in configuration space together with the middle box orbit.

When approaching the box-loop frontier, the surfaces of section for box orbits change rapidly near the $\dot{x}$-axis, $\dot{x}$ tending to zero at $x = 0$ before the transition to loops. Especially here a proper choice of the box point transformation function $f(u)$ is very important, because the section of the toy torus behaves very differently, its $\dot{x}$ always increasing monotonically as $|x|$ decreases.

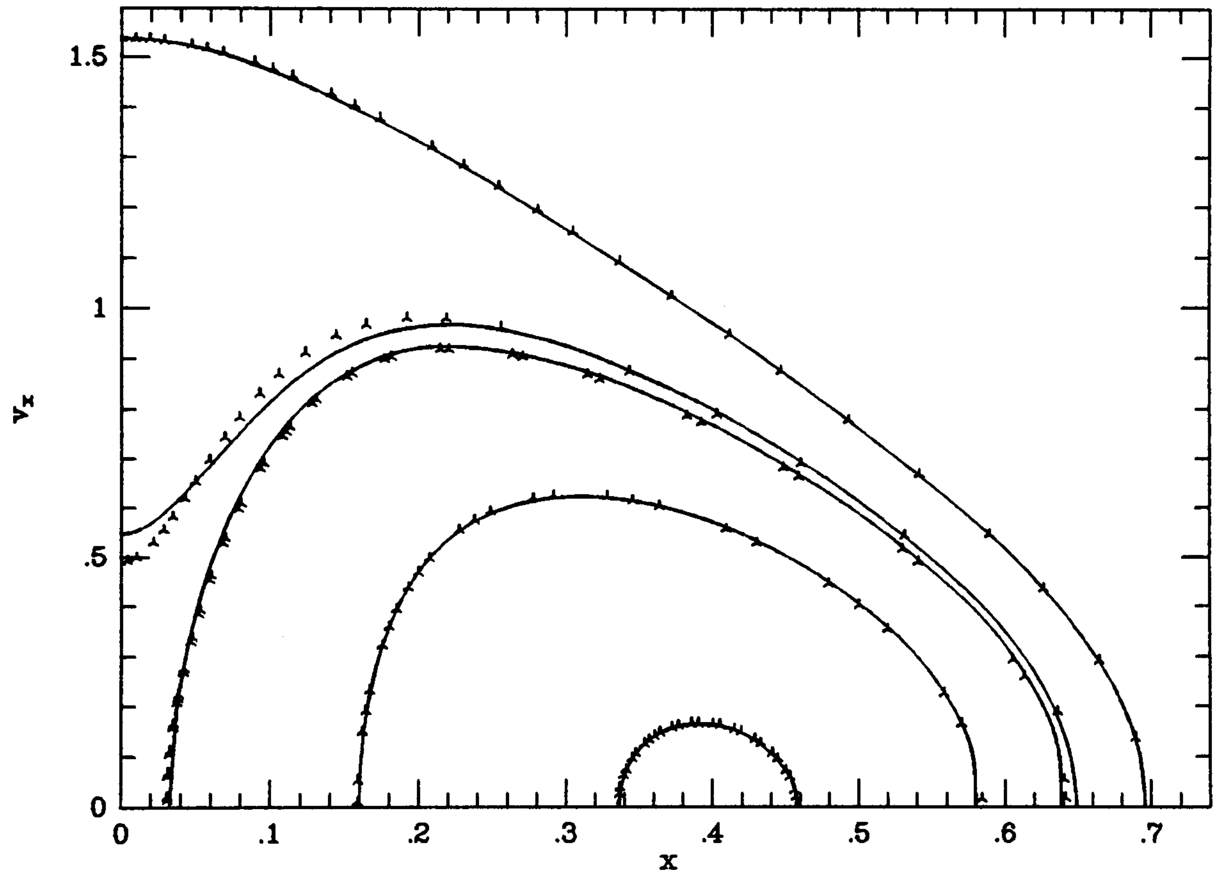

Figure 2.4: As Fig. 2.2, but the section is for the logarithmic potential with $q = 0.9$ at $H = -0.315$. The dots show consequents obtained by direct orbit integration starting from a point on each curve.

Next, we study the orbits in the logarithmic potential (2.5). We set $R_c = 0.14$ and varied q as in Binney & Spergel (1984; hereafter B&S), where the structure of action space is described for $q = 0.9$ and $q = 0.6$ (mildly and strongly barred potentials, respectively). In particular, the action diagrams in B&S, which were computed by the spectral method, can be used for choosing test orbits. As in B&S, we double the J_r-values of all loop orbits to ensure that the action spaces of the loops and the boxes fit neatly together to form the single action space explored in B&S. Here we consider the mildly barred nearly separable case; in chapter 4, we will study a strongly barred and non-separable potential.

In Fig. 2.4, a surface of section is shown for five orbits in the logarithmic potential with $q = 0.9$ at $H \approx -0.315$. Qualitatively, the orbits are similar to those in Fig. 2.2. Here, the points are consequents obtained by integrating the orbit from one point on the curve. The rms variation of the Hamiltonian along the trial torus was typically 10^{-3} or smaller. Starting from the uppermost box orbit, the actions are (0.450,0.040), (0.300,0.160), (0.266,0.190), (0.130,0.300), and (0.010.0.400). Again, the lowest loop was constructed using the point transformation for small J_r'.

Interface between the orbit families

For a given potential, one generally does not know a priori exactly where in action space the box-loop frontier lies. Therefore, near the transition line, torus construction must be carried out with both types of toy tori, and one fit chosen.

For a Stäckel potential, the location of the frontier is exactly known and thus we can rigorously test the performance of the torus-fitting machinery in the transition zone and establish criteria for choosing the correct family. It is natural to choose the fit that gives the smallest rms variation in the Hamiltonian, and for most orbits this criterion is usually sufficient. If one does not consider the fit obtained with one toy torus to be clearly better than that obtained with the other, comparing the surfaces of section can clarify the matter: the mutual consistency of the consequents obtained by integrating from various points on the curve obtained by torus fitting will often reveal the correct orbit family.

As an example, we show in Fig. 2.5 the results of using the wrong toy torus for the two orbits near the box-loop frontier that are shown in Fig. 2.2. For these fits, the rms variation of the Hamiltonian along the torus is a few times larger than with the correct toy torus, and in any case from the sections one can clearly see which toy tori fit the orbits best.

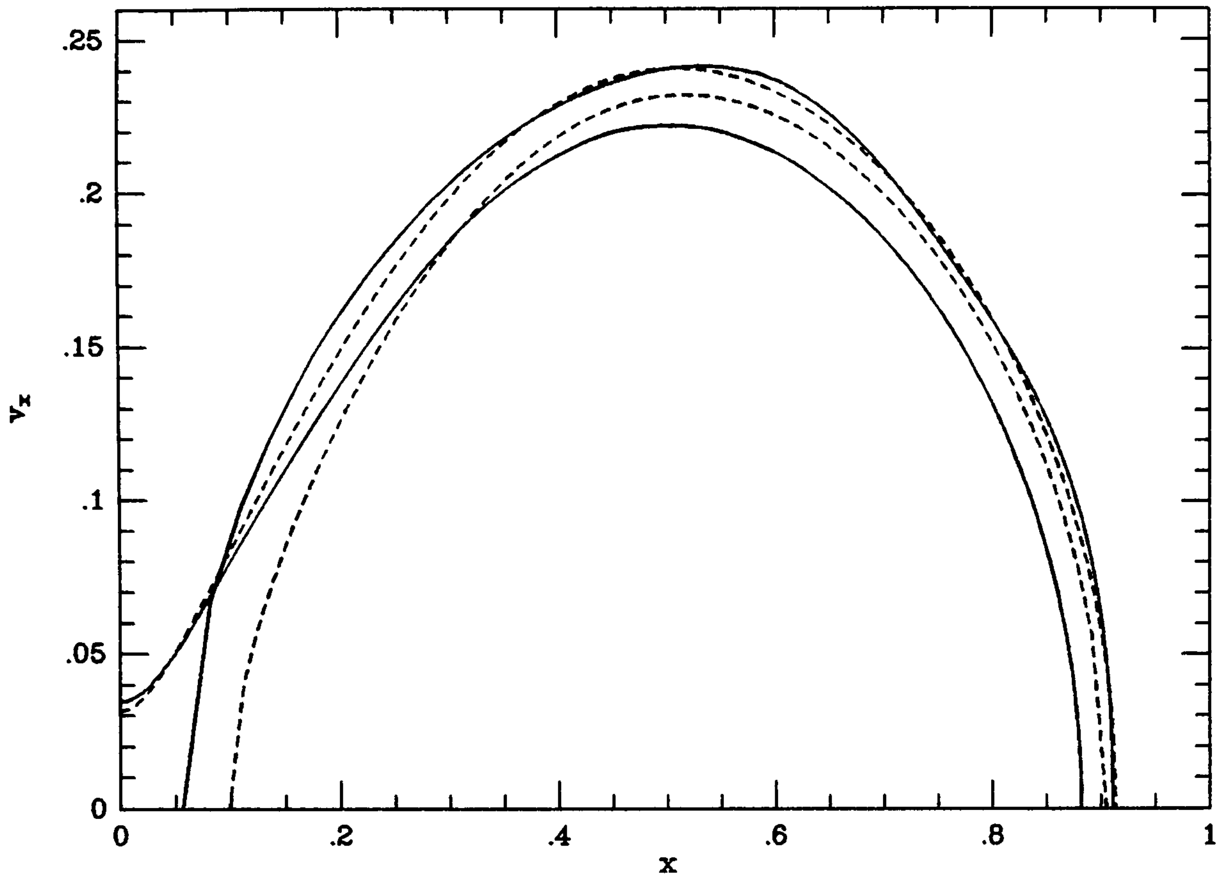

Figure 2.5: The same surface of section as for the two centremost orbits in Fig. 2.2, but now the tori are constructed with the wrong toy potentials.

Using the above criteria, we find that only very close to the frontier – roughly when $|I_2| \leq 0.0001$ – is there any uncertainty as to the correct orbit family from the torus fits. In ambiguous cases the two tori project to similar figures in configuration space, and differ principally in that the loop torus has a unique sense of rotation, while the box torus does not. In nearly but not precisely integrable potentials, these ambiguous orbits are generally stochastic, alternating erratically between box- and loop-like behaviours. We conclude that the torus machinery generates as clean a transition between boxes and loops as one is likely to observe in the real world.

Tori for minor-orbit families

Previously, we specifically tailored a point transformation to be used for orbits in the vicinity of a closed loop orbit. The loop orbits belong to a major-orbit family, but otherwise, in a surface of section, their invariant curves look like any islands around the single point(s) of a closed orbit. Therefore, it would seem natural to seek a generalized point transformation such that one could construct the invariant tori ('major' or 'minor') associated with any closed orbit. What is more, the motion of the orbits of a minor-type family is either librating (the banana orbits; see Fig. 2.6 and the corresponding surfaces of section Fig. 2.7 and Fig. 4.4) or circulating (the fish orbits; see Fig. 4.6 and the surface of section Fig. 2.1), just like with box and loop orbits, so the harmonic oscillator $H_{\rm H}$ and the isochrone $H_{\rm I}$ would be natural choices for the toy Hamiltonians also in this case.

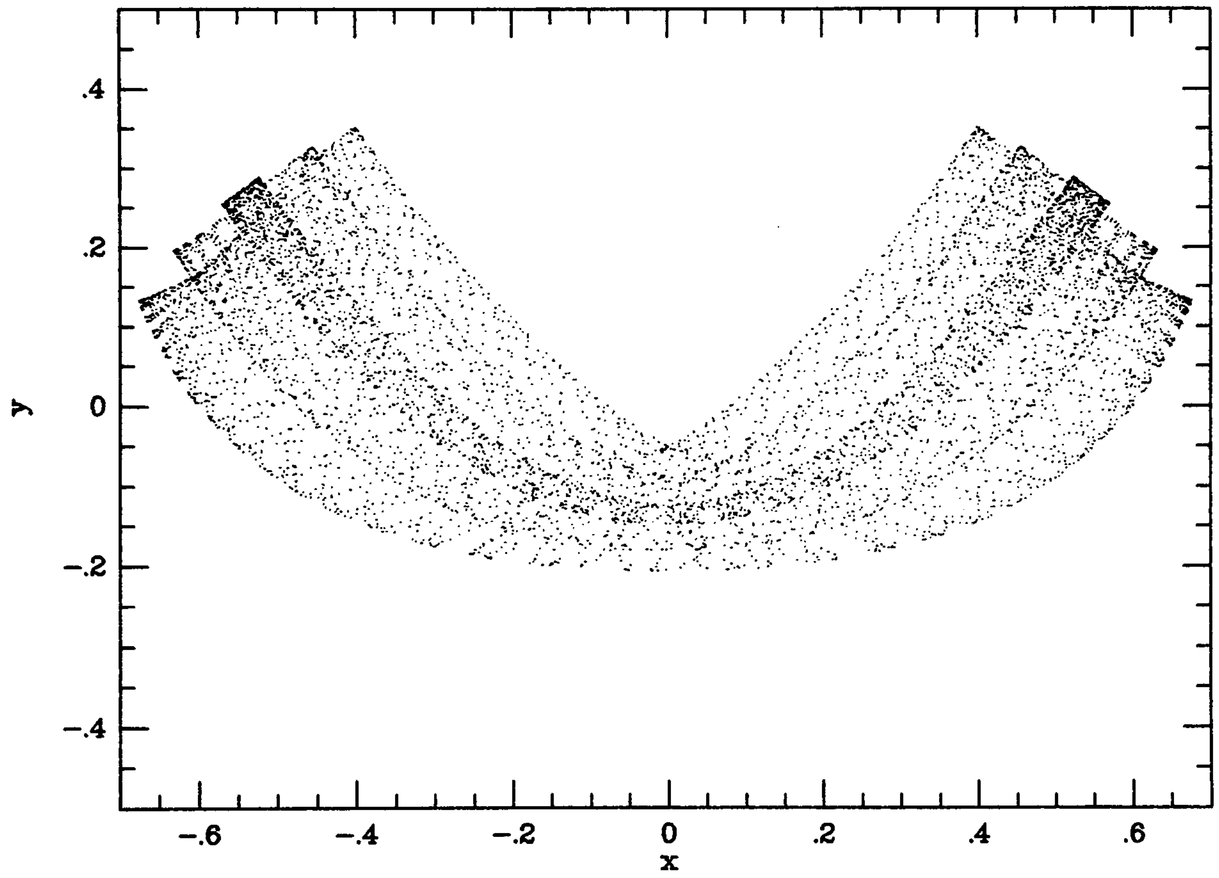

Figure 2.6: Three superposed 'banana' orbits in the logarithmic potential.

The basic principle is, then, first to define a point transformation such that a closed orbit in $H_{\rm T}$ is mapped to the closed orbit in the target Hamiltonian H. After this, one can use the torus-fitting scheme as above, provided the target tori and the transformed toy tori are of similar types. The main question is the choice of coordinates. For a closed orbit, one of the coordinates in the toy configuration space is fixed and the corresponding action J_i is zero. When $J_i > 0$, the variations in the i:th coordinate describe replacements perpendicular to the closed orbit. Correspondingly, it would be good to use a similar 'local' coordinate system in the target configuration space, where one coordinate always gives the perpendicular distance to the chosen closed orbit. In appendix D, we describe such a system and the corresponding point transformations. Such a system is arguably the best possible for the job at hand, mainly because it is a natural distortion of the toy coordinate system.

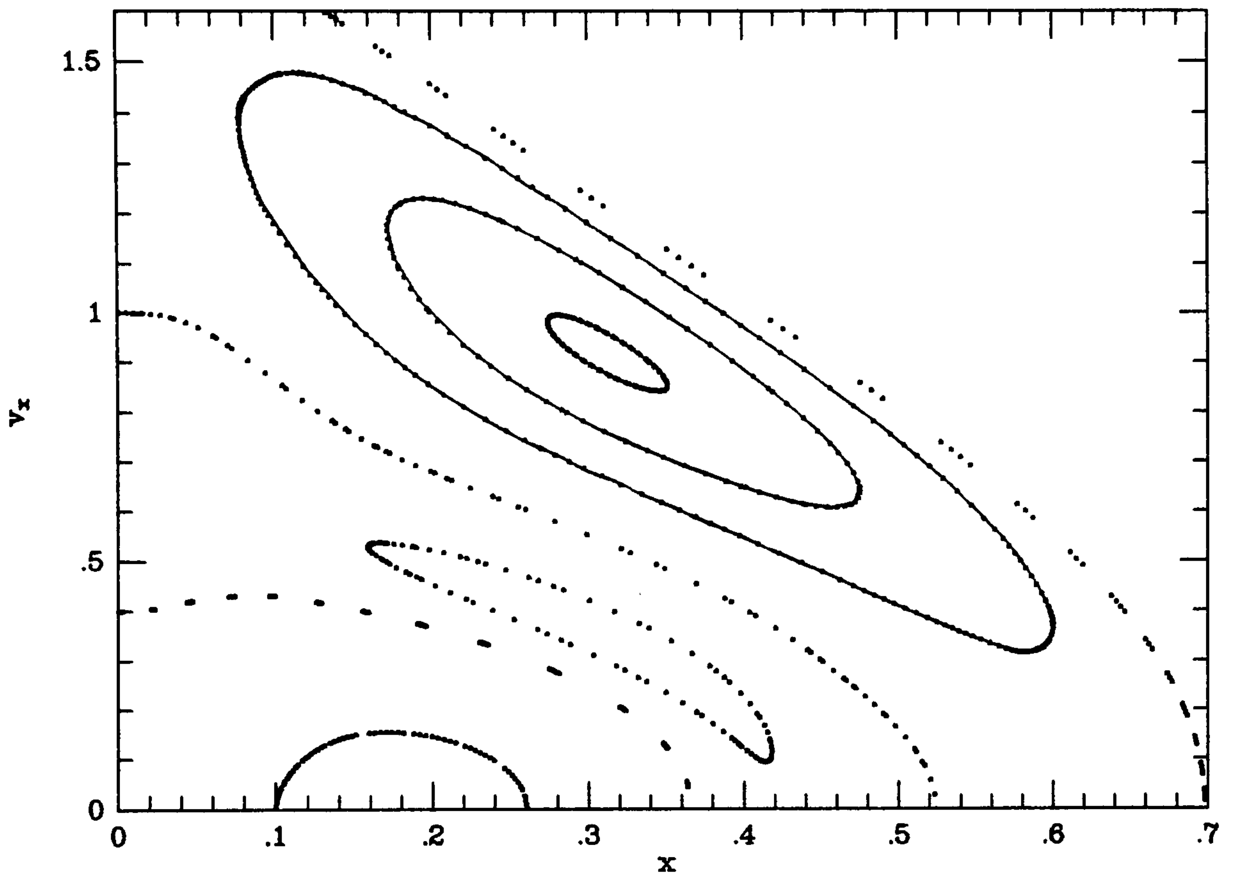

Figure 2.7: The $(x, \dot{x})$-surface of section corresponding to Fig. 2.6. Dots are integrated consequents; lines are the sections of the constructed tori.

In Fig. 2.7, the above approach has been applied to the banana orbits of Fig. 2.6, using the harmonic oscillator as $H_{\rm T}$. The actions of the tori are, starting from the outermost one, $\boldsymbol{J}' = (0.4303, 0.0344)$, $(0.4676, 0.0121)$, and $(0.4867, 8.03\,10^{-4})$, at $H = -0.3215$ and $q = 0.6$. The actions of the closed orbit are $\boldsymbol{J}' = (0.4880, 0)$. In this case, the symmetry of motion perpendicular to the closed banana orbit is lost, so the generating function is of the form

$$S(\boldsymbol{\theta}, \boldsymbol{J}') = \boldsymbol{\theta} \cdot \boldsymbol{J}' + 2 \sum_{\mathbf{n}>0} S_{\mathbf{n}}(\boldsymbol{J}') \sin[n_1\theta_1 + n_2(\theta_2 - \pi/2)], \tag{2.21}$$

where n_1 are even numbers, and n_2 both even and odd ones. The shift $\theta_2 - \pi/2$ occurs because the time-symmetric Fourier series (2.2) is evaluated in variables that are zero at the point

$\dot{\boldsymbol{J}} = 0$ (see McG&B); when the symmetry in the coordinate x_i is lost, $\dot{J}_i = 0$ at $\theta_i = \pm\pi/2$, as can be seen from (2.8). On the island contours of Fig. 2.7, θ_2 receives all the values in the range $0 \leq \theta_2 \leq 2\pi$, and θ_1 a range of values somewhere inside $0 < \theta_1 < \pi/2$. The number of coefficients $S_{\mathbf{n}}$ and the resolution in the $\boldsymbol{\theta}$-space were similar to those in the previous examples, and again the magnitudes of the coefficients decreased in a regular manner with increasing $|\boldsymbol{n}|$. More details of the point transformation used are given in appendix D. In chapter 4 we present another way of handling minor-orbit families by showing that one can derive them by perturbing an integrable Hamiltonian defined by the constructed major invariant tori.

Chapter 3

Hamiltonian dynamics and torus construction

3.1 Angle coordinates for a torus

In the previous chapter, torus construction was based on purely geometric principles when applying canonical transformations. The concept of actual Hamiltonian dynamics in the form of equations of motion was employed only when determining closed orbits in the target Hamiltonian. However, to complete the torus construction one needs consistent angle coordinates $\boldsymbol{\theta}'$ on a target torus, and this is where the equations of motion enter: the time development of an orbit on the torus labelled by $\boldsymbol{J}'$ is given by

$$\boldsymbol{\theta}' = \boldsymbol{\theta}'_0 + \boldsymbol{\omega}'(\boldsymbol{J}')t, \tag{3.1}$$

where $\boldsymbol{\omega}'$ is the orbital frequency vector, $\boldsymbol{\theta}'_0$ an initial phase and t the time.

On the other hand, from the definition (2.1) of the generating function $S(\boldsymbol{\theta}, \boldsymbol{J}')$ we obtain the relationship between the toy and target angles:

$$\boldsymbol{\theta}' = \frac{\partial S(\boldsymbol{\theta}, \boldsymbol{J}')}{\partial \boldsymbol{J}'} = \boldsymbol{\theta} + 2\sum_{\mathbf{n}>0} \frac{\partial S_{\mathbf{n}}(\boldsymbol{J}')}{\partial \boldsymbol{J}'} \sin(\boldsymbol{n} \cdot \boldsymbol{\theta}). \tag{3.2}$$

Thus we need to determine the derivatives $\partial S_{\mathbf{n}}/\partial \boldsymbol{J}'$ of the coefficients $S_{\mathbf{n}}$ to be able to use (3.2). Direct numerical differencing will be dominated by numerical noise because $S_{\mathbf{n}}$ are obtained independently for each $\boldsymbol{J}'$; therefore $\partial S_{\mathbf{n}}/\partial \boldsymbol{J}'$ should be obtained independently for each torus as well. From (3.2) and (3.1) we can obtain $\partial S_{\mathbf{n}}/\partial \boldsymbol{J}'$ and the frequencies $\boldsymbol{\omega}'$ as follows.

Once the generating function S is known, one has a mapping from any point described by the toy angles $\boldsymbol{\theta}$ on the torus of given $\boldsymbol{J}'$ to a phase-space point $\boldsymbol{w}$. We choose K initial points $\boldsymbol{\theta}^{(k0)} \equiv \boldsymbol{\theta}(t_{k0})$ and integrate the equations of motion to obtain, for each starting point, the particle's locations $\boldsymbol{\theta}^{(ki)} \equiv \boldsymbol{\theta}(t_{ki})$ at N subsequent times $t_{k1}, \ldots, t_{kN}$. Eliminating $\boldsymbol{\theta}'$ between (3.2) and (3.1), and enumerating a sufficiently large set of integer vectors $\boldsymbol{n}$ with α, it follows that

$$\boldsymbol{\theta}'^{(k0)} + \boldsymbol{\omega}' t_{ki} + \sum_{\alpha} s_{\alpha}^{ki} \frac{\partial S_{\mathbf{n}_\alpha}}{\partial \boldsymbol{J}'} = \boldsymbol{\theta}^{(ki)}, \tag{3.3}$$

where

$$s_{\alpha}^{ki} \equiv -2 \sin\left(\boldsymbol{n}_\alpha \cdot \boldsymbol{\theta}^{(ki)}\right). \tag{3.4}$$

Each component $j = r, \varphi$ (or x, y) of equation (3.3) may be regarded as one of a system of $K(N+1)$ linear equations for the components of a vector of unknowns

$$\boldsymbol{X}^{(j)} \equiv \left(\theta_j'^{(10)}, \ldots, \theta_j'^{(K0)}, \omega_j', \frac{\partial S_{\mathbf{n}_1}}{\partial J_j'}, \frac{\partial S_{\mathbf{n}_2}}{\partial J_j'}, \ldots\right)^T, \tag{3.5}$$

in terms of the known $K(N+1)$-dimensional vector

$$\boldsymbol{\Theta}^{(j)} \equiv (\theta_j^{(10)}, \theta_j^{(11)}, \ldots, \theta_j^{(1N)}, \theta_j^{(20)}, \ldots, \theta_j^{(KN)})^T. \tag{3.6}$$

In matrix form, equations (3.3) read

$$\mathsf{M}\, \boldsymbol{X}^{(j)} = \boldsymbol{\Theta}^{(j)}, \tag{3.7}$$

where M is defined by

$$\mathsf{M} \equiv \begin{pmatrix} 1 & 0 & \dots & t_{10} & s_1^{10} & s_2^{10} & \dots \\ 1 & 0 & \dots & t_{11} & s_1^{11} & s_2^{11} & \dots \\ \dots & & & \dots & & \dots & \dots \\ 0 & 1 & \dots & t_{20} & s_1^{20} & s_2^{20} & \dots \\ 0 & 1 & \dots & t_{21} & s_1^{21} & s_2^{21} & \dots \\ \dots & & & \dots & & \dots & \dots \\ 0 & \dots & 1 & t_{KN} & s_1^{KN} & s_2^{KN} & \dots \end{pmatrix} . \tag{3.8}$$

Note that M is independent of j.

The matrix M is well conditioned, and the solution of equations (3.7) is straightforward, provided the starting points $\boldsymbol{\theta}^{(k0)}$ as well as the subsequent integrated points $\boldsymbol{\theta}^{(ki)}$ are fairly uniformly distributed in the range $(0, 2\pi)$. Since the toy angles θ_j do not evolve uniformly in time, this goal will in general not be achieved by sampling at equal time intervals; it is better to aim for roughly equal sampling in, say, θ_r. Similarly, if all the $\boldsymbol{\theta}^{(ki)}$ form a single temporal sequence ($K = 1$), a resonant or near-resonant orbit will produce a distinctly non-uniform distribution of points $\boldsymbol{\theta}^{(1i)}$. This problem can be overcome by an intelligent choice of several different starting points $\boldsymbol{\theta}^{(k0)}$. One has only to bear in mind that each additional starting point introduces two new unknowns into the equations, so one must maintain a balance between having many starting angles, and gathering enough data points $\boldsymbol{\theta}^{(ki)}$ by following the orbit for long enough from each starting point.

If the created torus approximates an existing invariant torus of H, the motion (3.1) on the torus will now approximate the time development $\boldsymbol{w}(t)$ given by H. If the torus lies in a resonant or stochastic region of H (such tori will be constructed in the next chapters), one is left free to define how (3.1) projects into phase-space points $\boldsymbol{w}(t)$: obtaining $\partial S_{\mathbf{n}}/\partial \boldsymbol{J}'$ as the best fit of (3.3) to many short orbit integrations provides a controlled transition over such a region.

As an example, in Fig. 3.1(a) we show the trajectories in toy angle space of 36 orbit

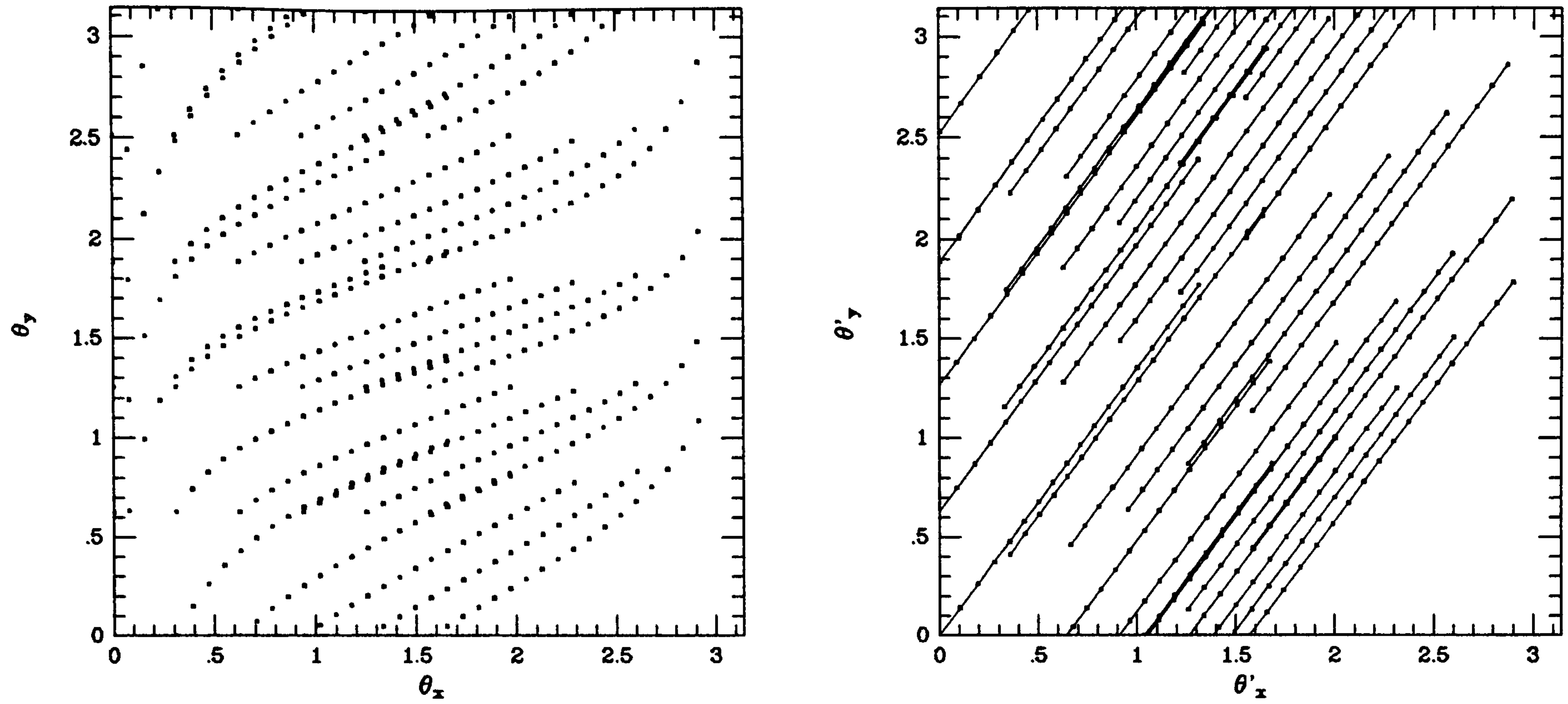

Figure 3.1: (a) Trajectories in toy angle space obtained by direct integration of the equations of motion in the potential of (2.5) with $q = 0.9$ from 36 starting points on the uppermost torus of Fig. 2.2. (b) The same trajectories in target angle space (dots), together with the semi-analytical linear trajectories (3.1) (straight lines).

integrations started from different points $\boldsymbol{\theta}^{(k0)}$ on the torus corresponding to the uppermost box orbit of Fig. 2.4 – this is an orbit in the logarithmic potential (2.5) with axial-ratio $q = 0.9$. The set of integer vectors $\mathbf{n}_\alpha$ used for $\partial S/\partial \boldsymbol{J}'$ was the same as that for S. Fig. 3.1(b) shows the same trajectories in target angle space as determined from equation (3.7). The dots in Fig. 3.1(a) and (b) show the input and output angles $\boldsymbol{\theta}^{(ki)}$ and $\boldsymbol{\theta}'^{(ki)}$, respectively, and the straight lines in Fig. 3.1(b) show the associated trajectories $\boldsymbol{\theta}'(t) = \boldsymbol{\theta}'^{(k0)} + \boldsymbol{\omega}'t$.

3.2 Generating function from orbit integration

In the same way as the angle equation (3.2), also the action equation (2.3) can be combined with the integration of the equations of motion. If the motion is restricted to a torus, substituting different values of $\boldsymbol{J}(t)$ and $\boldsymbol{\theta}(t)$, corresponding to integrated $\boldsymbol{w}(t)$, into (2.3) yields a set of linear equations for $\boldsymbol{J}'$ and $S_{\mathbf{n}}$. However, in contrast to the equations (3.7), such a set is usually hopelessly unstable to solve, even when the chosen angle values $\boldsymbol{\theta}(t_i)$ are evenly distributed in the $\boldsymbol{\theta}$-space.

To circumvent this difficulty, Warnock (1991) proposed a scheme where one replaces the equations for the coefficients $S_{\mathbf{n}}$ by equations for the values taken by $\boldsymbol{J}$ on a regular grid in $\boldsymbol{\theta}$. This amounts to obtaining the values of $\boldsymbol{J}$ needed for a discrete Fourier transformation (DFT), just like we did directly in the case of Stäckel potentials in the previous chapter, and on performing DFTs the values of $\boldsymbol{n}S_{\mathbf{n}}$ and $\boldsymbol{J}'$ are thus obtained from (2.3).

The set of linear equations is straightforward to obtain. First, recall that the two-dimensional DFT of a 2-D function $J(\theta_1, \theta_2)$ is

$$c_{mn} = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} J_{kl} \exp\left[-\mathrm{i}2\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)\right], \tag{3.9}$$

where M, N are the numbers of sampling points in each dimension, and

$$J_{kl} \equiv J\left(\frac{2\pi k}{M}, \frac{2\pi l}{N}\right). \tag{3.10}$$

Using c_{mn} as approximations for the coefficients of the 2-D Fourier series of J (and identifying the index $M - i$ with the Fourier coefficient index $-i$), we have

$$J(\boldsymbol{\theta}) = \frac{1}{MN} \sum_{k,l} J_{kl} \mathcal{S}_{kM}(\theta_1) \mathcal{S}_{lN}(\theta_2), \tag{3.11}$$

where

$$\mathcal{S}_{kM}(\theta) \equiv \sum_m \exp[\mathrm{i}m(\theta - 2\pi k/M)]. \tag{3.12}$$

We use a grid with an even number M of points in a dimension, to be able to employ the symmetries of the system (Warnock uses an odd number, resulting in a slightly different final expression for $\mathcal{S}$). The sum over m goes thus from $-(M/2-1)$ to $(M/2-1)$, which can be evaluated as a geometric series, and then the 'halfway' $M/2$-term is included. Denoting $\tilde{\theta}_{kM} \equiv \theta - 2\pi k/M$, the coefficients $\mathcal{S}_{kM}(\theta)$ are

$$\begin{aligned} \mathcal{S}_{kM}(\theta) &= \frac{1 - \exp[\mathrm{i}(M-1)\tilde{\theta}_{kM}]}{1 - \exp(\mathrm{i}\tilde{\theta}_{kM})} \exp[-\mathrm{i}(M/2-1)\tilde{\theta}_{kM}] + \cos(M\tilde{\theta}_{kM}/2) \\ &= \frac{\sin[(M-1)\tilde{\theta}_{kM}/2]}{\sin(\tilde{\theta}_{kM}/2)} + \cos(M\tilde{\theta}_{kM}/2) = \frac{\sin(M\tilde{\theta}_{kM}/2)\cos(\tilde{\theta}_{kM}/2)}{\sin(\tilde{\theta}_{kM}/2)}. \end{aligned} \tag{3.13}$$

(With an odd-valued M, the cosine-factor of the final expression of (3.13) is missing.)

From (3.11), the grid point actions J_{kl} can now be solved for by integrating the orbit long enough so that there is a point $\boldsymbol{\theta}^{(kl)}$ close to each grid point $\boldsymbol{\theta}_{kl}$. (In other words, we interpolate the actions J_{kl} from the integrated values $J(\boldsymbol{\theta}^{(ij)})$ in such a way that a 2-D Fourier series gives the observed $J(\boldsymbol{\theta}^{(ij)})$.) The equation matrix elements $\mathcal{S}_{kM}(\theta_1^{(ij)})\mathcal{S}_{lN}(\theta_2^{(ij)})$ of (3.11) are strongly peaked at kl (when $\boldsymbol{\theta} \rightarrow \boldsymbol{\theta}_{kl}$ for all grid points, the equation matrix becomes an identity matrix). The matrix is therefore sparse and the equations can quickly be solved iteratively (e.g. Gauss-Seidel method). The iteration usually converges robustly so long as there is a point in each grid cell. Also, the number of iteration steps needed for practical accuracy is usually only a few, often just two or three. Because a matrix element is separated in θ_1 and θ_2, it is sufficient to store the factors $\mathcal{S}_{kM}(\theta_1^{(ij)})$ and $\mathcal{S}_{lN}(\theta_2^{(ij)})$ separately instead of the huge matrix in its entirety.

This scheme is less general than ours in that it can only yield tori that are invariant tori of H, i.e., the types of possible target tori are limited. As discussed in chapter 2, the toy and target tori must be of the same type. Thus, for example, it would be impossible to create box-type major-orbit tori slicing through minor-orbit tori of the target Hamiltonian, as is done in the next chapter with the least-squares method, because these tori are not compatible in $\boldsymbol{\theta}$-space in the sense that an expression like (2.3) cannot be used.

When applicable, this scheme is more accurate, however, since with it there is no practical objection to employing every $S_{\mathbf{n}}$, whereas in our scheme computational tractability demands that one set to zero as many of the $S_{\mathbf{n}}$ as possible. In view of this superior accuracy, it is advantageous to combine our scheme with Warnock's: our scheme is first used to optimize a point transformation and the toy potential parameters and determine a crude generating function. Then Warnock's scheme is used to refine the generating function with the other parameters taken as optimized. As an example of this approach in action, in Fig. 3.2 we compare fits to the five tori of Fig. 2.4 in the potential (2.5). The dashed curves show sections of tori obtained by least-squares fitting, while the solid lines show tori constructed

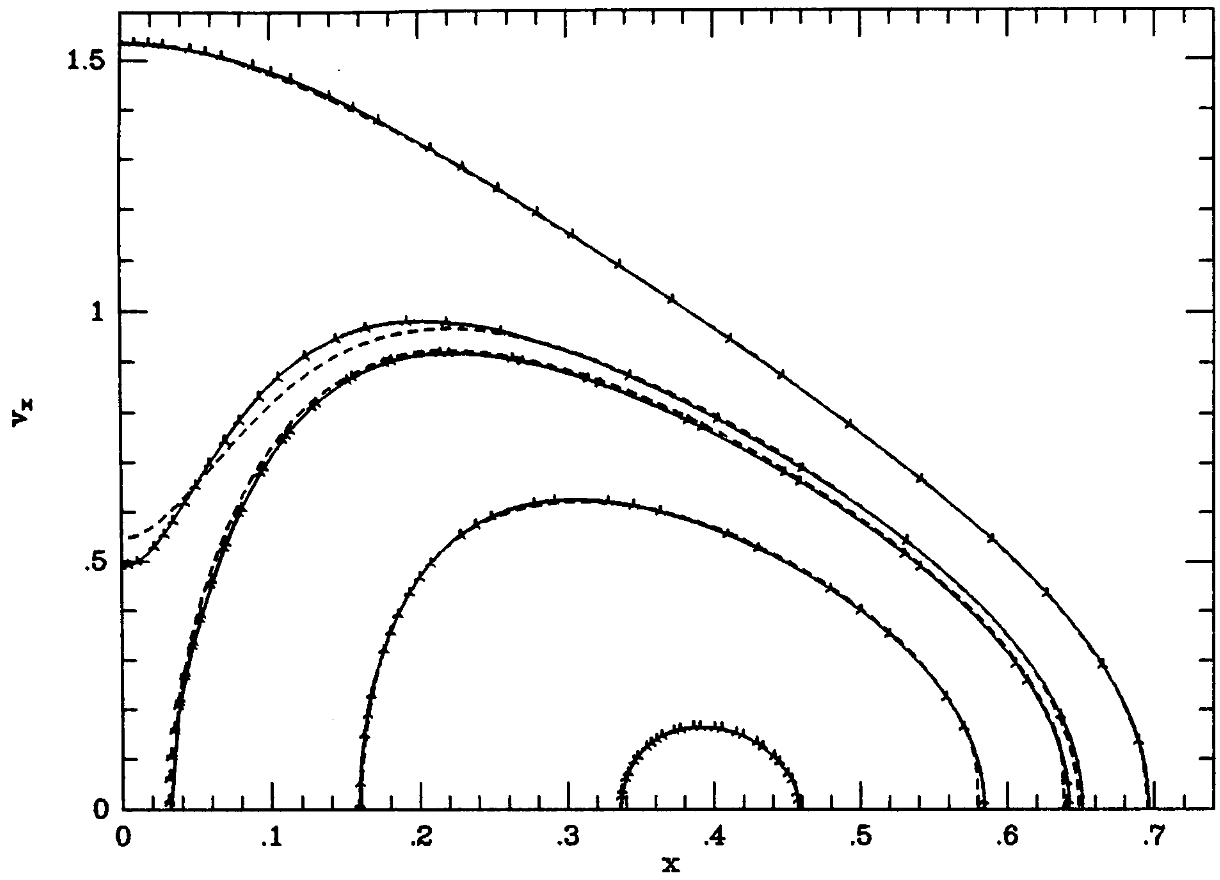

Figure 3.2: $(x, \dot{x})$ at $q = 0.9$ and $H = -0.315$. The points show consequents obtained by direct orbit integration, and the lines are the invariant curves obtained by two torus-construction algorithms.

by Warnock's method using the toy potential and point transformation parameters returned by the least-squares fits. The computational effort in systems of two degrees of freedom is approximately the same in both techniques for comparable resolution in $\boldsymbol{\theta}$-space.

If the generating function S has to be very long, there are practical difficulties just as with the least-squares method. For example, if the toy actions $\boldsymbol{J}(\boldsymbol{J}'; \boldsymbol{\theta})$ have somewhere large, abrupt 'jumps' (even if they are elsewhere smooth), a proper Fourier series representation can be hard to give. With a sparse grid, the obtained generating function produces too wiggly curves, and a dense grid requires long integration times. Also, the integrated trajectories can cluster very densely together in some parts of the toy angle $\boldsymbol{\theta}$-space, while leaving some parts practically empty, so the integration time does not necessarily grow just linearly with the grid density. Besides, when a very dense grid has to be used, the orbit must really fill a torus densely, i.e. the corresponding curve in a surface of section must really be invariant to a high accuracy.

Computational efficiency in more than two degrees of freedom

With two degrees of freedom, we have two quite efficiently working procedures: on an average-sized workstation, the accurate construction of a torus takes from, say, half a minute or a bit less to a few minutes or a bit more, depending on the difficulty of the situation. Ultimately we would like to do all this in three degrees of freedom, e.g. in the case of fully three-dimensional stellar systems, so first we have to check how long the machinery will have to grind to create a torus. Undoubtedly, at least a lunch-break is needed, but will we have to wait for a longer time?

As an example, let us consider a harmonic-oscillator mapping in a potential with the same symmetries as (2.5), and form crude estimates of computation times. Typically, coefficients $S_{\mathbf{n}}$ of at least the order 30 in each dimension are needed. For the orbit integration method, this means, say, 64 points for 2π in each dimension, so employing the symmetries, 32 points in the interval $[0,\pi]$ are sufficient. Thus the space of $\boldsymbol{\theta}$-cells is 32 times larger: this means 32 times longer integration times and 1024 times longer matrix iteration times than in two dimensions. For the least-squares method, the number of $S_{\mathbf{n}}$ needed will be, say, 10 times larger (assuming that most of the $S_{\mathbf{n}}$ can be set to zero as efficiently as in two dimensions). The matrix (Gauss-Jordan) solution part of the Levenberg-Marquardt algorithm is essentially a N^3-process; thus it will take about 1000 times longer. There are 32 times more $\boldsymbol{\theta}$-points; at each point the gradient $\partial H/\partial S_{\mathbf{n}}$ for the Fourier coefficients (and other parameters) will have to be computed, so that part will take 320 times longer.

In two degrees of freedom, the gradient computation takes up most of the time in the least-squares method; the matrix solution is very fast (but not negligible after a sizeable N^3-increase). In the other method the orbit integration is by far the most time-consuming part - each matrix iteration is usually a matter of seconds. Expanding the timescales to three degrees of freedom, we can form the following (very rough) estimates. The orbit integration method will take some $\frac{1}{2}$–3 hours to integrate the $\boldsymbol{\theta}$-points, and then each matrix iteration will take up to about one hour, resulting in an average timescale of a few hours. (Also, using the so-called successive overrelaxation methods (e.g. Press et al. 1986), the number of iterations needed is even smaller than for the Gauss-Seidel method.) The least-squares method will need roughly an hour or more for each iteration, and there are usually many more iterations needed than for the orbit integration method. Therefore, the least-squares method might take some 10 to 30 hours to construct a torus. In addition to this, there is the problem of obtaining a stable and accurate solution for such a large dense matrix (in orbit integration, this is no problem: even though the matrix is hideously large, it is sparse, and stable to iterative solution).

In view of this, it seems that, in three degrees of freedom, the orbit integration method could be more efficient when there are invariant (or 'essentially invariant') tori of H to approximate. What makes this method robust is that the only 'vulnerable' part of its algorithm is the orbit integration that must reach all $\boldsymbol{\theta}$-cells; once these are obtained, the iterative part is more or less foolproof. Of course, to use this method, one must first have the proper point transformation and toy Hamiltonian parameters; these could be obtained with the least-squares method using a greatly reduced $S_{\mathbf{n}}$-set. Even then, the integration part could take very much longer than estimated above: there can be some cells in the $\boldsymbol{\theta}$-space that will be hard to reach.

In any case, we can say we are lucky to live in a world where systems of a low number of degrees of freedom are of interest: in two dimensions, the machinery is efficient; in three dimensions, all right with a bit of patience and a powerful computer; in four or more, the project would be completely impossible in practice.

Chapter 4

Integrable Hamiltonians from invariant tori

4.1 Overview

In the previous chapters, we have shown how approximate invariant tori can be constructed for a general Hamiltonian H. Since invariant tori are the geometric manifestation of an integrable Hamiltonian, it is natural to use these tori to define such a Hamiltonian H_0 that closely approximates H (e.g., Warnock & Ruth 1991, 1992 and Binney & Kumar 1993). Here we give a detailed formulation of such a method. More specifically, the main points are: (i) to show how H_0 may be constructed anywhere in phase space, either by direct torus construction or by interpolation using previously constructed tori; (ii) to introduce a modified form of the pendulum analysis of standard secular perturbation theory that is necessitated by the unusually close fit of H_0 to H, and to show that minor-orbit families can be derived by perturbing constructed H_0s. Also, we show that the direct construction scheme of chapter 2 can often be succesfully used in regions of phase space where there are no existing invariant tori of H of a required type, e.g., major invariant ones. (In the next chapter, we study more closely the transition to global stochasticity, i.e., the final breakup of all 'KAM-tori' of H.) In the following discussion, tori in a toy potential are not needed, so we drop the earlier convention that primes are used for quantities associated with the target Hamiltonian.

The torus-construction scheme proceeds by numerically constructing a canonical transformation that maps the invariant tori of some 'toy' Hamiltonian into the phase space of the given 'target' Hamiltonian. Each image torus is null in the sense that the Poincaré invariant evaluated on any part of its surface vanishes, and it is usually an approximate orbital torus of H in the sense that some orbit of H keeps close to its surface. A set of these tori, together with their frequencies, essentially defines a new Hamiltonian H_0. The H_0 we construct behaves like a separable Hamiltonian in that all its orbits are quasi-periodic and the tori form a global sequence, smoothly parametrized by the actions $\boldsymbol{J}$ or the orbital frequencies $\boldsymbol{\omega} = \partial H_0/\partial \boldsymbol{J}$.

When H is near-integrable, a portion of the phase space is still occupied by invariant tori. Ideally, the torus-construction scheme 'finds' these target tori, and thus on them $H_0 = H$ in principle; the residual Hamiltonian

$$\delta H \equiv H - H_0 \tag{4.1}$$

is noise from the non-linear least-squares fit procedure. In the regions of phase space where there are no invariant tori of H of the desired (major or minor) type, H_0 and H must necessarily differ on fitted tori.

The difference δH is generally small, and dynamically insignificant away from the resonances of H_0. Near these resonances, minor-orbit families and/or chaos occur. It is natural to study these phenomena by treating $H = H_0 + \delta H$ as a perturbation on H_0.

This scheme raises two problems.

(i) The scheme employed above for filling phase space with null tori is liable to fail on encountering a sufficiently large minor-orbit family (when crossing over small minor families, the least-squares method can be expected to create consistent major-type tori that 'ignore' the minor families, as first described in Binney & Kumar 1993). Until this failure has been remedied, H_0 cannot be constructed, and the offending family cannot be treated as comprising resonantly trapped orbits.

(ii) δH often effectively vanishes through much of phase space, whereas the standard treatment of resonant trapping (e.g. Lichtenberg & Lieberman 1983) assumes that δH is significant far from the resonance. Also, constructing approximating tori is usually easier for some orbits than for others; this can break the symmetry (assumed in the standard theory) of the variation of δH over a resonance.

4.2 Driving tori through resonant islands

First, we address the question of how to obtain the tori which allow us to define the integrable Hamiltonian H_0. Generally, one needs to be able to obtain the phase-space coordinates $\boldsymbol{w}$ at an arbitrary value of $\boldsymbol{J}$. Unfortunately, the torus-fitting technique proceeds orbit by orbit, with the result that one initially knows $\boldsymbol{w}(\boldsymbol{J}, \boldsymbol{\theta})$ only on a grid of $\boldsymbol{J}$-values. It is natural to fill the gaps in our knowledge by interpolation between grid points. While one can interpolate directly between the values of $\boldsymbol{w}(\boldsymbol{J}, \boldsymbol{\theta})$, it is preferable to interpolate between the parameters that define the tori – only in this way can one create a general recipe for providing a torus for any $\boldsymbol{J}$. A generating function defined by interpolated parameters is of the same general form as one created by the torus-fitting procedure, and hence generates a valid canonical transformation. It is important to note that our integrable Hamiltonian H_0 is now essentially *defined* (not approximated) by whatever interpolation scheme we choose to employ. Therefore, the tori created by the torus-fitting machinery are used as guides; they alone do not define H_0.

Here we shall apply an interpolation scheme when using perturbation theory. Since it is usually employed in a regime where the frequencies $\boldsymbol{\omega}$ are close to a resonance of H_0 and the value of H_0 stays nearly constant, we introduce new action-angle coordinates that are more useful in the vicinity of a resonance than the old ones. The new and old coordinates are simple linear combinations of each other.

Let there be a resonance of H_0 at some $\boldsymbol{J}$ such that the ratio of the orbital frequencies $\boldsymbol{\omega} = \partial H_0/\partial \boldsymbol{J}$ is

$$\frac{\omega_2}{\omega_1} \equiv \gamma = \frac{r}{s}. \tag{4.2}$$

We transform to new coordinates $(J_\psi, \psi), (J_\xi, \xi)$ via the canonical transformation with generating function

$$F = (\boldsymbol{k} \cdot \boldsymbol{\theta}) J_\psi + (\boldsymbol{l} \cdot \boldsymbol{\theta}) J_\xi, \tag{4.3}$$

where

$$\boldsymbol{k} \equiv (r, -s), \tag{4.4}$$

and $\boldsymbol{l}$ is some other vector. Here, we shall use $\boldsymbol{l} \equiv (0, 1)$; this is useful in applying perturbation theory. From (4.3), we have

$$\begin{pmatrix} \psi \\ \xi \end{pmatrix} = \mathsf{M} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \quad \begin{pmatrix} J_\psi \\ J_\xi \end{pmatrix} = \left(\mathsf{M}^T\right)^{-1} \begin{pmatrix} J_1 \\ J_2 \end{pmatrix}, \quad \mathsf{M} = \begin{pmatrix} \boldsymbol{k} \\ \boldsymbol{l} \end{pmatrix}. \tag{4.5}$$

The above definition of $\boldsymbol{k}$ causes ψ to increase much more slowly than ξ near the resonance, since its frequency, $\omega_\psi = \boldsymbol{k} \cdot \boldsymbol{\omega}$, vanishes on resonance. Also, for infinitesimal deviations $\mathrm{d}H_0$ measured from the resonance,

$$\mathrm{d}H_0 = \boldsymbol{\omega} \cdot \mathrm{d}\boldsymbol{J} = (\boldsymbol{l} \cdot \boldsymbol{\omega})\, \mathrm{d}J_\xi. \tag{4.6}$$

In the new coordinates, moving along J_ψ while keeping J_ξ constant moves us along the tangent to the curve H_0=const. at the resonance point. With the tangents, we can approximate that curve piecewise by employing a set of different values of γ and the corresponding new coordinates. By using these local coordinates, we operate effectively in one dimension when keeping close to a constant value of H_0. When moving in the vicinity of a given resonance, we shall denote the value of J_ψ at which H_0 has the resonance by $J_{\psi 0}$, and employ the coordinate

$$\delta J_\psi \equiv J_\psi - J_{\psi 0}. \tag{4.7}$$

Weak resonances

'Fish' orbits in the potential (2.5), associated with a closed orbit in 3:2 resonance, provide an instructive example. In Fig. 4.1, we show the section $(x, \dot{x})$ at $q = 0.8$ and $H = -0.1992$, the dots marking numerically integrated consequents. The fish minor-orbit family is associated with the chain of islands; for this value of q, this is the largest of the minor-orbit families. Dashed curves show sections of some numerically constructed tori (otherwise this is the same as Fig. 2.1). Since the latter are the invariant tori of an integrable Hamiltonian that supports only major-orbit families, the dashed curves behave as though the island chain were not there. They define the unperturbed Hamiltonian H_0 by the rule that its value at any phase-space point $\boldsymbol{w}$ is simply the mean value of the full Hamiltonian H on the torus through $\boldsymbol{w}$:

$$H_0 = \overline{H} \equiv (2\pi)^{-2} \int \mathrm{d}^2\boldsymbol{\theta}\, H. \tag{4.8}$$

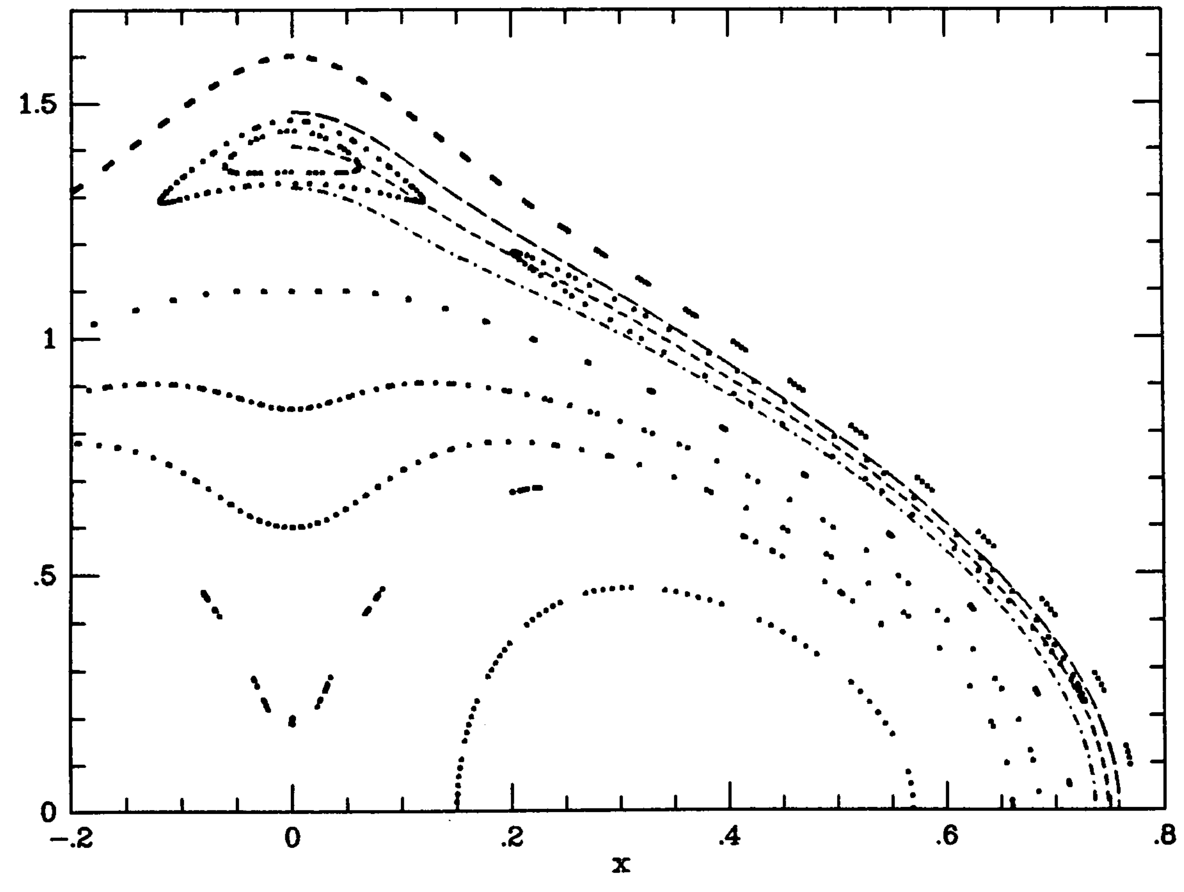

Figure 4.1: $(x, \dot{x})$ at $q = 0.8$ and $H = -0.1992$. Dots are integrated consequents, and the dashed lines are curves of constructed tori.

We show dashed curves that nearly touch the islands on each side, and one that slices right through the middle of all the islands in the chain. This last is the section of the torus $\boldsymbol{J} = (0.4324, 0.0934)$. The frequency-determining algorithm of chapter 3 assigns it frequencies $\boldsymbol{\omega} = (1.5477, 2.3215)$, giving $\omega_\psi = \boldsymbol{k} \cdot \boldsymbol{\omega} = 0.0001$ with $\boldsymbol{k} = (3, -2)$. Thus this torus is indeed a resonant torus of H_0. The two dashed curves that nearly touch the islands are the sections of tori with actions that differ from that of the resonant torus by $-0.008\,\boldsymbol{k}$ (the lower curve) and $0.007\,\boldsymbol{k}$, i.e., they are placed at $\delta J_\psi = -0.008$ and $\delta J_\psi = 0.007$.

Thus, when a minor-orbit family is sufficiently small as here, the torus-fitting machinery is able to drive perfectly reasonable tori through the portion of phase space occupied by the family, and the orbital frequencies and target angles for these tori can be computed as in chapter 3. Although there is no reason to give any special preference to these tori, it is natural to seek an interpolation scheme that produces similar tori when interpolating between those constructed outside the minor-orbit family's domain.

Consider now Fig. 4.2, which shows ω_ψ evaluated for a set of tori constructed at constant J_ξ and varying J_ψ at values that correspond to the fish family. The measured points clearly lie on a virtually straight line, the slope of which yields $\partial\omega_\psi/\partial J_\psi \equiv G = -10.4$. This implies that H_0 varies parabolically (between -0.1992 and -0.1997 for the points in the figure), reaching its maximum when $\delta J_\psi = 0$. The values corresponding to any given δJ_ψ, computed by adding $(G/2)(\delta J_\psi - \delta J_\psi^0)^2$ to the value of $\overline{H}$ on a torus constructed at any δJ_ψ^0 in the δJ_ψ-range of Fig. 4.2, do indeed agree with the independently computed values of $\overline{H}(\delta J_\psi)$.

Fig. 4.2 suggests that we could try interpolating linearly other quantities as well. Indeed, after having tried various interpolation schemes, we have found it practical to adopt the robust method of linearly interpolating the quantities S_{n}, $\partial S_{\mathrm{n}}/\partial \boldsymbol{J}$, ω_ψ, and all parameters $\boldsymbol{P}$ used in the torus-fitting technique, i.e. the parameters describing the toy potentials and possible additional point transformations. Employing this interpolation gives results that are in the closest agreement with those obtained by direct torus fitting.

Interpolating both S_{n} and $\partial S_{\mathrm{n}}/\partial \boldsymbol{J}$ linearly, even though these quantities are not in-

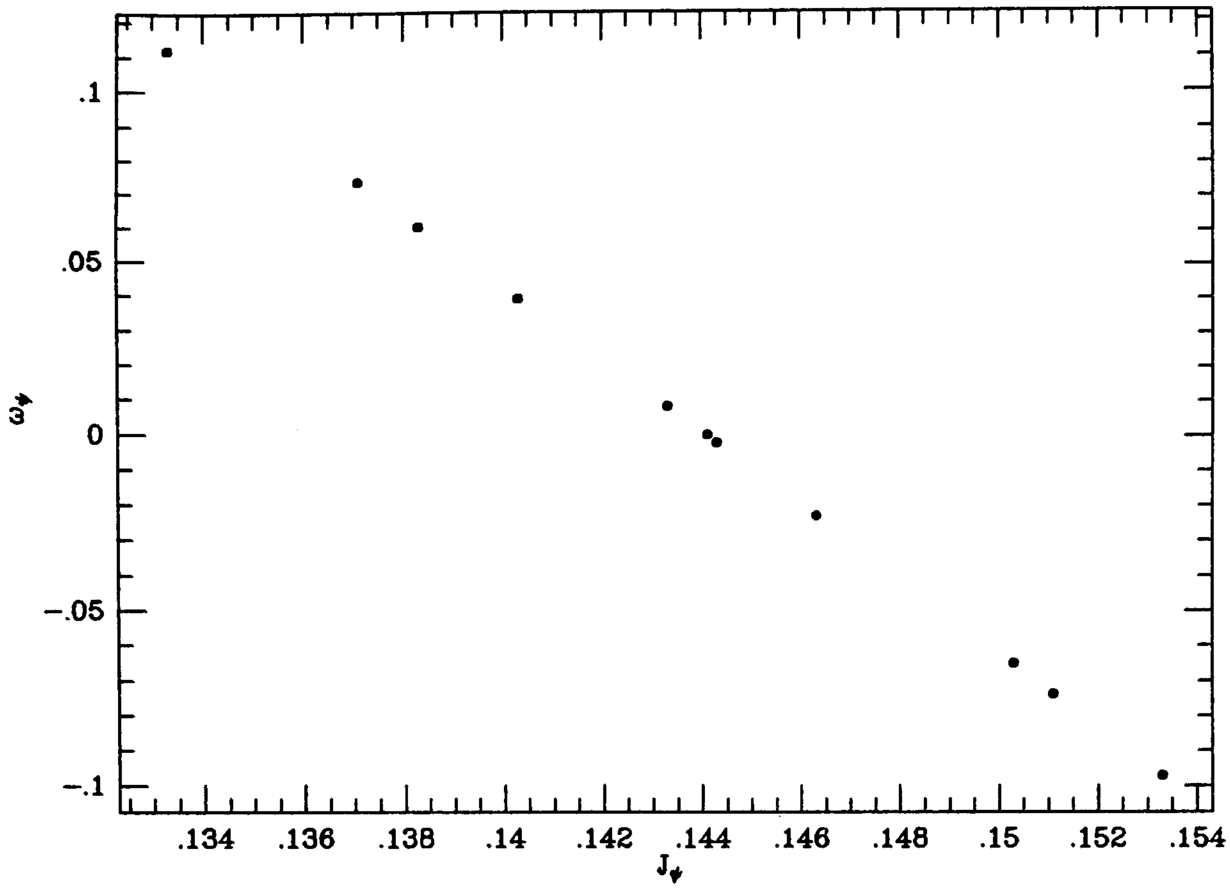

Figure 4.2: ω_ψ plotted against J_ψ for tori constructed at J_ξ = const. near the 3:2 resonance.

dependent, is quite justifiable. Taking into account the parameters $\boldsymbol{P}$, one can write $S_{\rm n} = S_{\rm n}(\boldsymbol{J};\boldsymbol{P}(J_\psi))$ along J_ψ. Thus we have

$$\frac{\partial S_{\rm n}}{\partial J_\psi} = \boldsymbol{k}\cdot\left(\frac{\partial S_{\rm n}}{\partial \boldsymbol{J}}\right)_{\boldsymbol{P}} + \left(\frac{\partial S_{\rm n}}{\partial \boldsymbol{P}}\right)_{\boldsymbol{J}}\cdot\left(\frac{\partial \boldsymbol{P}}{\partial J_\psi}\right), \tag{4.9}$$

where the subscript $\boldsymbol{P}$ has been attached to $\partial S_{\rm n}/\partial \boldsymbol{J}$ to emphasize that it is evaluated at constant $\boldsymbol{P}$. If we interpolate $S_{\rm n}$ and $\boldsymbol{P}$ linearly, it follows that

$$C = \boldsymbol{k}\cdot\left(\frac{\partial S_{\rm n}}{\partial \boldsymbol{J}}\right)_{\boldsymbol{P}} + \left(\frac{\partial S_{\rm n}}{\partial \boldsymbol{P}}\right)_{\boldsymbol{J}}\cdot\boldsymbol{c}, \tag{4.10}$$

where C and $\boldsymbol{c}$ are constants. $\partial S_{\rm n}/\partial \boldsymbol{P}$ is not explicitly determined for a torus constructed with the torus-fitting technique. In fact, we are at liberty to assign to its components widely differing values, since we move only along a line in $\boldsymbol{P}$-space (which is usually at least three-dimensional). On this line, which is parametrized by J_ψ, the condition (4.10) has to be fulfilled at each J_ψ for the linearly interpolated $S_{\rm n}$ and $\partial S_{\rm n}/\partial \boldsymbol{J}$ to be mutually consistent. Outside this line, there are no explicit constraints for $\partial S_{\rm n}/\partial \boldsymbol{P}$; the only condition is that

$S_{\mathrm{n}}(\boldsymbol{J};\boldsymbol{P}+\delta\boldsymbol{P})$ should be 'well behaved' and form a generating function S that gives sensible results. For each J_ψ, we have thus one constraint equation and as many unknowns as there are parameters $\boldsymbol{P}$, if $\boldsymbol{c}$ does not vanish. Therefore there should usually be various acceptable solutions for $\partial S_{\mathrm{n}}(J_\psi)/\partial\boldsymbol{P}$. (We are not interested in their actual values because we do not employ points off the given line in $\boldsymbol{P}$-space.) If the parameters $\boldsymbol{P}$ were the same (or very close to the same) for the interpolating tori, so that $\boldsymbol{c} = 0$, the resulting condition $\boldsymbol{k}\cdot\partial S_{\mathrm{n}}/\partial\boldsymbol{J} \doteq C$ would generally not be satisfied. Hence, the fact that the $\boldsymbol{P}$ are allowed to vary in the torus-fitting technique enables us to use linear interpolation for both S_{n} and $\partial S_{\mathrm{n}}/\partial\boldsymbol{J}$.

Strong resonances

What happens when a minor-orbit family is very large, and the corresponding island occupies a large portion of a surface of section? Our experiments have shown that, inside the island, least-squares minimization no longer works even qualitatively: the invariant curves it predicts are liable to cross each other and generally fail to follow the outline of the island. The reason for this is that the function minimized, i.e. the magnitude of the residual Hamiltonian, no longer represents what we want: the family is large precisely because the residual Hamiltonian is large. An interpolation scheme is now absolutely necessary, because there is no other way of constructing tori inside the island.

Fortunately, linear interpolation along the lines described above provides an entirely satisfactory route to the requisite tori. Fig. 4.3 is a plot of $\omega_\psi(J_\psi)$ for orbits in the potential (2.5) with $q = 0.6$ that lie near the banana minor-orbit family associated with a 2:1 closed orbit. Each point not between those marked by crosses connected with a line is obtained from a torus whose actions correspond to an actual non-resonant box orbit. The 'bad' points between the crosses are obtained from fitted tori that lie inside the island of the minor-orbit family. By contrast with the analogous frequency plot for the fish orbits, Fig. 4.2, the points

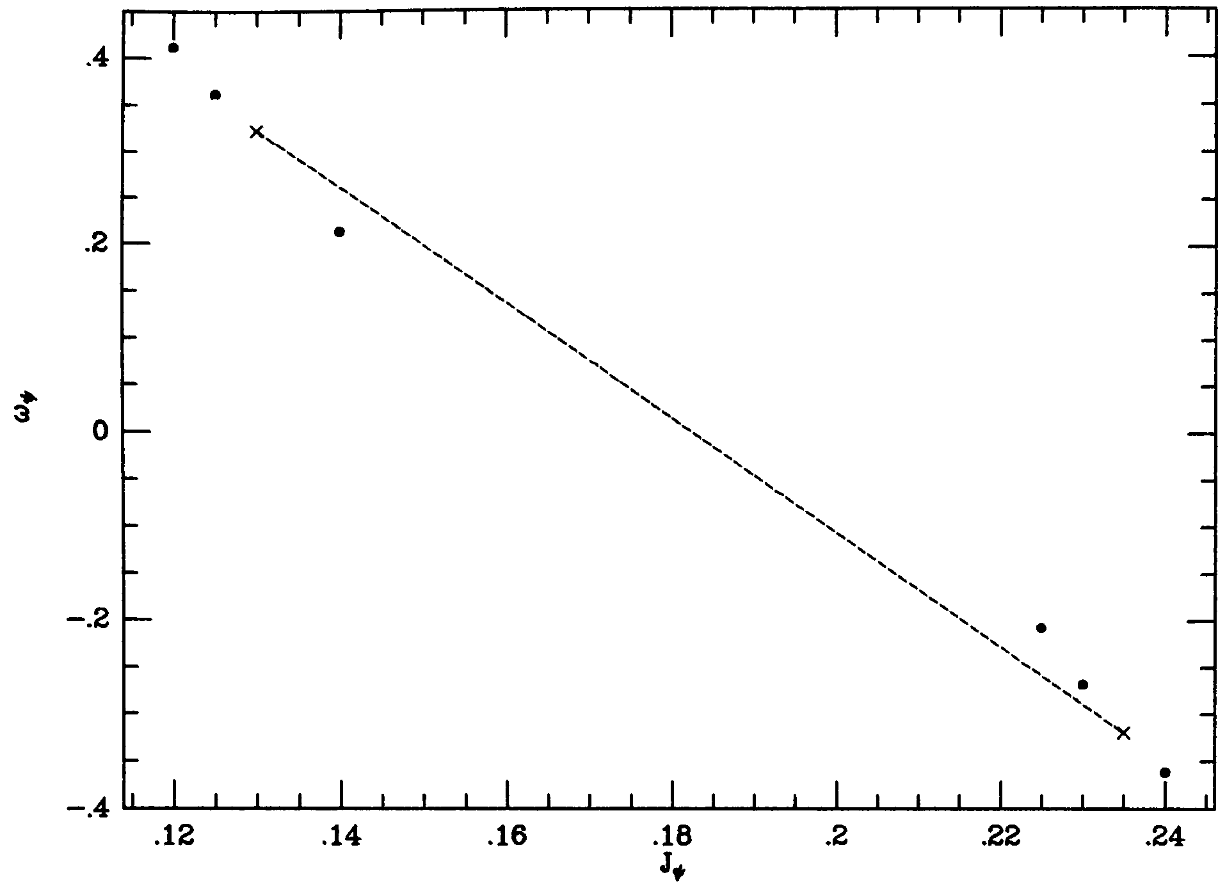

Figure 4.3: A $(\delta J_\psi, \omega_\psi)$-plot analogous to Fig. 4.2 near the 2:1 resonance at $q = 0.6$. The crosses mark the tori used for interpolation inside the resonant island.

in Fig. 4.3 do not lie on a straight line, but form two straight lines outside the crosses, one at each side of the plot. Between these two lines lies the range of actions associated with the island.

The linearly interpolated values of $\boldsymbol{\omega}(J_\psi)$, together with the value of H_0 at one end of the range of interpolation, *determine* $H_0(J_\psi)$. Continuity of H_0 at the other end of the range of interpolation can be ensured by choosing as endpoints tori on which H_0 (there defined to be $\overline{H}$) takes the same value at the two ends, and ω_ψ is equal in magnitude and opposite in sign. (Such pairs can always be found, because we are moving in two degrees of freedom and have two conditions.) The crosses in Fig. 4.3 mark two such points; the corresponding sections and that of the resulting resonant torus are shown in Fig. 4.4, which is the $(x, \dot{x})$ surface of section of (2.5) with $q = 0.6$ at $H = -0.3215$. Interpolation places the resonant torus at $\boldsymbol{J} = (0.3654, 0.0673)$. (A plot of three banana orbits corresponding to Fig. 4.4 is shown in Fig. 2.6.)

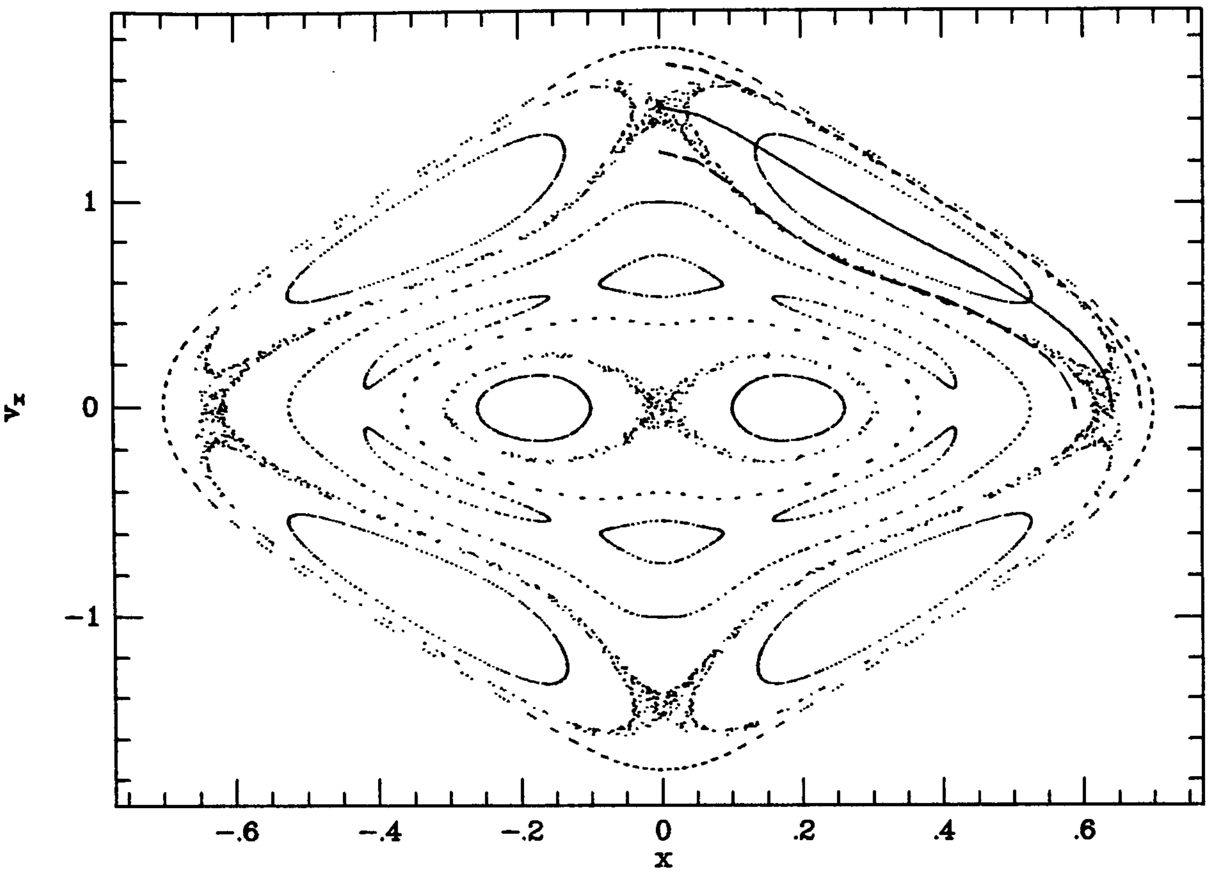

Figure 4.4: $(x, \dot{x})$ section at $H = -0.3215$ and $q = 0.6$ analogous to Fig. 4.1; the dashed lines represent the resonant torus and the two tori used in constructing it.

4.3 Defining H_0

Let us sum up here the procedure for defining an integrable Hamiltonian H_0. Once tori have been constructed and furnished with action variables for a set of $\boldsymbol{J}$-values, it is a simple matter to define an H_0 for which these are invariant tori. Globally, the orbits of most non-trivial Hamiltonians fall into several distinct families. For each significant family, one seeks a combination of a point transformation and a generating function (2.1) that maps the invariant tori of a toy Hamiltonian into approximately invariant tori of H. After fitting tori to every family one wishes H_0 to possess, we obtain all the quantities needed to define a torus for any given $\boldsymbol{J}$ by interpolating smoothly in $\boldsymbol{J}$ between the fitted tori. In particular the phase-space point $(\boldsymbol{x}, \boldsymbol{p})$ corresponding to given $(\boldsymbol{J}, \boldsymbol{\theta})$ is immediately computable, as is $H_0(\boldsymbol{J})$. For example, the values of $\boldsymbol{J}$ of the box- and loop-type invariant (or effectively invariant) tori of (2.5) fill much of the $\boldsymbol{J}$-space in an orderly manner and without overlapping, and a global H_0 can be constructed from these. The gaps between such $\boldsymbol{J}$-values correspond to minor-family or stochastic orbits (B&S). The former can be treated as resonantly trapped by a perturbation to H_0.

It is of interest to note that, if a Hamiltonian is near-integrable in the sense of the KAM theorem, it is integrable on a Cantor set consisting of the 'surviving' invariant (or KAM) tori (Pöschel 1982). In other words, the actions (or frequencies) assigned to the invariant tori indeed parametrize a global, ordered set, as we implicitly assume when constructing the global $H_0(\boldsymbol{J})$.

The value of H_0 on the created tori that approximate existing invariant tori of H is defined to be the average value of H (i.e. $H_0 = H$ in principle). When interpolating for H_0 between these tori, their frequencies $\boldsymbol{\omega} = \partial H_0/\partial \boldsymbol{J}$, independently determined by the method of chapter 3, must be taken into account; generally, H_0 is not necessarily the average value of H on a torus between the interpolating ones.

In the examples of this chapter, one-dimensional interpolation in J_ψ will be sufficient for applying perturbation theory. To obtain a completely general Hamiltonian $H_0(\boldsymbol{J})$, one must interpolate in two dimensions. This can be done by using spline functions, for example (Warnock & Ruth 1991,1992). If the $\boldsymbol{J}$-space contains merely weak resonances, one can use any action points $\boldsymbol{J}$ to create tori for an equally spaced interpolation grid. If there are strong resonances, one should adjust the grid so as to avoid regions where tori cannot be constructed directly. Also, one should note that if the parameters defining the toy system are the same for all $\boldsymbol{J}$, one necessarily needs the gradients $\partial S_{\mathbf{n}}/\partial \boldsymbol{J}$ for each grid point if $H_0(\boldsymbol{J})$ is to be consistent. In general, interpolation using only the values of $S_{\mathbf{n}}(\boldsymbol{J})$ does not produce the correct gradients, with the result that the angle coordinates $\boldsymbol{\theta}$ are not consistent with (3.1). Often it may be more convenient to interpolate using e.g. the values of H_0 and the rotation number $\gamma = \omega_2/\omega_1$ (parametrizing a curve of constant H_0) as the two interpolation variables, rather than $\boldsymbol{J}$. A grid in (H_0, γ) is natural in the sense that one can define tori at many fixed values of H_0 as above, and then fill the action space with the isoenergetic curves.

Also, it should be noted that H_0 is usually not derivable from a potential: on performing the test described in Binney & Kumar (1993), we did not find our constructed H_0 to be derivable from any potential.

4.4 Secular perturbation theory

Now we turn our attention to the problem of treating $H = H_0 + \delta H$ as a perturbation on H_0. (A detailed description of Hamiltonian perturbation theory and the removal of resonances can be found, for example, in chapter 2 of Lichtenberg & Lieberman 1983.)

Consider a Hamiltonian of the form

$$H(\boldsymbol{J},\boldsymbol{\theta}) = H_0(\boldsymbol{J}) + \delta H(\boldsymbol{J},\boldsymbol{\theta}), \tag{4.11}$$

where H_0 is a function of the actions $\boldsymbol{J}$ only, and δH is periodic in the angles $\boldsymbol{\theta}$:

$$\delta H(\boldsymbol{J},\boldsymbol{\theta}) = \sum_{\mathbf{n}} H_{\mathbf{n}}(\boldsymbol{J}) \exp(\mathrm{i}\boldsymbol{n}\cdot\boldsymbol{\theta}). \tag{4.12}$$

In our case, the coefficients $H_{\mathbf{n}}$, $\boldsymbol{n} \neq 0$, are the Fourier coefficients describing the fluctuation of the target Hamiltonian H on the constructed torus.

Let there be a resonance of H_0 at some $\boldsymbol{J}$ such that the ratio of the orbital frequencies is $\gamma = r/s$. Using the generating function (4.3) with $\boldsymbol{k} = (r, -s)$ and $\boldsymbol{l} = (0, 1)$, we have

$$\begin{aligned} J_\psi &= \frac{J_1}{r}, \qquad J_\xi = J_2 + \frac{s}{r} J_1, \\ \psi &= \boldsymbol{k}\cdot\boldsymbol{\theta}, \qquad \xi = \theta_2. \end{aligned} \tag{4.13}$$

Near the resonance, ψ increases much more slowly than $\xi = \theta_2$. The rapid circulation of ξ causes its conjugate action J_ξ to be effectively constant during the slow evolution of ψ. (From (4.12) one can see that $\dot{J}_\xi = -\partial\delta H/\partial\xi$ is small and rapidly alternating since the derivative $\partial/\partial\xi$ eliminates all terms not periodic in ξ.) Consequently, over the long timescale associated with the evolution of ψ we may treat J_ξ as a constant (adiabatic invariant) rather than a dynamical variable.

Similarly, J_ψ varies little in the characteristic time $2\pi/\omega_\xi$ associated with changes in ξ, so we may treat it as a constant when we substitute (4.13) into (4.11) and average H over ξ to obtain an effective Hamiltonian, at a constant value of which the motion in the phase plane (J_ψ, ψ) occurs:

$$\widehat{H} = H_0(J_\psi) + 2\sum_{m=1}^{\infty} H_{m\mathbf{k}}(J_\psi)\cos m\psi, \tag{4.14}$$

where the exponential form of (4.12) is given in the same cosine-form for time-reversal-symmetric Hamiltonians as in the previous chapters. If the target Hamiltonian is quadrant-symmetric, as with the logarithmic potential (2.5), $m = 2$ gives the first non-vanishing Fourier coefficient. The next ones are given by $m = 4, 6, \ldots$, and the corresponding coefficients in (4.14) decrease fast.

We have now transformed the problem into one with a single degree of freedom and thus into an integrable form, valid when $\omega_\psi \ll \omega_2$. Note that it does not matter which one of the original angles assumes the role of the fast angle ξ; here we choose the convention $\xi = \theta_2$. (The surfaces of section look the same with either θ_1 or θ_2 as ξ; the only difference is in the time development of the orbit, because the chosen angle is taken to increase linearly. This difference is very small so long as we are in the valid region $\omega_\psi \ll \omega_\xi$.)

The stationary points $(J_\psi^\pm, \psi^\pm)$ in the (J_ψ, ψ)-plane are found from $\partial\widehat{H}/\partial J_\psi = 0$ and $\partial\widehat{H}/\partial\psi = 0$. Evidently $(\partial\widehat{H}/\partial\psi) = 0$ for $m\psi = n\pi$, where $n = 0, 1, \ldots$. Also, we know that $(\partial H_0/\partial J_\psi) = \omega_\psi = 0$ on resonance. $J_{\psi 0}$, the value of J_ψ at which H_0 has a resonance, may not be identical with $J_\psi^\pm$, because the derivatives $(\partial H_{m\mathbf{k}}/\partial J_\psi)$ do not necessarily vanish at $J_{\psi 0}$. With (4.14), the nature of a fixed point depends on the sign of $(\partial^2\widehat{H}/\partial J_\psi^2)(\partial^2\widehat{H}/\partial\psi^2)$; for stable (elliptic) fixed points (J_ψ^+, ψ^+) it is positive, and for the unstable (hyperbolic) ones (J_ψ^-, ψ^-) it is negative.

The standard pendulum equation

Equation (4.14) is the master equation of secular perturbation theory. Since the underlying system has one degree of freedom, it is integrable regardless of the dependence of H_0 and the $H_{\mathbf{n}}$ on J_ψ. In practice, one works with a simplified form of (4.14) (above all to avoid endless computing of the Fourier coefficients at different J_ψ). Conventionally, (4.14) is limited to second order in J_ψ and to the leading value of $m\boldsymbol{k}$, because the excursions in J_ψ are small and the Fourier coefficients drop off fast. The standard form is obtained by ignoring the constant terms, expanding H_0 to second order in $\delta J_\psi \equiv J_\psi - J_{\psi 0}$, and dropping the first derivatives because the resonance condition $\partial/\partial J_{\psi 0} = 0$ is assumed to apply to all terms at $J_{\psi 0}$. This gives the pendulum equation characterizing motion near a resonance (Chirikov 1979):

$$\Delta\widehat{H} \equiv \frac{1}{2}G(\delta J_\psi)^2 + F\cos m\psi, \tag{4.15}$$

where the non-linearity parameter G is

$$G \equiv \left.\frac{\partial^2 H_0}{\partial J_\psi^2}\right|_{J_{\psi 0}} = \left.\frac{\partial \omega_\psi}{\partial J_\psi}\right|_{J_{\psi 0}} \tag{4.16}$$

and

$$F \equiv 2H_{m\mathbf{k}}(J_{\psi 0}). \tag{4.17}$$

(Note that we do not use a minus sign in our definition of F; some authors prefer to use it to make (4.15) look similar to the 'real' pendulum equation.) The nature of the fixed point $(J_{\psi 0}, \psi^{\pm})$ is given by GF. For $GF < 0$, $m\psi^- = \pm\pi$, and $m\psi^+ = 0$; for $GF > 0$, vice versa. In the following discussion, $GF < 0$ and $\psi^+ = 0$. (It is trivial to transform a system with $GF > 0$ to one in which $GF < 0$.)

Equation (4.15) predicts libration inside a resonant island bounded by the two branches of the separatrix, and circulation outside the island (see the phase plane in Fig. 4.5a). Qualitatively it produces islands resembling those in a surface of section such as Fig. 4.1, but quantitatively it fails because of special characteristics of our H_0 that we now discuss.

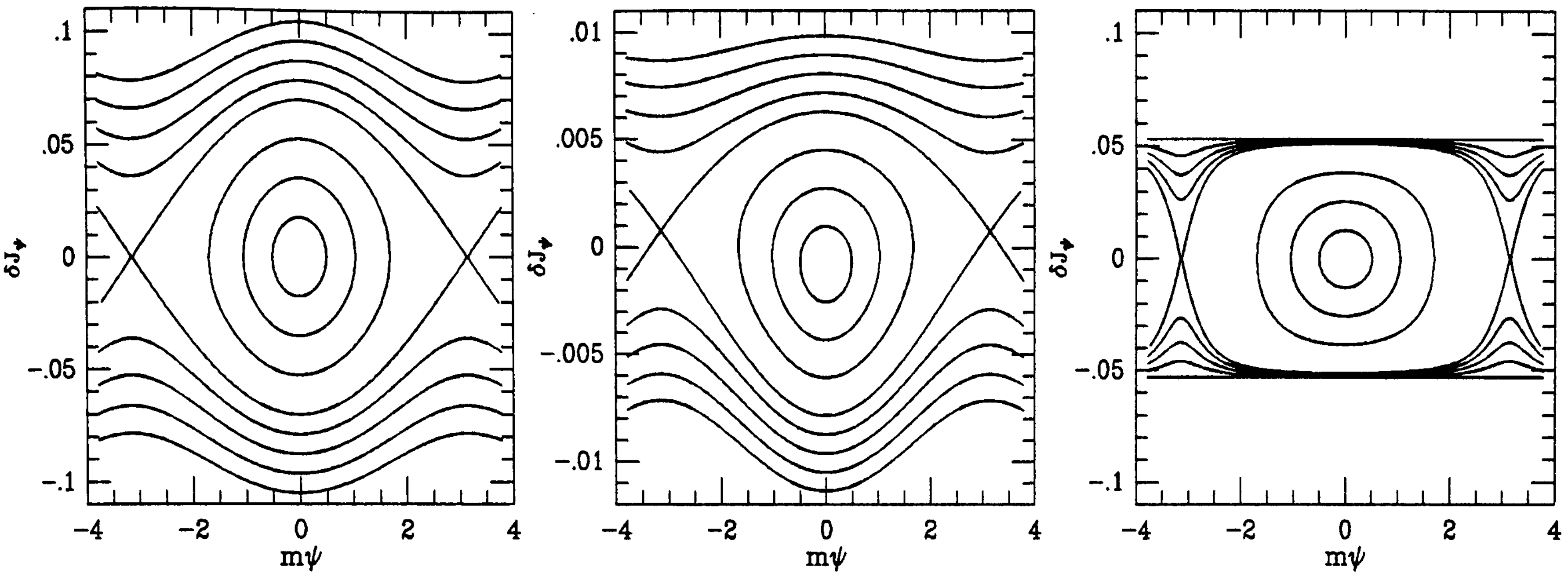

Figure 4.5: The (J_ψ, ψ) phase planes of different pendulums. (a) The standard pendulum. (b) The modified pendulum with $\alpha \neq 0, \beta = 0$. (c) The modified pendulum with $\alpha = 0, \beta \neq 0$. In (a) and (c) all parameters other than β are the same.

Outside the islands in Fig. 4.1, the surface of section looks very much like that of an integrable potential. In fact, perfect torus-fitting machinery would yield $H_0 = H$ soon outside the island, since there the tori describe orbits whose motion can (most probably) be described by invariant tori: from the theory of near-integrable potentials, we know that such tori exist outside the islands. It is known that, associated with the island boundary, there is usually a stochastic layer (see e.g. Lichtenberg & Lieberman 1983). In this example the layer is extremely small; thus soon outside the islands we would expect it to be possible to achieve $H_0 = H$.

Consider now the neighbourhood of an island in Fig. 4.1 as the counterpart of the (J_ψ, ψ) phase planes of Fig. 4.5. If one drew a horizontal line $J_\psi =$ const. in Fig. 4.5(a), it would be represented in Fig. 4.1 by a section of torus at constant $\boldsymbol{J}$, such as the dashed curves. On the other hand, the trajectories $\Delta\widehat{H} =$ const. in Fig. 4.5(a) correspond to the dotted curves formed by the consequents in Fig. 4.1. Outside the island there is a significant difference between the two figures. In Fig. 4.1 we know that, soon outside the island, perfect torus machinery would give curves that align accurately with the consequents because it is possible to have $H_0 = H$. But in Fig. 4.5(a) the pendulum equation (4.15) never produces compeletely horizontal lines – they align with the straight lines J_ψ = const. only asymptotically. The required phase plane should look like that in Fig. 4.5(c). Thus (4.15) cannot be the correct equation of motion for our perturbed system.

In practice, the torus-fitting machinery cannot perfectly achieve $H_0 = H$. In the case of Fig. 4.1, δH on tori constructed near the fish islands is larger on the side closer to the origin than on the side further away. This asymmetry is understandable, because it is naturally easier to fit tori to narrow box orbits than to wider ones. However, it gives rise to a need for an asymmetric term in the analysis as well. The asymmetric effect can be seen in the phase plane of Fig. 4.5(b).

There are thus two reasons why the standard perturbation theory has to be modified: one follows from a matter of principle, and the other because the torus-fitting machinery is imperfect. The formal reason for the failure of the standard equation (4.15) is that in it F is treated as a constant. One arrives at this constancy by arguing that the first-order (in δJ_ψ) term vanishes because of the assumed symmetry, and that the term quadratic in δJ_ψ from the expansion of $H_{m\mathbf{k}}(J_\psi)$ in (4.14) is smaller than the corresponding term in the expansion of $H_0(\psi)$ by a factor of order $\delta H/H_0$, and should therefore be negligible for sufficiently small perturbations. However, what matters is not so much its size relative to terms in the expansion of H_0, but its size relative to other coefficients of $\cos m\psi$. Using constant values of $H_\mathbf{n}$ is not consistent with the requirement that somewhere outside the resonant island the $H_\mathbf{n}$ should (ideally) vanish because $H = H_0$, and that in practice the $H_\mathbf{n}$ often exhibit asymmetric behaviour. Thus, the variation of the Fourier coefficients in δJ_ψ should be included in our formalism.

Modified pendulum

The complete second-order form for our analysis is given by approximating the variation of $H_{m\mathrm{k}}$ (or any other Fourier coefficient) in the neighbourhood of $J_{\psi 0}$ as

$$H_{m\mathrm{k}}(J_{\psi 0} + \delta J_\psi) = H_{m\mathrm{k}}(J_{\psi 0}) + \frac{\alpha}{2}\delta J_\psi + \frac{1}{2}\frac{\beta}{2}(\delta J_\psi)^2, \tag{4.18}$$

where α and β are constant parameters to be determined: α is associated with the asymmetry described above, while β determines how quickly $H_{m\mathrm{k}}$ decreases away from the resonance of H_0.

When the derivatives of $H_{m\mathrm{k}}$ are taken from (4.18) rather than set to zero, (4.15) is replaced by

$$\Delta\widehat{H} \equiv \frac{1}{2}(G + \beta\cos m\psi)(\delta J_\psi)^2 + (\alpha\delta J_\psi + F)\cos m\psi. \tag{4.19}$$

The fixed points lie at

$$\delta J_\psi^\pm = \mp\frac{\alpha}{G \pm \beta}. \tag{4.20}$$

Thus the α-term shifts them to the opposite sides of the line $\delta J_\psi = 0$. The nature of a given fixed point is determined by the sign of $(G \pm \beta)H_{m\mathrm{k}}(J_{\psi 0} + \delta J_\psi)$ which is essentially $\mathrm{sign}(GF)$, as for the normal pendulum.

Fig. 4.5(b) shows the phase plane of a modified pendulum for which $\alpha \neq 0$ and $\beta = 0$, while in Fig. 4.5(c) $\beta \neq 0$ and $\alpha = 0$. The asymmetric nature of Fig. 4.5(b) is evident. In Fig. 4.5(c), the key feature as opposed to Fig. 4.5(a) is striking: at the value $|\delta J_\psi| = \Delta J_\psi = \sqrt{2|F|/|\beta|}$, $H_{m\mathrm{k}}$ vanishes and further outside the rotation is uniform with the curves of constant $\Delta\widehat{H}$ aligning with lines of constant δJ_ψ. When $\alpha = 0$ and the resonant island is very large, so that $|H_{m\mathrm{k}}|$ outside the island is much smaller than in its centre, one can thus interpret β as a parameter giving roughly the first torus outside the island (at $\pm\Delta J_\psi$), for which $H_0 = H$ can be achieved.

The size and shape of the island in Fig. 4.5(a) are quite different from those of the islands in Fig. 4.5(b) and (c). The two separatrix branches comprise the contour of $\Delta\widehat{H}$ that passes through the unstable fixed points $(J_\psi^-, m\psi = \pm\pi)$. On this contour $\Delta\widehat{H} = -F - (\alpha/2)\delta J_\psi^-$. The values of δJ_ψ at $\psi = 0$ on this contour are the maximum deviations of δJ_ψ in libration. They are given by $\delta J_\psi^+ \pm \delta J^{\mathrm{sx}}$; the island half-width δJ^{sx} can be solved for from (4.19):

$$\delta J^{\mathrm{sx}} = \frac{2}{G+\beta}\sqrt{-F(G+\beta) - \frac{\alpha^2\beta}{2(G-\beta)}}. \tag{4.21}$$

A full account of the dynamics of the modified pendulum will be found in appendix C.

Generally, the choice of the values of α and β has to be made more or less by hand, evaluating $H_{m\mathrm{k}}$ for some constructed tori to judge its variation and comparing the sizes and shapes of the islands predicted by (4.19). Naturally, an extensive fit to $H_{m\mathrm{k}}$ on many tori is both impractical (and somewhat ambiguous because of numerical uncertainties) and against the basic philosophy of our approach. We have found that many cases are well described by using only one of α and β, and setting the other to zero. An important factor is the level of the 'background noise' in the torus fitting. When the island is very small, the fluctuation of H is of the same magnitude both inside and outside the island; only with very large islands can it be considered negligible close to the island. Thus α and β are typically functions of the size of the island relative to the total area of the surface of section, having separate regions of greatest influence: $|\alpha|$ decreases as the island grows, while $|\beta|$ grows with the island's size.

We now apply the above approach to the fish family of Fig. 4.1. The non-linearity parameter G is obtained from Fig. 4.2; this yields $G = -10.4$. From the Fourier analysis of the residual Hamiltonian on the resonant torus, we have $F = 0.000\,13$. On computing $H_{m\mathrm{k}}$ for tori in the neighbourhood of the resonance, we found that $H_{m\mathrm{k}}$ decreases effectively in a linear manner over the island with $\alpha = -0.008$, and β can well be set to zero. Equation (4.21) now places the island between the δJ_ψ-values $\delta J_\psi^+ \pm \delta J^{\mathrm{sx}} = -0.000\,75 \pm 0.007\,05$, inside the two tori closely bounding the island in very good agreement with Fig. 4.1. The phase plane of Fig. 4.5(b) corresponds to this example.

In Fig. 4.6(a) we show the closed fish orbit at the energy of Fig. 4.1, together with a

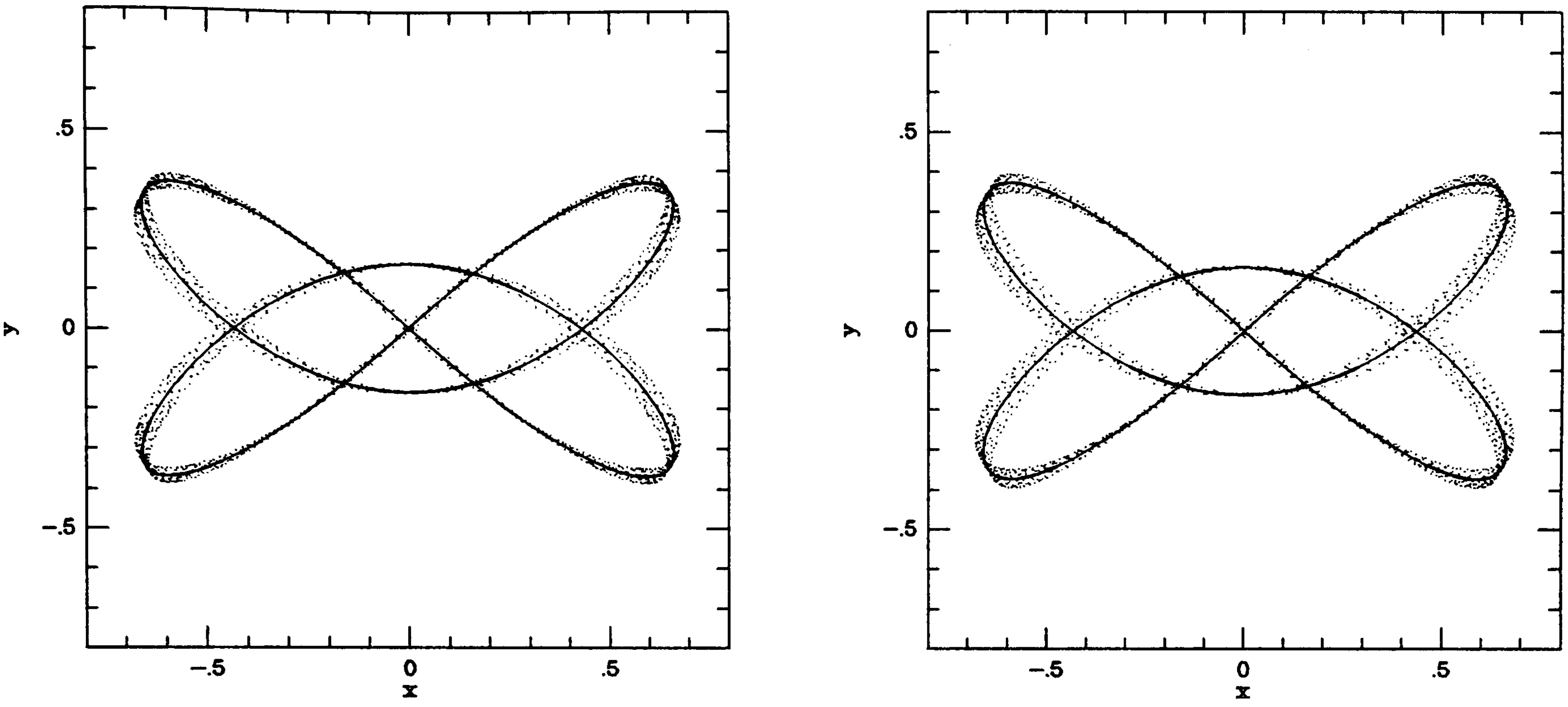

Figure 4.6: (a) The solid curve shows the closed 3:2 orbit; the dots mark a numerically integrated orbit. (b) The same as (a) except that the closed orbit has been obtained by setting the resonant phase ψ to zero on the δJ_ψ^+ -torus, while the non-closed orbit has been obtained from the modified pendulum Hamiltonian (4.19).

non-closed fish orbit. Fig. 4.6(b) shows the approximation to the closed orbit provided by the torus at $\delta J_\psi^+ = -0.000\,75$, and a non-closed orbit obtained by using the modified pendulum Hamiltonian.

The full curves in Fig. 4.7 show some fish islands obtained using the previously given values for the modified pendulum, the largest deviation δJ_ψ being -0.0075 for the larger triangle-shaped island and the long, thin one; for the small triangle inside the larger, it is -0.0052. Numerically integrated consequents, marked as dots, show that the invariant curves predicted by the modified pendulum agree well with the actual invariant curves. Also, the elliptic fixed points $(J_\psi^\pm, \psi^\pm)$ of perturbation theory, corresponding to the stable closed 3:2 orbit, lie close to the centres of these islands, as they should; similarly, the hyperbolic fixed points, corresponding to the unstable closed orbit, are correctly placed in between the islands.

Next, let us study the banana family of Fig. 4.4. From Fig. 4.3 we infer $G = -6.11$, and

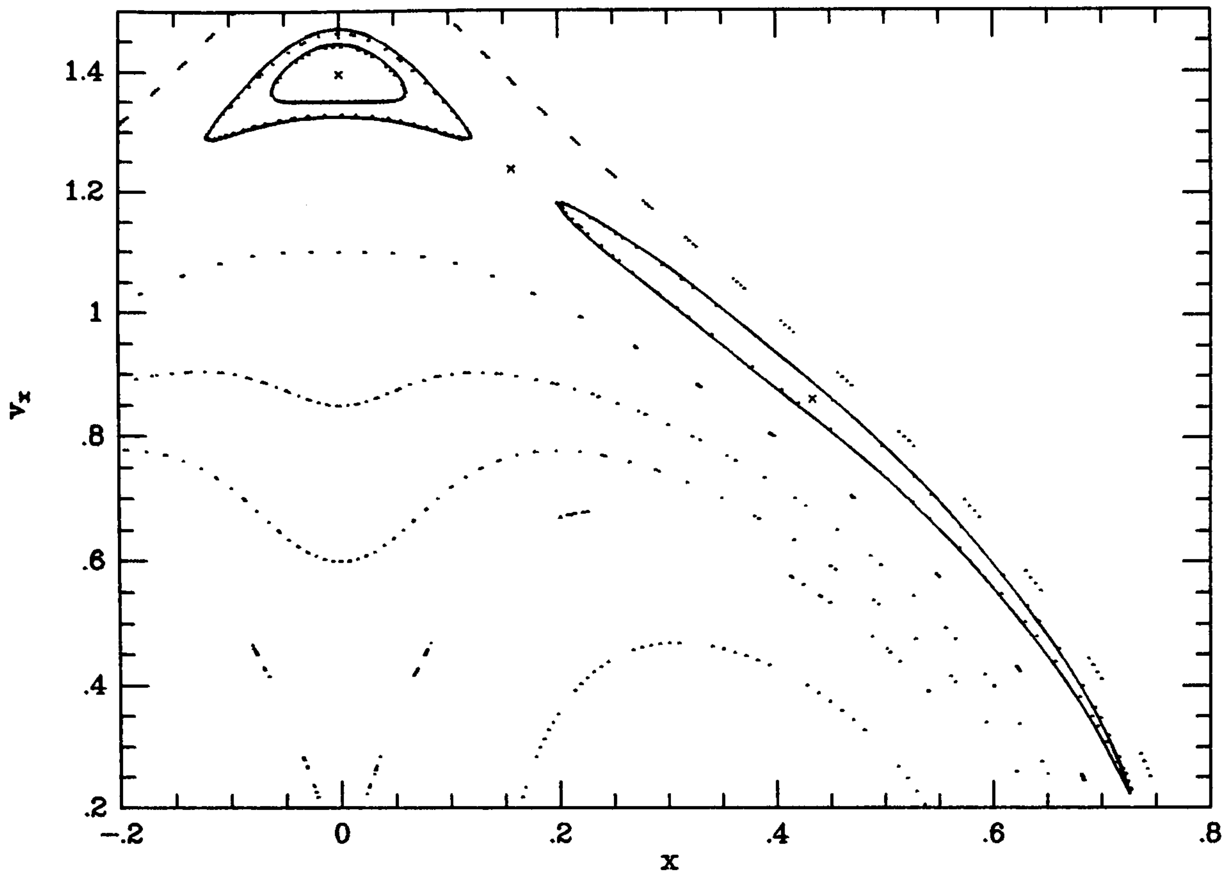

Figure 4.7: A surface of section as in Fig. 4.1; the solid curves show islands obtained using the modified pendulum. Crosses mark the fixed points $(J_\psi^\pm, \psi^\pm)$.

Fourier analysis of the residual Hamiltonian at resonance gives $F = -0.0075$. This is much larger than outside the island; indeed, now $H_{m\mathrm{k}}$ essentially vanishes for tori constructed just outside the island (and the decrease is parabolical to a very good approximation). We did not find any notable asymmetry, and could thus set $\alpha = 0$. A modified pendulum with $\beta = 5.32$ gives now the island half-width $|\delta J^{\mathrm{sx}}| = 0.051$, while the two tori close to the island boundaries lie at $\delta J_\psi = \pm 0.0525$, in very good agreement. The phase plane of Fig. 4.5(c) corresponds to this example ($\Delta J_\psi = 0.053$), and Fig. 4.5(a) is the standard pendulum with the values of F and G used here.

The value of the integrable Hamiltonian $H_0(J_\psi)$, being determined by the interpolated $\omega_\psi(J_\psi)$, is not identical with the average value $\overline{H}$ of H on the torus corresponding to J_ψ. If they differ significantly, an additional coefficient $H_{0,0}(J_\psi)$ equal to this difference must be included in the analysis; this means only adding a constant to G in (4.19). Here the change in H_0 across the island, $\Delta H_0 = |\frac{1}{2}G(0.0525)^2| = 0.008$, happens to equal the difference between the average value of the residual Hamiltonian on the resonant torus and that on the two interpolation tori.

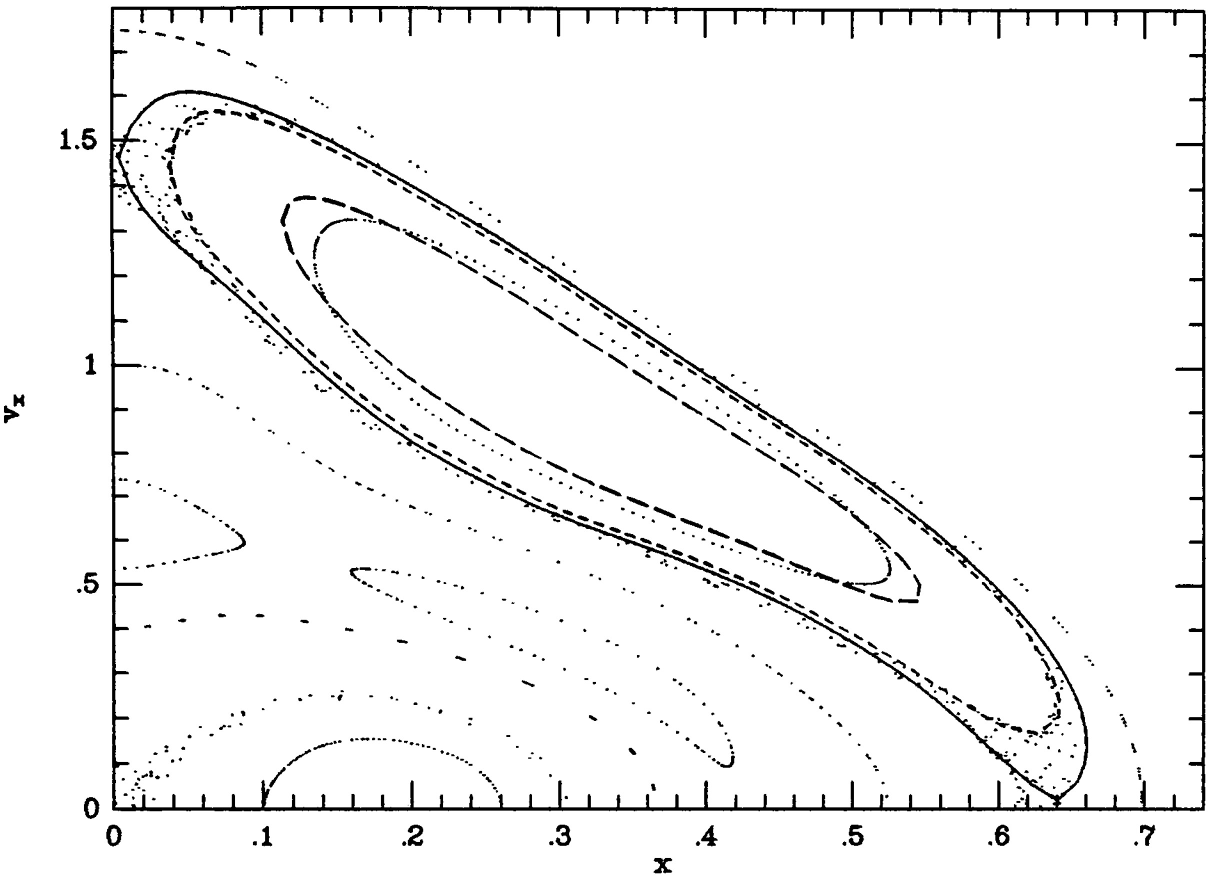

Figure 4.8: The same section as in Fig. 4.4, zoomed in on one quadrant. The dots are numerically integrated consequents, and the dashed lines are contours obtained using perturbation theory.

Fig. 4.8 shows the structure of the bananas' island from perturbation theory and by direct integration of the equations of motion. The contours from the two approaches agree well towards the island's edge, even though the region between the outermost two contours is chaotic, and the motion there spreads to the stochastic layers of all four islands in the chain seen in Fig. 4.4. The inner contours do not agree quite as well, but the location and orientation of the contour from perturbation theory are correct.

Generally speaking, it seems that, when a practical and concise approximation of (4.14) is needed, the modified pendulum is a consistent choice even for the large minor-orbit families. For example, adding one or two further H_n-terms in the case of the inner island contours of Fig. 4.8 does not refine the result. Improvement of the fit requires adding higher orders to the second-order approximation to the δJ_ψ-variation. In fact, in the second-order approximation the successive contours approaching the island centre approximately retain the shape of the island, whereas the directly integrated contours become less elongated closer to the centre, implying that $H_{m\mathbf{k}}$ of the ideal H_0 would decrease according to a higher than quadratic power law. But if one uses a higher-order approximation, then the interpolation would surely have to be of higher order as well, especially because in linear interpolation the decrease of $H_{m\mathbf{k}}$ seems to be essentially parabolical. We have experimented with some higher-order interpolations, but they undermine the simplicity of the model – their main fault is that they produce complicated δJ_ψ-dependence in $H_{m\mathbf{k}}$ and sometimes some instability problems. In principle, the tori defining H_0 can be obtained by any interpolation scheme so long as they collectively behave like those of a bona fide integrable Hamiltonian. Applying (4.14) should then produce proper results, but at the price of losing a simple pendulum-type model. We have presented here the best second-order model, and in all likelihood it seems that for a better description one would need a considerably more complicated model.

As was shown in chapter 2, the tori for banana orbits or other minor-orbit families can be constructed directly by employing the general point transformation of appendix D. Thus, we have two complementary methods at our disposal for studying the phenomena associated with minor-orbit families.

Chapter 5

From regular to stochastic motion

One of the examples of the previous chapter (Fig. 4.4) contained a non-negligible layer of stochasticity associated with an island boundary; in this chapter, we experiment with the torus-construction scheme in larger chaotic regions that eventually spread to global stochasticity. The main point is to examine whether it is possible to create invariant tori in the chaotic regions of systems where fewer and fewer KAM-tori exist. In conjunction with this, we shall extend the torus-construction scheme to time-irreversible Hamiltonians – chaos can be created in a controllable way by using rotating frames of reference. Also, we shall study the transition from local to global stochasticity in the light of the resonance-overlap criterion (Chirikov 1979).

We continue using the logarithmic potential (2.5), but in a modified form. Motion in the unadorned potential (2.5) is notoriously regular: when the parameter q is made smaller, more and more resonant islands appear, but stochasticity is always contained in local layers, never spreading to larger regions of phase space. One way to create chaos is to set (2.5) rotating with an angular speed $\boldsymbol{\Omega} = \Omega\hat{\boldsymbol{z}}$, where $\hat{\boldsymbol{z}}$ is the z-axis unit vector, and study the Hamiltonian in the rotating frame of reference. Astrophysically this is of interest, because from observations it is known that the potential figures of galaxies typically rotate.

5.1 Torus construction in a rotating frame of reference

First, we write the Hamiltonian in a rotating frame of reference. This can be done by using a time-dependent canonical transformation between the inertial and the rotating frame. Let us denote the canonical variables in the rotating frame by $(\boldsymbol{p}, \boldsymbol{x})$ and those in the inertial frame by $(\boldsymbol{p}', \boldsymbol{x}')$. Now

$$\boldsymbol{x}'(\boldsymbol{x}, t) = \boldsymbol{x} \cos \Omega t + \hat{\boldsymbol{z}} \times \boldsymbol{x} \sin \Omega t. \tag{5.1}$$

(In the three-dimensional case, (5.1) holds for x- and y- coordinates; z and p_z are, of course, invariant.) This transformation can be generated by the function $F(\boldsymbol{p}', \boldsymbol{x}, t) = \boldsymbol{p}' \cdot \boldsymbol{x}'(\boldsymbol{x}, t)$; from $\partial F / \partial \boldsymbol{p}'$ we have (5.1), and

$$\boldsymbol{p} = \frac{\partial F}{\partial \boldsymbol{x}} = \boldsymbol{p}' \cos \Omega t - \hat{\boldsymbol{z}} \times \boldsymbol{p}' \sin \Omega t. \tag{5.2}$$

The Hamiltonian H in the rotating frame follows from the inertial Hamiltonian H' by

$$H = H' - \frac{\partial F}{\partial t} = \frac{1}{2} p^2 + \boldsymbol{\Omega} \cdot (\boldsymbol{p} \times \boldsymbol{x}) + \Phi(\boldsymbol{x}). \tag{5.3}$$

The Coriolis term makes this Hamiltonian asymmetric under time reversal. Also, the canonically conjugate momenta $\boldsymbol{p}$ are no longer just the Cartesian coordinate velocities as in the inertial space; from the Hamiltonian equations of motion we have

$$\dot{\boldsymbol{x}} = \frac{\partial H}{\partial \boldsymbol{p}} = \boldsymbol{p} - \boldsymbol{\Omega} \times \boldsymbol{x}. \tag{5.4}$$

The Newtonian equations of motion can be obtained by taking the time derivative of (5.4) and combining that with $\dot{\boldsymbol{p}} = -\partial H / \partial \boldsymbol{x}$:

$$\ddot{\boldsymbol{x}} = -\frac{\partial \Phi_{\text{eff}}}{\partial \boldsymbol{x}} - 2(\boldsymbol{\Omega} \times \dot{\boldsymbol{x}}), \tag{5.5}$$

where

$$\Phi_{\text{eff}} = \Phi - \frac{1}{2} \Omega^2 (x^2 + y^2). \tag{5.6}$$

The value of the Hamiltonian (5.3) is called the Jacobi constant E_{J}, which can also be written in terms of the Cartesian non-canonical variables as $E_{\text{J}} = \frac{1}{2} \mid \dot{\boldsymbol{x}} \mid^2 + \Phi_{\text{eff}}$.

What are the major-orbit families now, and what kind of a set of toy Hamiltonians is needed? As is well known (e.g. Binney & Tremaine 1987, Martinet & Udry 1990), rotating barred potentials produce four kinds of major families, usually labelled x_1, x_2 and x_4 corresponding to loop orbits around stable closed ones (x_3 corresponds to an unstable orbit), plus the box orbits; the usual discussion of their shapes, 'characteristic curves' and other properties of astrophysical interest can be found in the aforementioned references and many others. We shall not pay much attention to such details here, because we want to concentrate on the general aspects of the problem.

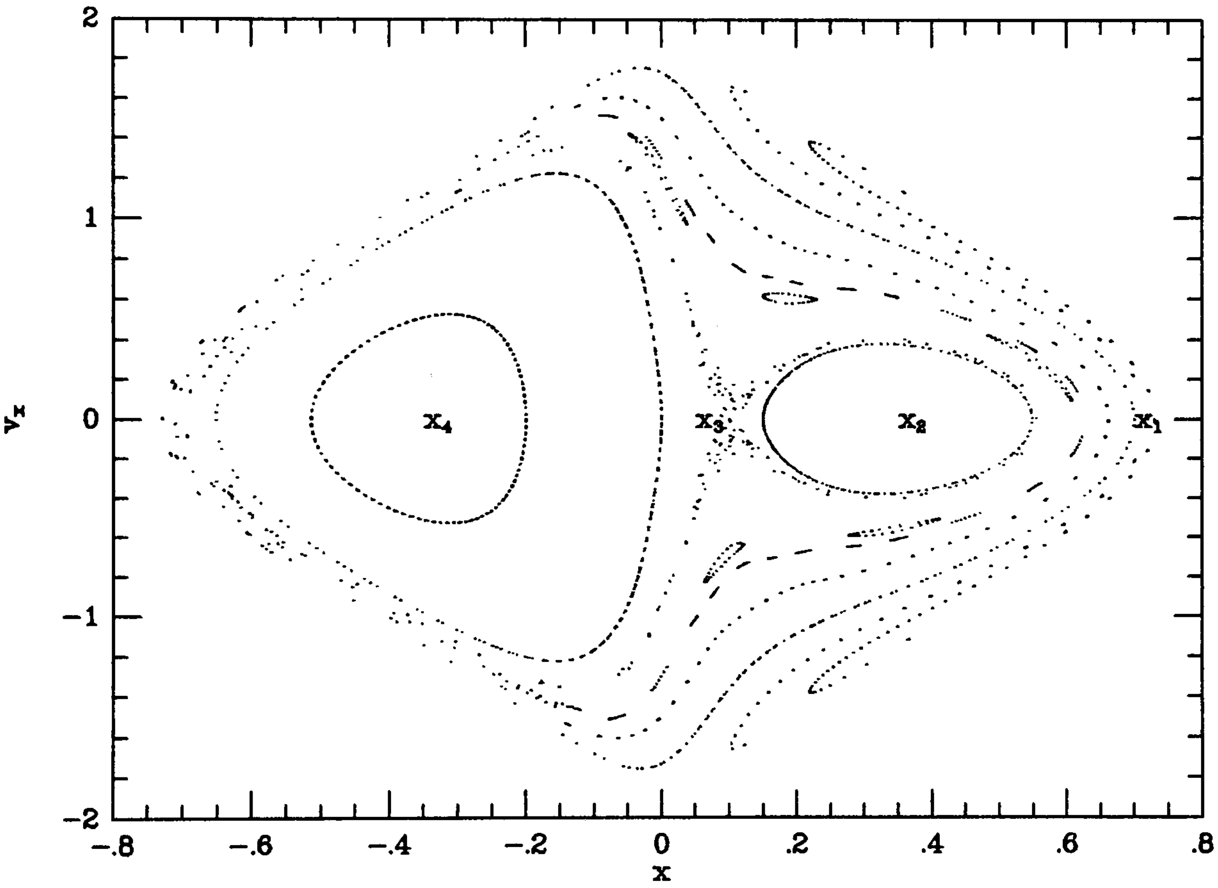

Figure 5.1: $(x, \dot{x})$ at $H = -0.315$ with $\Omega = 0.3$. The orbit families x_1–x_4 (x_3 at an unstable point) are marked.

The x_1-family orbits often cross themselves in a rather inconvenient manner that limits even the 'optimal' point transformation technique of appendix D to a small set of orbits near the parent closed orbit. Fortunately these orbits usually occupy a very small portion of the phase space, and in surfaces of section, such as Fig. 5.1, their curves form a separate group of their own, starting at the 'borderline' of the curves of the other families. Fig. 5.1 shows the $(x, \dot{x})$-section of the logarithmic potential with $q = 0.9$ at $H = -0.315$ and rotation speed $\Omega = 0.3$. (At $y = 0$, the momentum p_1 coincides with $\dot{x}$.) The x_2-family (prograde loops) and x_4-family (retrograde loops) can be handled with the isochrone H_I, and the boxes with the harmonic oscillator H_H. Previously, the target action values for the sets of major-type tori could be arranged to join smoothly in a natural way in the action space; now, the corresponding three action-patches overlap and leave gaps (Binney & Spergel 1984). Thus, if one creates an integrable Hamiltonian from the constructed tori, it applies to each major family separately.

The form of the generating function S

Since we have lost the time symmetry, the generating function (2.1) reads now

$$S(\boldsymbol{\theta}, \boldsymbol{J}') = \boldsymbol{\theta} \cdot \boldsymbol{J}' + 2 \sum_{\mathbf{n}>0} [S^R_\mathbf{n}(\boldsymbol{J}') \sin(\boldsymbol{n} \cdot \boldsymbol{\theta}) + S^I_\mathbf{n}(\boldsymbol{J}') \cos(\boldsymbol{n} \cdot \boldsymbol{\theta})], \tag{5.7}$$

where the superscripts R and I refer to, respectively, the real and the imaginary parts of the coefficients $S_\mathbf{n}$ used in the exponential form (2.1). The action formula (2.3) becomes

$$\boldsymbol{J} = \frac{\partial S(\boldsymbol{\theta}, \boldsymbol{J}')}{\partial \boldsymbol{\theta}} = \boldsymbol{J}' + 2 \sum_{\mathbf{n}>0} \boldsymbol{n}\, [S^R_\mathbf{n}(\boldsymbol{J}') \cos(\boldsymbol{n} \cdot \boldsymbol{\theta}) - S^I_\mathbf{n}(\boldsymbol{J}') \sin(\boldsymbol{n} \cdot \boldsymbol{\theta})], \tag{5.8}$$

and the angle formula (3.2) changes accordingly.

Naturally, the geometric symmetries affect the set of coefficients $S_\mathbf{n}$ as before. Let us see what happens under the harmonic-oscillator and isochrone mappings. In $(x, \dot{x})$ and $(y, \dot{y})$ surfaces of section, we can see that the invariant curves are reflection symmetric about the coordinate axis (it is rather obvious that this should be the case). Under the harmonic-oscillator mapping, such a section is created when one of the toy angles θ_j is zero. In terms of the other toy angle θ_i, the symmetry about the coordinate axis corresponds to a reflection symmetry about the 'axis' of $\theta_i = \pm\pi/2$, i.e. to terms of the form $\cos[n(\theta_i + \pi/2)]$, when considering the Fourier series (5.8) for the toy action. When such a series is given in terms of $n\theta_i$, we have $\cos(n\theta_i)$ for even-valued n and $\sin(n\theta_i)$ for odd-valued n. Combining this result for the two dimensions, we see that $S(\boldsymbol{\theta}, \boldsymbol{J}')$ contains only terms $S^R_{\mathbf{n}}$ with both indices of $\boldsymbol{n}$ even (as before), and terms $S^I_{\mathbf{n}}$ with both indices of $\boldsymbol{n}$ odd. Such a series gives a rather curious pattern of repeated $\boldsymbol{\theta}$-patches, combining reflection and translation symmetries. For example, the values of the Fourier series in the patch $[-\pi/2, \pi/2] \times [-\pi/2, \pi/2]$ are invariant under the transformations

$$\boldsymbol{\theta} \rightarrow \begin{cases} \boldsymbol{\theta} + \boldsymbol{\pi} \\ \pi - \theta_x, -\theta_y \\ -\theta_x, \pi - \theta_y . \end{cases} \tag{5.9}$$

The 'tiling pattern' from these fills the entire $\boldsymbol{\theta}$-space, so only a quarter of the whole $2\pi \times 2\pi$-space need be explored (and only a half of the border points of the repeating section). Under the isochrone mapping, time-asymmetry adds $S^I_{\mathbf{n}}$ that are like $S^R_{\mathbf{n}}$ from the time-symmetric case, i.e., coefficients with odd-valued indices n_φ vanish; this can be seen by considering e.g. a $(\varphi, J_\varphi = p_\varphi)$-plot at $p_r = 0$ $(\theta_r = 0, \pi)$.

5.2 Creating tori in chaotic regions

Now we are in a position to create tori for (5.3) using the least-squares or orbit integration methods. (Note that the Hamiltonian that is computed in the process is specifically (5.3) as a function of the canonical variables $(\boldsymbol{p}, \boldsymbol{x})$, not the Jacobi constant as a function of $(\dot{\boldsymbol{x}}, \boldsymbol{x})$.) In Fig. 5.2 we show the $(x, \dot{x})$-section at $H = -0.253, q = 0.8$ and $\Omega = 0.4$. In the chaotic region $x < 0.6$, the solid lines are the invariant curves of box-type tori constructed at (from right to left) $\boldsymbol{J}' = (0.294, 0.19)$, $(0.28, 0.20)$, $(0.26, 0.215)$, and $(0.235, 0.235)$. The outermost solid line corresponds to a torus constructed with the orbit integration method of chapter 3, which assigns to it the actions $\boldsymbol{J}' = (0.4305, 0.0932)$. In this case, the box orbits are almost rectangular in configuration space, so we do not use an additional point transformation as earlier. In the Fourier series for S, the coefficients $S^R_{\mathbf{n}}$ and $S^I_{\mathbf{n}}$ decrease in $\boldsymbol{n}$ in a similar fashion – the number of coefficients is thus about two times that needed for time-symmetric

Hamiltonians. The need for many $S^I_{\mathbf{n}}$ is obvious from the asymmetry of the surface of section with respect to the momentum axis.

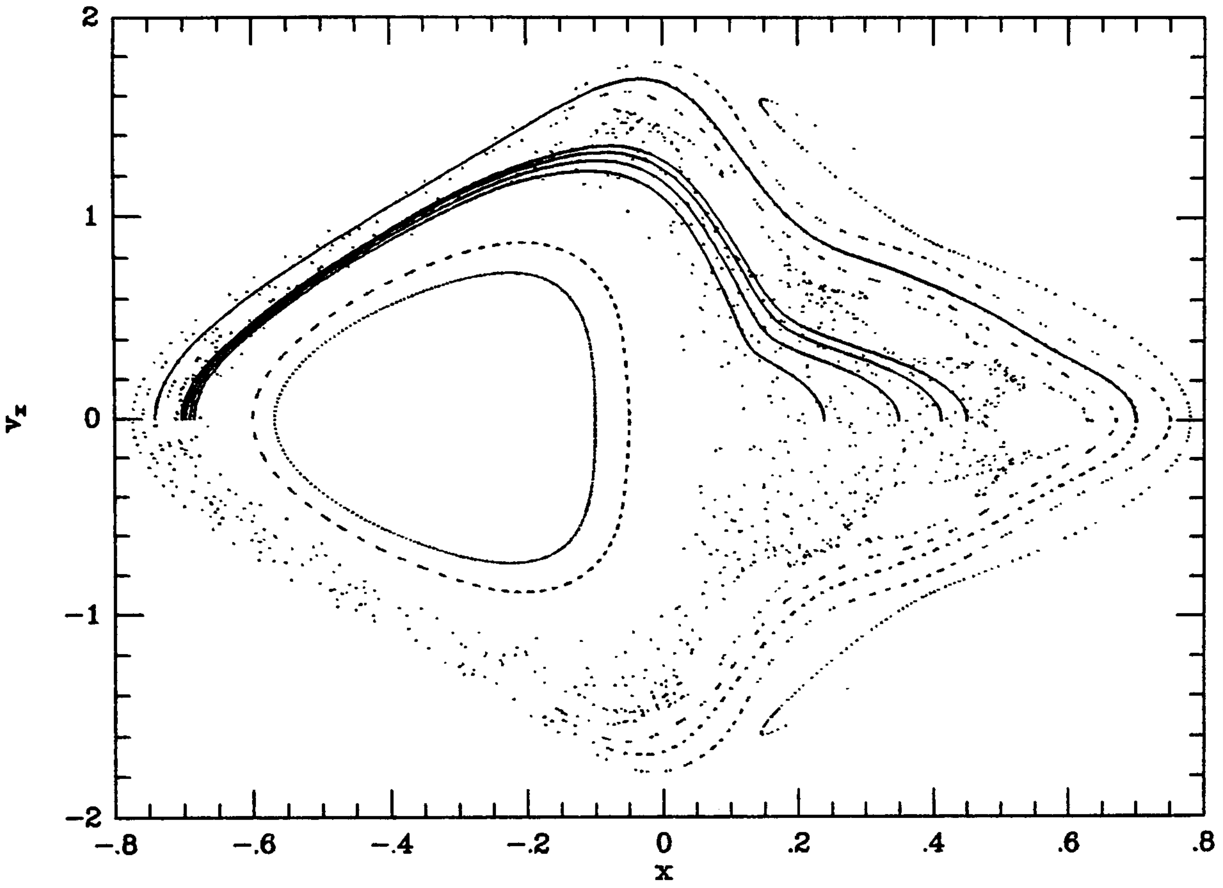

Figure 5.2: $(x, \dot{x})$ at $H = -0.253$ with $\Omega = 0.4$. The solid curves are those of constructed tori.

The least-squares method is thus able to construct good tori even in wide chaotic regions. Note how the curves of the constructed tori mimic well the behaviour of the invariant curves that 'used to' lie in the now chaotic region before the fast rotation dissolved them (cf. Fig. 5.1). Also, the target angles and frequencies can be assigned with the method of chapter 3, as usual. Note that even though the motion on the integrated strips certainly does not occur on a torus, and cannot thus coincide with a linear increase in the created target angles, the linear strips in Fig. 5.3, corresponding to $\boldsymbol{J}' = (0.28, 0.20)$, are still very close to target angle points corresponding to the integrated strips. In other words, locally or for short time periods, motion on our constructed torus still approximates well motion in H.

An interesting thing about the chaotic region is that it stops very sharply at the outer

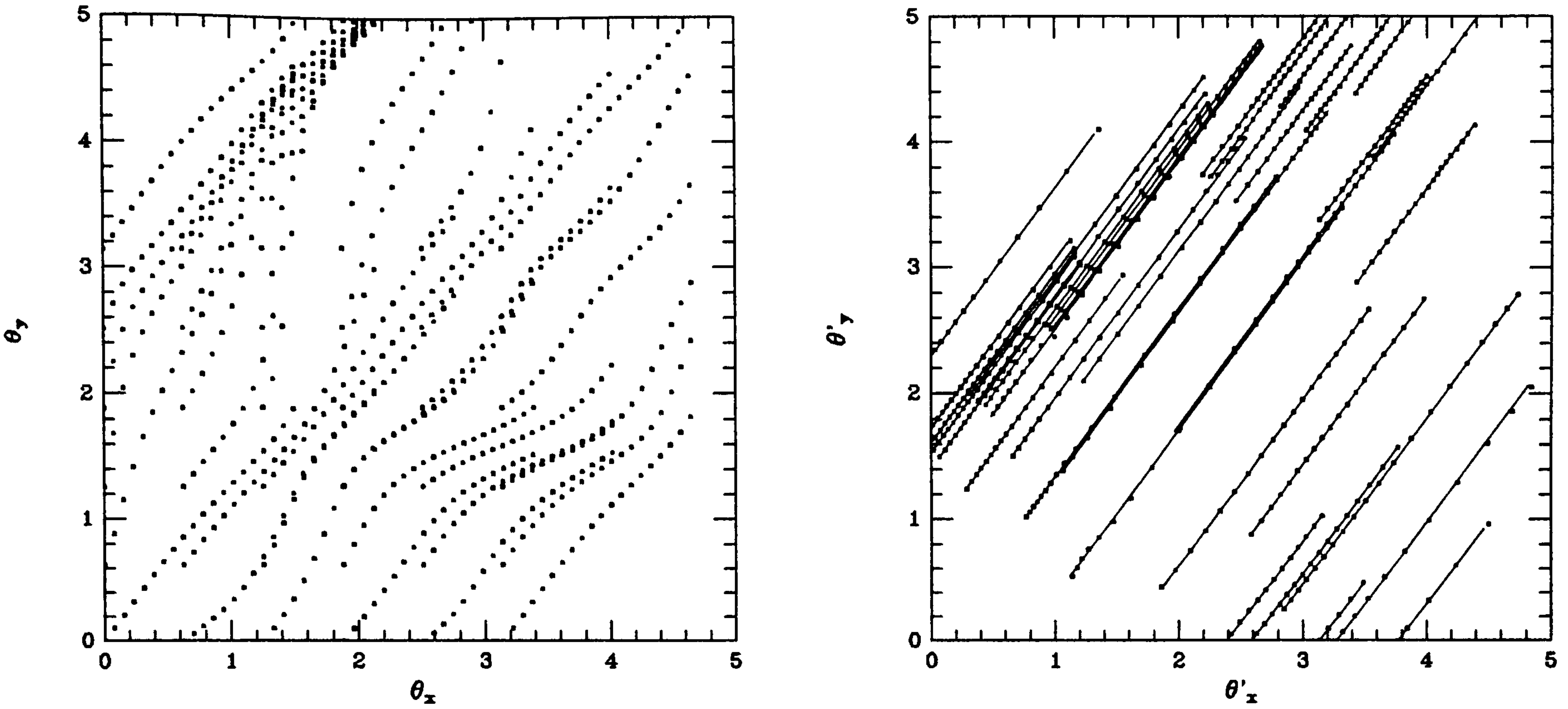

Figure 5.3: The toy and target angles for one of the tori in Fig. 5.2, shown as in Fig. 3.1.

edge of the large 7:5 islands seen in Fig. 5.2. A series of invariant curves starts immediately after the islands, going then over to the separate region of the x_1-family. This transition produces difficulties in torus construction. The x_1-orbits cannot be mapped as boxes with the harmonic oscillator any more, because they leave a gap around the origin in configuration space. Naturally, the box orbits become increasingly difficult to handle close to x_1-orbits, in the same way as with the family transition from boxes to loops earlier. In the box/loop-case, we knew a useful point transformation. In this case, however, no such transformation is known. In configuration space, the boundary of a box orbit close to the x_1-family is roughly rectangular all right, but the Coriolis force causes motion *inside* it to be far from that of the harmonic oscillator. Thus there is no convenient point transformation, like in the earlier case of box orbits, that would map a point in toy configuration space to one in target space (one-to-one) and lighten the work of the generating function S. Again, we are reminded of the fact that in theory S can be infinite, but in practice it must converge rapidly. With similar resolution and number of coefficients as for the four tori above, the tori produced by the least-squares method do not properly coincide with the invariant curves between the

7:5-islands and the x_1-family.

Fortunately, the orbit integration method can be employed for the invariant curves. In this case, we had no difficulty in obtaining a perfect fit to the integrated consequents, as can be seen from the outermost solid curve in Fig. 5.2. The torus was constructed with 32 × 32 points in $\pi \times \pi$, i.e. maximum order 31; sparser grids produced slightly wiggly invariant curves. The Fourier coefficients of the obtained generating function S decrease steadily in both dimensions, so there are no 'surprise coefficients'. In this case, the least-squares method is in practice inefficient simply because S contains many coefficients. Again, this example shows the useful complementary nature of the two torus-construction techniques.

5.3 Transition to global chaos

The resonance overlap criterion

And, as they say, now for something completely different. Chirikov's overlap criterion (Chirikov 1979) is often mentioned when discussing transition from local to global chaos. Proper analytic criteria for this transition are not easy to invent, few have been proposed (see e.g. Lichtenberg & Lieberman 1983), and the overlap criterion is the most pleasing one because of its simplicity and intuitive appeal. However (as Chirikov already emphasized), where accuracy is concerned, it is often a rather rough one. In this section, we first test the overlap criterion with a specially prepared near-integrable 'laboratory' Hamiltonian that is well under our control: we consider a system with two degrees of freedom and just two resonances, resembling the old problem of two coupled oscillators. Similar systems were studied by Walker & Ford (1969) (though not really using the concept of overlap), and Chirikov (1979) gives an example with two resonances that are much larger than any others in $1\frac{1}{2}$ degrees of freedom. I have attempted to create the absolutely simplest possible nontrivial near-integrable system, so that we always know where we stand.

The reason for this digression is partly that I have always wanted to know just how accurate the overlap criterion is in its simplest two-resonance form; and partly, that below we will try to show how chaos spreads in our model systems by merging one resonance after another in a rather orderly fashion. Also, by computing how large the islands of perturbed resonant tori would be if they were left alone, and whether they seem to overlap, we check if the chaotic regions are consistent with our created integrable Hamiltonian in the same way that minor-orbit families were shown to be.

Let us now construct a near-integrable Hamiltonian H exhibiting easily 'controllable' resonant trapping and stochasticity. We take any integrable Hamiltonian $H_0(\boldsymbol{J})$, and add terms of the form used in secular perturbation theory:

$$H = H_0(\boldsymbol{J}) + 2\sum_{\mathbf{n}} H_{\mathbf{n}} \cos(\boldsymbol{n} \cdot \boldsymbol{\theta}), \tag{5.10}$$

where the factor 2 is added because we have used it earlier, and $H_{\mathbf{n}}$ are constants; here we use two different $\boldsymbol{n}$. The actions and angles $(\boldsymbol{J}, \boldsymbol{\theta})$ are those of any integrable system of two degrees of freedom; by definition, they are action-angles of H_0 as well. It does not matter very much how $(\boldsymbol{J}, \boldsymbol{\theta})$ are transformed to phase-space points $\boldsymbol{w}$; here we use (surprise, surprise) the harmonic-oscillator and isochrone tori. In this sense, one can imagine two oscillators (either in the Cartesian directions or in radial and azimuthal directions) living lives of their own, until a coupling term traps them to a coexistence near a resonance. Using the form (5.10) with only a limited number of resonance terms saves us from the complex consequences of a general 'KAM-style' perturbation with an infinite Fourier series. For example, with only one resonant term there will not be myriads of little islands corresponding to each resonant torus popping up everywhere, and no second-order resonant islands around the first-order island. With two resonances, the only interesting region will be between the two resonant islands; outside them, phase-space looks regular and the original invariant tori are only somewhat distorted by the appearance of the islands.

From the Hamiltonian equations of motion it follows that

$$\begin{aligned}
\dot{\boldsymbol{\theta}} &= \frac{\partial H}{\partial \boldsymbol{J}} = \frac{\partial H_0}{\partial \boldsymbol{J}} \equiv \boldsymbol{\omega}(\boldsymbol{J}(t)), \\
\dot{\boldsymbol{J}} &= -\frac{\partial H}{\partial \boldsymbol{\theta}} = 2\boldsymbol{n} \sum_{\mathbf{n}} H_{\mathbf{n}} \sin(\boldsymbol{n} \cdot \boldsymbol{\theta}(t)).
\end{aligned} \tag{5.11}$$

These equations determine the development of the system; the transformation $(\boldsymbol{J}, \boldsymbol{\theta}) \rightarrow \boldsymbol{w}$ projects it to ordinary phase-space coordinates, if so wished. (Since $\boldsymbol{J}$ and $\boldsymbol{\theta}$ are separated in (5.10), its numerical integration can be quickly done by using a symplectic integrator just as when integrating orbits in potentials.) As described in chapter 4, if there is a resonance with the ratio $\gamma \equiv \omega_2/\omega_1 = r/s$, the island size (half-width) δJ_ψ for $\boldsymbol{k} = (r, -s)$, in units of $\boldsymbol{k} \cdot \boldsymbol{J}$, is

$$\delta J_\psi = 2\sqrt{2\,|\,H_{\boldsymbol{k}}/G\,|}, \tag{5.12}$$

where

$$G = \boldsymbol{k} \cdot \frac{\partial}{\partial \boldsymbol{J}}(\boldsymbol{k} \cdot \boldsymbol{\omega}). \tag{5.13}$$

The overlap criterion tells us to expect chaotic behaviour when two resonant islands are large enough for their boundaries to touch each other. Of course, this statement is not rigorous – we know that in near-integrable systems the island boundary is more or less fuzzy: generally, there is a local (sometimes negligible) layer of stochasticity associated with the island separatrix. Graphically, however, this concept is quite clear: it relates directly to what we see in a plotted surface of section. Analytically, the island widths (5.12) are derived by considering each resonance separately by using the averaging principle. Thus, the real overlap condition for the complete problem can be somewhat different from that derived from separate resonance calculations. Also, in two or more degrees of freedom, the overlapping of resonance widths (5.12) is not uniquely defined, because the widths are given in $\boldsymbol{J}$-space along lines (tangents to the surface of constant H_0) whose directions are determined by the indices $\boldsymbol{n}$: one must define a surface onto which the widths are projected. In this chapter, we follow a simple convention: because our $\boldsymbol{J}$-space is two-dimensional, we simply calculate where the lines corresponding to the two resonances cross, and how large the coefficients H_{n} must be to reach that point. This prediction is then compared with the numerical solution of (5.11) to see how accurate the overlap criterion is. Thus, from (5.12) we have that the critical value for H_{k} of a resonance $\gamma^{(1)}$ to reach the overlap point with another resonance at $\gamma^{(2)}$ is

$$H_{\mathrm{k}} = \frac{G}{8\,|\boldsymbol{k}\,|^2}(J_{\mathrm{c}} - J_2^{(1)})^2(1+\gamma^{(1)^2}), \tag{5.14}$$

where

$$J_{\mathrm{c}} = \frac{J_2^{(2)}\gamma^{(2)} - J_2^{(1)}\gamma^{(1)} + J_1^{(2)} - J_1^{(1)}}{\gamma^{(2)} - \gamma^{(1)}}, \tag{5.15}$$

and $(J_1^{(i)}, J_2^{(i)})$ refer to the value of $\boldsymbol{J}$ at the resonance i.

To keep the matters simple, we employ an underlying integrable Hamiltonian of the form

$$H_0(\boldsymbol{J}) = a(J_1 + J_a)^{\alpha} + b(J_2 + J_b)^{\beta}, \tag{5.16}$$

where all the parameters $a, b, J_a, J_b, \alpha, \beta$ are larger than zero, and $\alpha, \beta \neq 1$, i.e., the harmonic oscillator will not do. This is because we want to have a nonlinear system with no intrinsic resonance and $G \neq 0$. This Hamiltonian serves well as a general one: with $\alpha, \beta < 1$ we have concave isoenergetic curves in $\boldsymbol{J}$-space, and with $\alpha, \beta > 1$ convex ones. Also, from the determinant (Arnold 1989)

$$\begin{vmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{vmatrix} \neq 0, \tag{5.17}$$

where subscripts denote differentiation, we have that H_0 is nondegenerate, and from

$$\begin{vmatrix} H_1 & H_2 & H_0 \\ H_{11} & H_{12} & H_1 \\ H_{21} & H_{22} & H_2 \end{vmatrix} \neq 0 \tag{5.18}$$

(when $(\alpha - 1)(\beta - 1) \neq 0$), we have that H_0 is isoenergetically nondegenerate.

We use here $\alpha = \beta = 2$, so $\gamma_{\mathrm{min}} = bJ_b/\sqrt{a(H_0 - bJ_b^2)}$, and $\gamma_{\mathrm{max}} = \sqrt{b(H_0 - aJ_a^2)}/aJ_a$ for a given curve of constant H_0. The frequencies are linear in $\boldsymbol{J}$ and $G(\boldsymbol{k}) = 2ar^2 + 2bs^2$.

For this choice, the second-order series expansion of the Hamiltonian in secular perturbation theory is exact, so the *only* approximation in applying perturbation theory to (5.10) is the averaging over the fast angle variable (chapter 4). Therefore we can say that the overlap principle is tested under well controlled conditions. We use the harmonic-oscillator action-angles to plot the surfaces of section; this is essentially the same as plotting the surfaces of section $\theta_j = 0$ using (J_i, θ_i) as polar coordinates. In the unperturbed case, each invariant curve is thus a harmonic-oscillator curve, but their 'packing density', when being nested at a given value of H_0, is different from the harmonic-oscillator one.

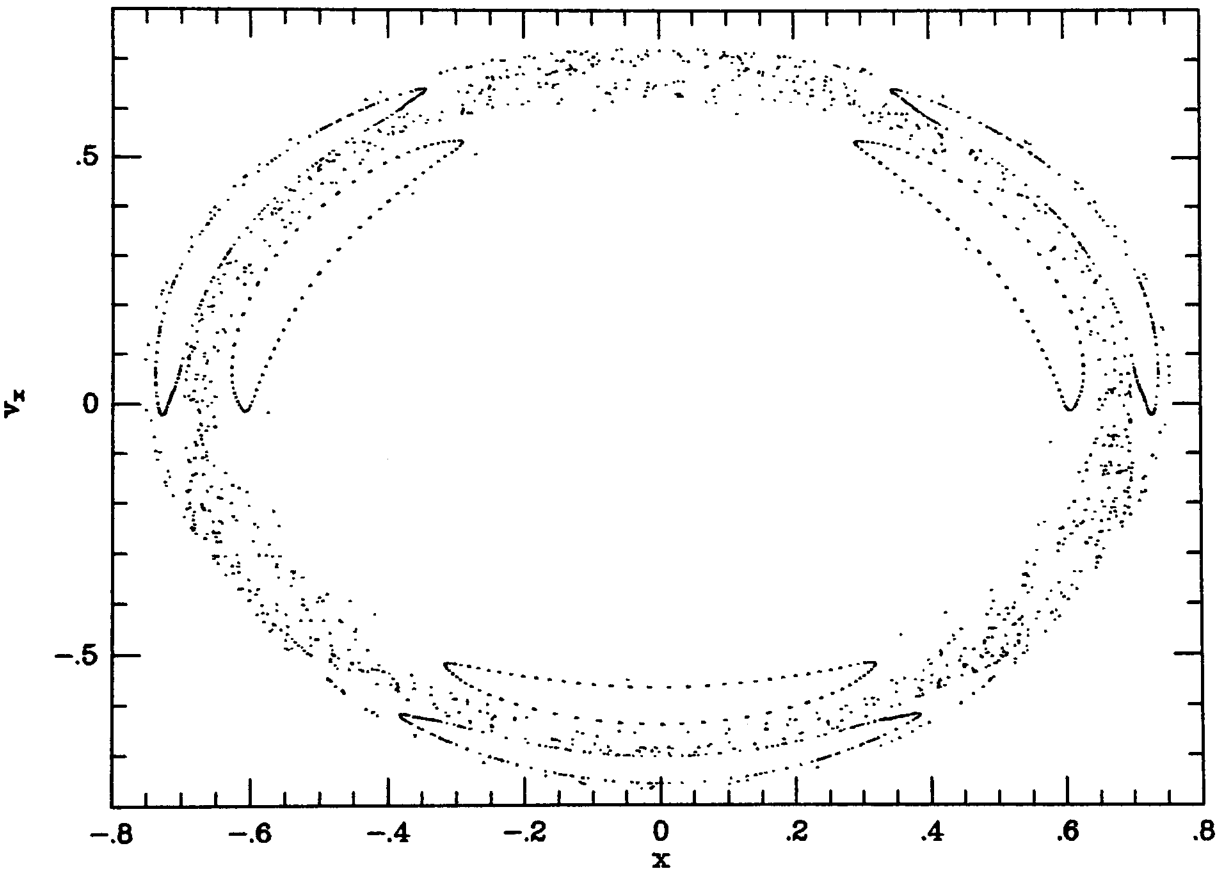

Figure 5.4: $(x, \dot{x})$ of (5.16) at 2/3 of the predicted critical perturbation strength.

Let us study the interaction of 3:4 and 3:5 resonances at $a = 2, b = 1, J_a = J_b = 0.5$ and $H = 2$. From (5.14) we obtain $H_{3,-4} = 0.0015 = H_{3,-5}$ for the critical values. However, the chaos starts with smaller values for H_{k} than the estimated ones: from about $H_{\mathrm{k}} = 0.0008$, i.e. island sizes 75% of the estimated ones, a joint stochastic region can be seen to form, although at first the spread is very slow and long integration times are needed to reveal the nature of the motion. In Fig. 5.4 we show the resonant islands and a chaotic region between them in an $(x, \dot{x})$-surface of section at $H_{\mathrm{k}} = 0.001$.

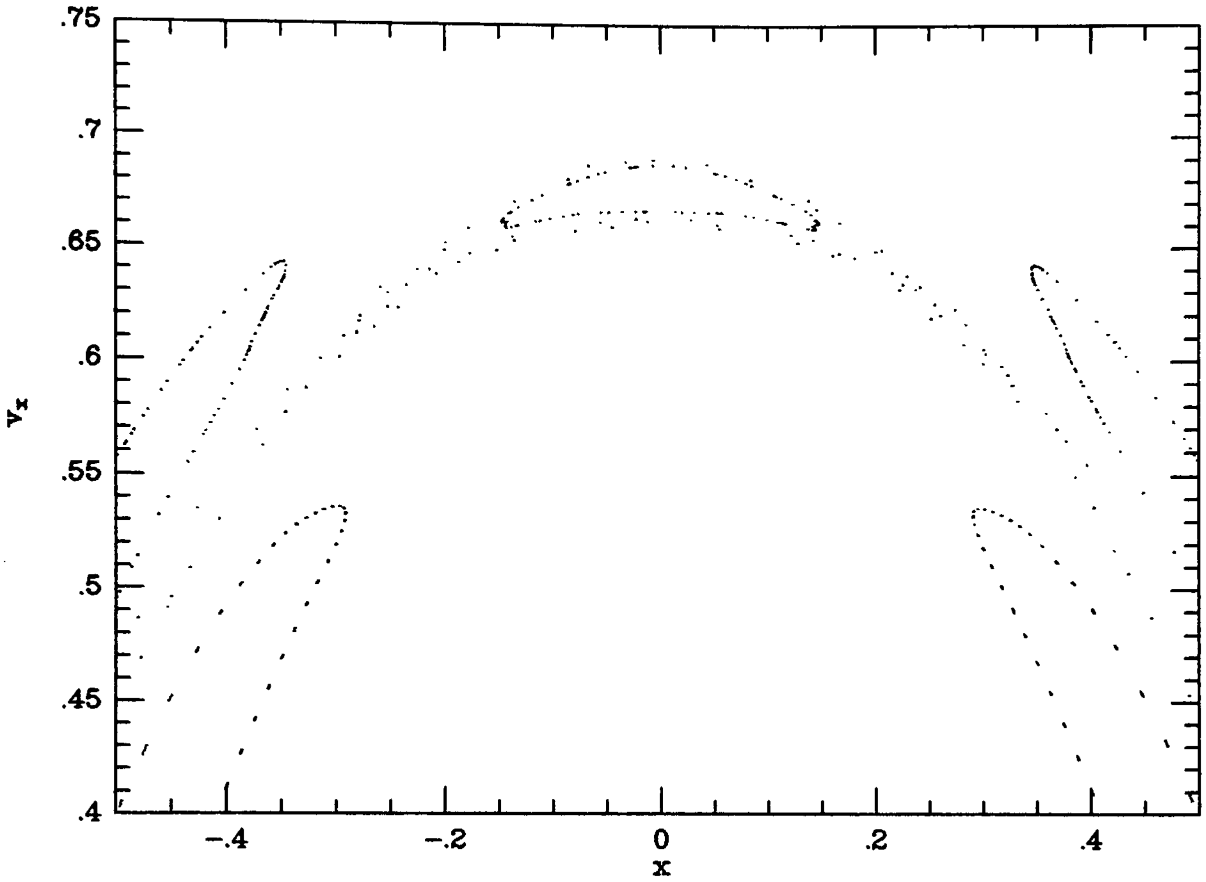

Figure 5.5: A detail of an $(x, \dot{x})$ with a smaller perturbation strength, producing the 2:3 ghost resonance.

A very interesting phenomenon is apparent even before chaos sets in. In Fig. 5.5, we show an enlarged part of the $(x, \dot{x})$ at $H_{\mathrm{k}} = 0.0008$: between the two main resonances, there is suddenly a new island chain. This chain corresponds to a triplet of 2:3 resonant orbits (this can easily be seen by comparing $(x, \dot{x})$ with the corresponding $(y, \dot{y})$-section – the complete chain corresponds to the multiple 6:9, while a starting point near the center of an island produces a single 2:3 group). But according to secular perturbation theory, such a chain needs the term $H_{6,-9}$ to exist, and there is no such term in our series. (Indeed, by using just the term $H_{6,-9}$ in (5.10), an exactly similar triplet of 2:3 islands can be created.) This 'ghost resonance' has formed because of the interplay between our two main resonances, not because of a driving term of its own. The value 2/3 is very close to the mean value 0.675 of the two resonances; also, in action space, the actions corresponding to 2:3 are very close to those of the point of overlapping associated with (5.14). When the two resonances are large enough, secular perturbation theory fails to take into account the complex behaviour of an orbit being pulled to and fro between the attracting resonances. Sometimes, before the onset of actual chaos, the 'frequency uncertainty' of the orbit results in its going in circles around a new ghost resonance, not because that resonance traps the orbit but because it just cannot make up its mind. In the spirit of the overlap criterion, one can see why the chaos starts earlier than estimated: the two resonant islands actually overlap with the ghost island first. (Similarly, with general perturbed Hamiltonians there are always small resonant islands between any two under consideration.) There can be several ghost resonances. For example, changing the sign of one $H_{\mathbf{k}}$ of Fig. 5.4 causes the corresponding island chain to rotate an angle π, and again (very thin) ghost resonances form: the 2:3 ghost chain rotates $\pi/2$ and close to it there is a 9:13 chain (and $9/13 \approx 0.69$ is again close to 0.675).

This example illustrates why the overlap criterion gives just a rough estimate – the onset of chaos is a dynamical process rather than the nice graphical one associated with the concept of overlapping. In the complete overlap picture one takes into account all islands and their local stochastic layers around the separatrices – the widths of these can be calculated using the Melnikov-Arnold integral (appendix B, Gerhard 1985). Even then, the last KAM tori between the resonances can break up before the overlap prediction because of the interaction between the resonances that secular perturbation theory overlooks. In the example case, the local stochastic layers associated with the islands were always very narrow; also, the islands computed separately look like those computed together, so there is no shape distortion because of the combined computation. Therefore, the main reason for the difference with Chirikov estimate is precisely the interplay between the resonances.

The natural home space of phenomena associated with resonances is, of course, frequency space. For example, resonant trapping can be seen as being caused by $\dot{\boldsymbol{J}}$ from (5.11) acting long enough in one direction near a resonance to form an island; further away the direction change is more rapid and this leads eventually to chaos between two competing resonances.

However, it is not easy to think of a quantitative 'frequency criterion' for the transition to chaos, because the motion between the resonances is complex.

To consider a slightly more complicated example, let us perturb the isochrone Hamiltonian, $H_0 = H_{\rm I}$ (and use the isochrone action-angles). Now the resonance ratio γ alone determines what value the action J_φ has at the resonance point:

$$J_\varphi = \sqrt{\frac{bk}{\gamma(1-\gamma)}}(2\gamma - 1), \tag{5.19}$$

where b, k are the isochrone potential parameters. Also, $\omega_r = (-2H_0)^{3/2}/k$ depends only on the value of the Hamiltonian. The nonlinearity parameter G is now given by

$$G(H_0, \boldsymbol{k}) = \omega_r \, | \, k_1 k_2 \, | \, \frac{1 - J_\varphi/C}{C}, \tag{5.20}$$

where $C \equiv \sqrt{4bk + J_\varphi^2}$.

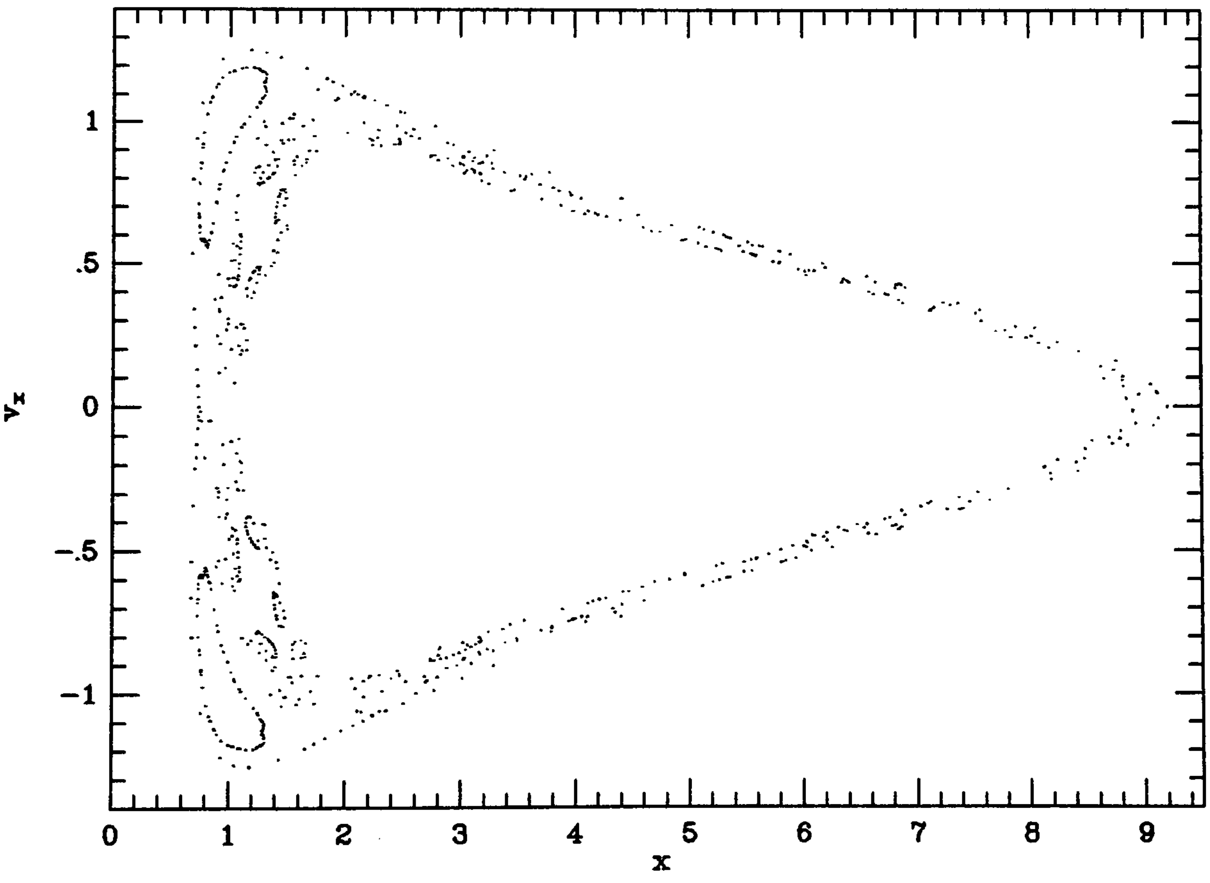

Figure 5.6: A two-resonance example with the isochrone Hamiltonian.

In Fig. 5.6, we show $(x, \dot{x})$ at $H = -0.3$ and $k = 3, b = 0.5$. The islands of 3:4 and 4:5 resonances are created by the perturbing terms $H_{3,-4} = 0.0001$ and $H_{4,-5} = 0.00009$; global chaos is first seen slightly under these values. Note the four secondary islands in the 4:5-region: using only two small resonance terms locally produces the main characteristics of a strongly non-integrable Hamiltonian! The overlap estimates from (5.14) are $H_{3,-4} = 0.00011$ and $H_{4,-5} = 0.0001$, so again the chaos starts somewhat earlier than estimated. In general, in our two-resonance systems the transition to global chaos seems to occur at island sizes roughly 70–90% of those predicted. Chirikov (1979) and others (Channell 1978, Lichtenberg & Lieberman 1983) have obtained similar results with various systems. Thus the overlap criterion serves quite well as a practical estimate, and certainly it always gives an upper limit for the perturbation. Besides, it is in fact the only existing proper criterion, in the sense that one can really predict (analytically) the critical values of the parameters in a straightforward manner.

Examples

Let us now explore the route to global chaos in two ways. First, we create global stochasticity by increasing the angular speed of the rotating frame for the logarithmic potential (2.5). In Fig. 5.1 and Fig. 5.2, we already saw chaotic regions, islands and major invariant curves. To give an example of a case where all islands have finally overlapped and produced global chaos, we set the system of Fig. 4.4 ($q = 0.6$ and $H = -0.3215$) rotating at the speed $\Omega = 0.28$. This is shown in Fig. 5.7: all families have mixed, x_2 has disappeared, and secondary islands of the banana 1:2 islands and x_4 can be seen.

The route to global chaos can be seen as an effect of progressive resonance overlapping. With other potential parameters, similar behaviour can be created with increasing rotation speed (sometimes quite complex and with very interesting shapes of invariant curves, but to show that properly we would have to create a video animation to accompany this thesis). This view differs somewhat from that of Martinet & Udry (1990). They argue that the instability of the x_3-family (the unstable companion of x_2) is the main cause in triggering

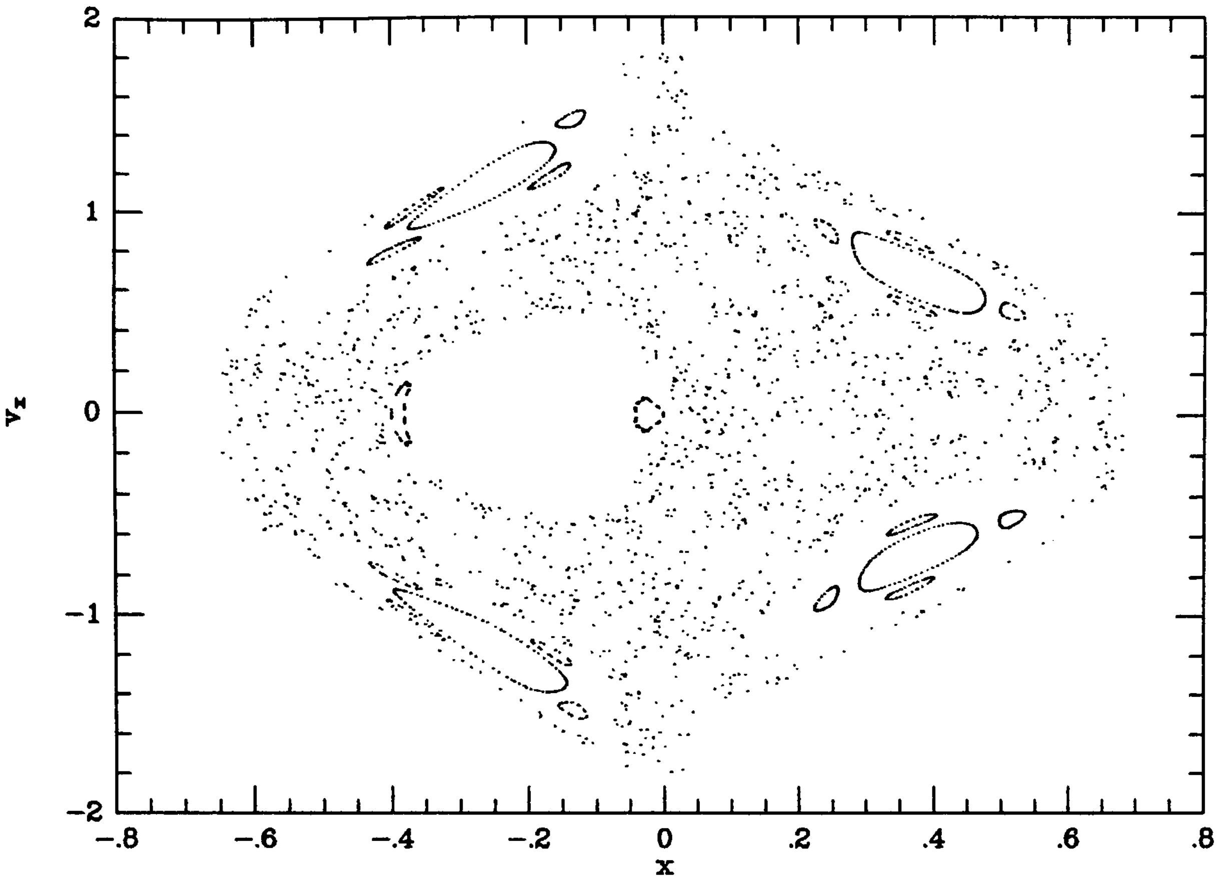

Figure 5.7: $(x, \dot{x})$ at $q = 0.6$, $H = -0.3215$ and $\Omega = 0.28$.

chaos, and that chaos is reduced when the x_2, x_3-pair vanishes. But there is nothing very special about the role of the x_3-family. It lies at the unstable point of a separatrix joining the loop pair. Accordingly, there is a local stochastic layer for the loop pair, and the interaction with a nearby chain of islands will produce a merged chaotic region (the width of the local layer is discussed in Gerhard 1985). This is just one resonance interaction, not a major source: other resonances away from the centre can overlap without knowing what is going on in the centre, although often the overlapping process spreads from the centre outwards. The rotation brings these resonances together – in the nonrotating case, islands stubbornly refuse to merge, even though there can be plenty of local chaotic layers (see Fig. 4.4). Also, in our example chaos grows steadily as the rotation speed is increased, regardless of the existence of the x_2, x_3-families. In other cases, the chaotic regions may finally become smaller and global chaos will never be reached, but the compression of the chaotic zone does not always seem to be connected with the disappearance of the x_2, x_3-families: rather, it is because some regular families start occupying larger regions of phase space. For example, in the earlier example of Fig. 5.2, the chaotic zone is kept under control because of the strong region of regularity starting right from the 7:5-islands. Inside, chaos reigns, but it does not stand a chance of breaking through the wall of KAM-curves, seemingly because these curves do not want to produce a series of islands that could then merge.

As an another way of creating increasing stochasticity, we employ a specific stochasticity parameter by adding an irregularity term to (2.5) (Binney & Tremaine 1987, chapter 3.5):

$$\Phi_{\mathrm{irr}} = \frac{1}{2}\ln\left[x^2 + \frac{y^2}{q^2} + R_{\mathrm{c}}^2 + \left(\frac{R}{R_{\mathrm{e}}}\right)(x^2 - y^2)\right], \tag{5.21}$$

where $R \equiv \sqrt{x^2 + y^2}$, and R_{e} is the stochasticity parameter.

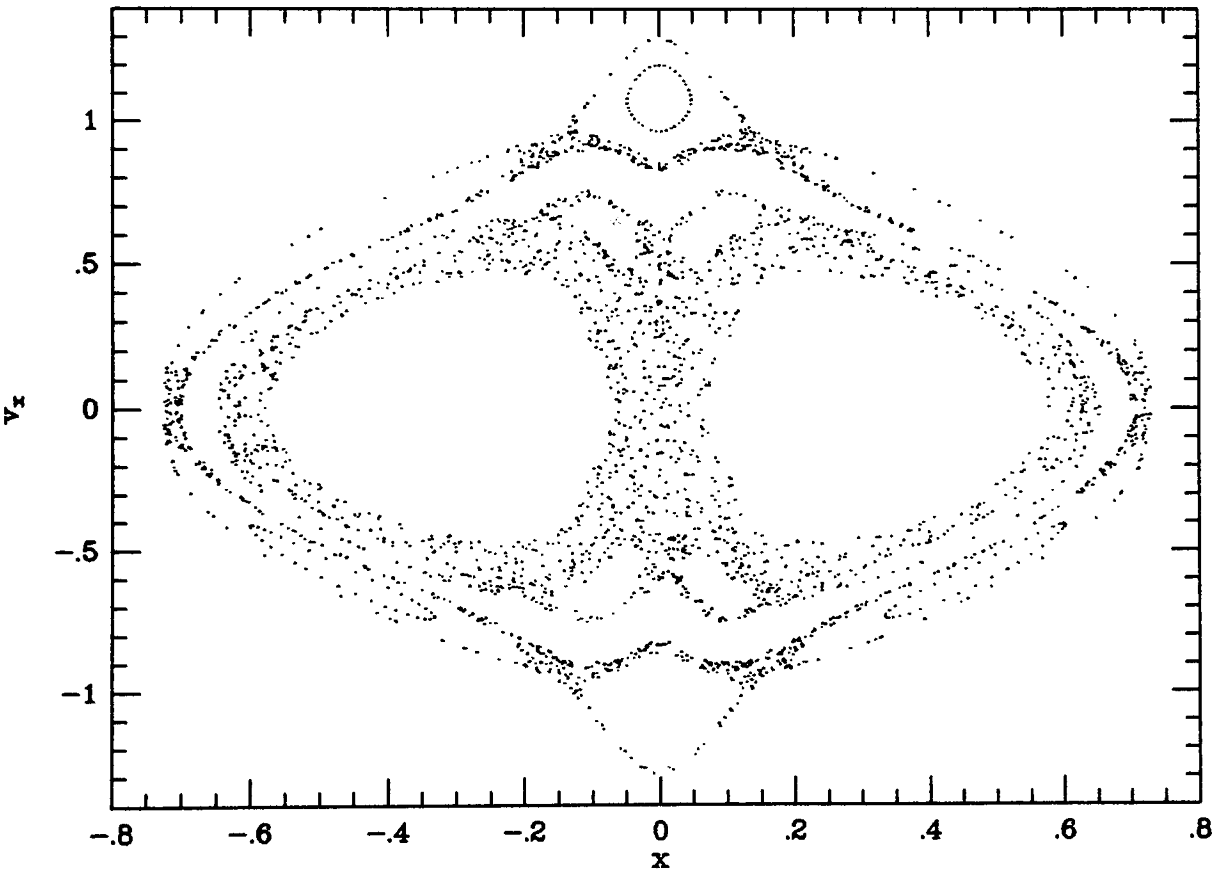

Figure 5.8: $(x, \dot{x})$ of (5.21) at $R_{\mathrm{e}} = 5$ and $q = 0.8, H = -0.20$.

Let us now see what happens when the irregularity term in (5.21) is increased. This is a more straightforward and 'smooth' way of creating chaos than using the rotating potential: the larger the irregularity term, the larger the regions of stochasticity. In the process, we construct invariant tori and estimate the widths of resonant islands: it is then easy to check whether they are consistent with the overlap concept. A typical example is shown in Fig. 5.8 and Fig. 5.9. In Fig. 5.8, the $(x, \dot{x})$-section is shown at $R_e = 5$ and $q = 0.8, H = -0.20$ (as with the fish islands example of chapter 4). The dominant islands are, towards the centre, 3:2 and 4:3; these have not yet overlapped, whereas the centre loops have mixed with 4:3. For clarity, we have left intentionally a gap between the two regions (it is occupied by some invariant curves and very small islands).

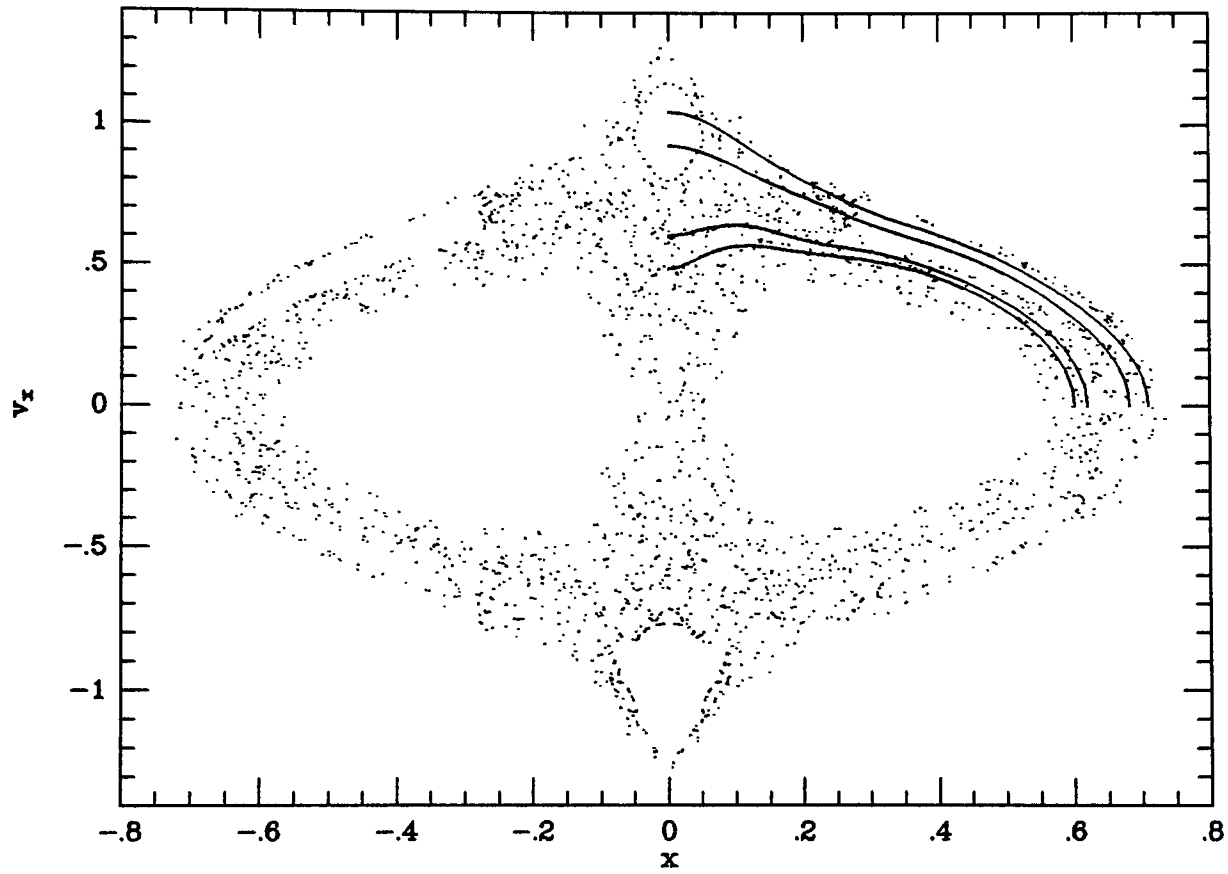

Figure 5.9: As above, but $R_e = 3.5$. The solid lines correspond to fitted tori.

At $R_e = 3.5$ in Fig. 5.9, 4:3 and 3:2 have just 'recently' mixed. In fact, chaos has spread so quickly that the 4:3 islands have vanished. It is interesting to see whether a perturbation of an integrable Hamiltonian, to be constructed, still 'remembers' the vanished island – it should, because the barrier between 3:2 and 4:3 was the last one to be broken, to allow global spread of stochasticity. The perturbation coefficients H_n at the resonant tori should not have changed much from the slightly 'earlier' transition to chaos (and numerical checking shows that they have not). As solid lines, we show the curves corresponding to two pairs of constructed tori, each pair bounding a resonance. Towards the centre, these tori are $\boldsymbol{J} = (0.29, 0.20), \boldsymbol{\omega} = (1.476, 2.248)$ and $\boldsymbol{J} = (0.26, 0.22), \boldsymbol{\omega} = (1.513, 2.232)$ for 3:2, and $\boldsymbol{J} = (0.195, 0.265), \boldsymbol{\omega} = (1.607, 2.197)$ and $\boldsymbol{J} = (0.175, 0.28), \boldsymbol{\omega} = (1.642, 2.182)$ for 4:3. (This is another example of the least-squares torus-fitting procedure and the angle recovery method working fine within chaotic/resonant regions.)

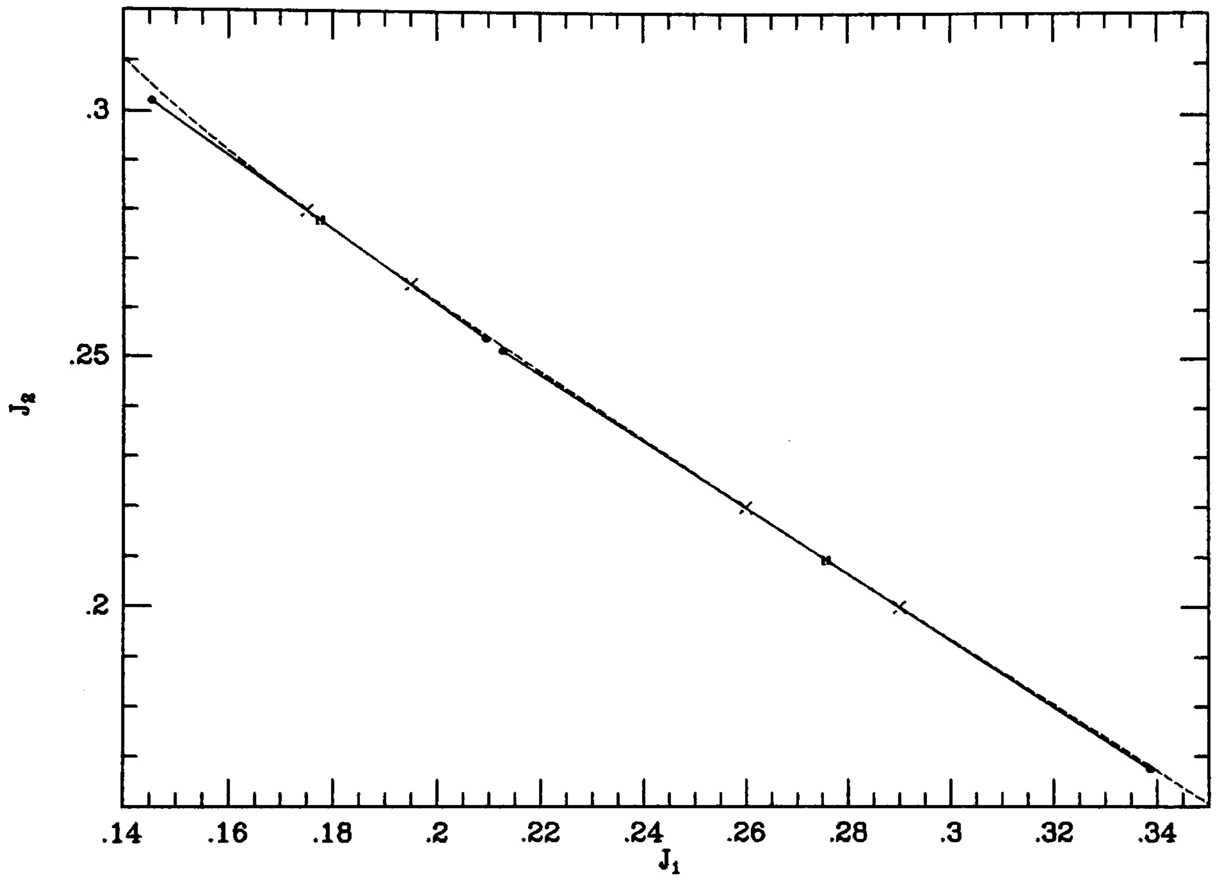

Figure 5.10: The calculated island widths of Fig. 5.9 in action space. The centre- and endpoints of the islands are marked with dots, and the interpolating tori with crosses. The dashed line shows (schematically) a curve of constant H_0.

Interpolation between these tori places the resonant tori at $\boldsymbol{J} = (0.2757, 0.2095)$ and $\boldsymbol{J} = (0.1775, 0.2781)$. Proceeding as in chapter 4, we obtain the approximate island widths (in units of $\boldsymbol{k} \cdot \boldsymbol{J}$, where $\boldsymbol{k}$ is the resonance index, see above); because the islands are small, we can set $\beta = 0$ and from (4.21) the half-width is then approximately $2\sqrt{|F/G|}$. This yields the half-widths 0.021 and 0.008 for, respectively, 3:2 and 4:3. The corresponding lines in action space thus almost overlap, as shown in Fig. 5.10, where the 4:3 resonance region is to the left of 3:2. (The slightly concave curve of constant H_0, defined by the constructed tori, would run through the actions of the interpolating tori marked with crosses.) One does not actually have to see the 4:3 island to conclude that it must have been there, and that the interaction between 3:2 and 4:3 has played a part in the transition to chaos. Other test cases have produced similar results; thus the scheme of defining H_0 by torus construction is consistent with the (somewhat vague) overlap concept.

Chapter 6

Conclusions, discussion and some speculation

Well, here we are. I hope you have found the journey interesting. Let us now briefly sum up what it was all about and fit the concepts in the big picture. In a nutshell or two:

Goal:

Construct an integrable Hamiltonian H_0 close to a given one H.

Method:

$$\begin{pmatrix}\text{invariant tori in the phase space}\\ \text{of a known integrable } H_\mathrm{T}\end{pmatrix} \xrightarrow{\text{map (non-perturbatively)}} \begin{pmatrix}\text{phase space}\\ \text{of target } H\end{pmatrix}$$

$\Rightarrow$ The product: $\left(\text{action-angles of } H_0 \leftrightarrow \text{phase-space point } \boldsymbol{w}\right)$

That was the methodological bottomline. In the process, lots of interesting things were discovered:

- Tori can be fitted to orbits of different families by using different H_Ts. In our examples, two completely independent mappings matched beautifully at the interface between box- and loop-orbits of barred potentials.
- Generally, the canonical transformation consists of (i) a point transformation designed to map orbits in H_T into crude representations of the desired orbit in H, and (ii) a generating

function $S(\boldsymbol{\theta}, \boldsymbol{J}')$. Item (i) is needed e.g. when dealing with box orbits or with loop orbits that are so narrow and so elliptical that a circle cannot be inscribed within them.

- The torus construction procedure can be employed in large chaotic regions of phase space and with Hamiltonians in rotating frames of reference as well.
- If needed, tori that coincide with existing invariant tori of H can be constructed using the orbit integration scheme of Warnock (1991), with the optimized parameters of $H_{\rm T}$ and point transformations given by the least-squares fitting method.
- Once the canonical transformation is known, the angle variables and frequencies of fitted tori can be recovered robustly from a set of orbit integrations. This method produces credible angle variables and frequencies not only for tori that are good fits to real orbits, but also for ones that slice through resonant islands or chaotic regions.
- From the constructed tori, one can create an integrable H_0 by using the fitted tori as interpolation points, and the chosen interpolation method then defines H_0. H_0 is usually not derivable from a potential.
- Minor-orbit families can be interpreted as families of trapped orbits created by perturbing H_0 at a resonance. This perturbation can be accurately described by a model which is a modification of the pendulum analysis of standard secular perturbation theory.
- With the aid of a point transformation, tori for minor-orbit families can be constructed directly, using the same $H_{\rm T}$s as for major-orbit families.

One could say that we have demonstrated yet again the usefulness of the old geometric principle of choosing the coordinates best representing the geometry of the problem. The action-angle coordinates and geometric phase-space structures are, of course, a sublime manifestation of this principle. Also, in our case, mappings between e.g. the Cartesian or polar coordinates of the toy tori and the elliptic coordinates used for the target tori have proved to be the keys to success. In a similar vein, Ollongren (1967) used canonical transformations to map the orbits of harmonic oscillators into typical meridional-plane orbits, and Kent & de Zeeuw (1991) fitted separable potentials to individual orbits.

The existence of tailor-made integrable Hamiltonians, which do not derive from potentials and will therefore probably prove difficult to construct by other means, opens up a number of promising lines of enquiry into astrophysically interesting cases in which minor orbit families play important rôles. The possibility of using perturbation theory for wider classes of systems than those hitherto available can be applied to numerous problems, for example in studying fine structure in early-type stellar systems or other phenomena associated with resonances. Also, a general consequence of the possibility of creating a global H_0 close to a given Hamiltonian is that all of the action space can be employed in problems involving functions expressed in actions, e.g. building self-consistent galactic models using distribution functions of Jeans type.

In various dynamical problems of a general nature, a detailed study concentrating on the neighbourhoods of resonances is of great interest, especially in studying the transition to chaos in near-integrable systems. We have confirmed that Chirikov's overlap criterion is a practical rule-of-thumb; however, one would like to get a deeper understanding of the transition, more along the dynamical lines than along the intuitive graphical ones of the overlap concept. In this respect, it would be interesting to see if one could gain more insight by using constructed H_0s with small, 'controlled' perturbations. For example, why exactly does rotation or an irregularity term destroy the regularity of the logarithmic potential (and why is it so regular in the first place – even when very elongated, producing a large number of sizeable resonant islands)?

When it comes to the vast realm of general dynamics, the general applicability of the presented methods is still an open question. Without doubt, there are lots of 'pathological' systems out of our reach – excluding those, one would like to know whether or not we have just been lucky so far with our successful case studies. Therefore, a natural step in the future is to experiment with the techniques using systems that are more general than (or different from) those studied here. One cannot say "if you've seen one Hamiltonian, you've seen them all". The biggest theoretical limitations are the coordinate systems (point transformations) and toy Hamiltonians available. Even if torus construction is possible in theory, the practical requirement of a rapidly converging series for the generating function S must be met. For example, the Hénon-Heiles Hamiltonian (Hénon & Heiles 1964, and any textbook in modern dynamics) possesses two interesting-looking major-orbit families of box- and loop-type (they are plotted in both configuration space and surfaces of section in Martens & Ezra 1987). Can we create suitable point transformations (appendix D) for handling the system?

It would be interesting to see how the methods presented here could be applied to problems such as predicting stability of particle orbits over very long time intervals (e.g. beams in storage rings, plasma and Solar system physics) in systems with more than two degrees of freedom. Such systems exhibit orbital (or Arnold) diffusion in phase space (e.g. Lichtenberg & Lieberman 1983), and conventional numerical integrations are not always completely reliable in determining long-time stability: there are many examples of orbits apparently behaving well for a long time before displaying instability. Also, as discussed in chapter 3, the $\boldsymbol{\theta}$-space in three degrees of freedom will be much larger than that in two dimensions. Therefore, general nonautonomous two-dimensional ('$2\frac{1}{2}$' degrees of freedom) systems, or those with three degrees of freedom, are even more challenging from both theoretical and computational points of view.

Appendix A

Action-angle variables for the isochrone Hamiltonian

For completeness, the equations for the fully three-dimensional case are given here (because as such I have not seen them elsewhere); the calculational details are mostly from Saha (1989) and Goldstein (1980) – the full derivation is presented here to serve as an example of the procedure of solving the Hamilton-Jacobi equation.

In spherical coordinates, the full Hamiltonian reads

$$H(r,\vartheta,\varphi,p_r,p_\vartheta,p_\varphi) = \frac{1}{2}(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\varphi^2}{r^2\sin^2\vartheta}) + \Phi(r), \tag{A.1}$$

where the isochrone potential is

$$\Phi(r) = -\frac{k}{b+\sqrt{b^2+r^2}}. \tag{A.2}$$

In order to obtain expressions establishing the transformations between $(\boldsymbol{J},\boldsymbol{\theta})$ and the momenta and coordinates $(\boldsymbol{p},\boldsymbol{x})$ of (A.1), the generating function $S(\boldsymbol{J},\boldsymbol{x})$ of the Hamilton-Jacobi equation must be solved for; this gives the necessary relations $\boldsymbol{\theta} = \partial S/\partial \boldsymbol{J}$ (and $\boldsymbol{p} = \partial S/\partial \boldsymbol{x}$). The Hamilton-Jacobi equation is now

$$\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \vartheta}\right)^2 + \frac{1}{r^2\sin^2\vartheta}\left(\frac{\partial S}{\partial \varphi}\right)^2 + 2\,\Phi(r) = 2\,H(\boldsymbol{J}). \tag{A.3}$$

Next we write S in the separated form $S(\boldsymbol{J},\boldsymbol{x}) = S_r(\boldsymbol{J},r) + S_\vartheta(\boldsymbol{J},\vartheta) + S_\varphi(\boldsymbol{J},\varphi)$; this gives

$$p_\varphi^2 = \left(\frac{\partial S}{\partial \varphi}\right)^2 = L_z^2, \tag{A.4}$$

$$p_\vartheta^2 = \left(\frac{\partial S}{\partial \vartheta}\right)^2 = L^2 - \frac{L_z^2}{\sin^2\vartheta}, \tag{A.5}$$

$$p_r^2 = \left(\frac{\partial S}{\partial r}\right)^2 = 2H - \frac{L^2}{r^2} - 2\Phi(r), \tag{A.6}$$

where the separation constants L and L_z are, respectively, the magnitude and the z-component of the angular momentum vector, as can directly be seen from (A.4) and (A.5). The azimuthal action J_φ is trivially

$$J_\varphi = \frac{1}{2\pi}\oint p_\varphi(\varphi)\,\mathrm{d}\varphi = L_z. \tag{A.7}$$

From (A.5) we have

$$S_\vartheta = \int_{(\vartheta_{\min})}^{(\vartheta)} p_\vartheta(\vartheta)\,\mathrm{d}\vartheta, \tag{A.8}$$

and $\sin\vartheta_{min} = L_z/L$. S_ϑ could be represented as an indefinite integral, since S admits an arbitrary additional constant; here this constant is fixed by choosing ϑ_{min} to be the 'natural' lower integration limit. The latitudinal action J_ϑ is

$$J_\vartheta = \frac{2}{\pi}S_\vartheta\Big|_{\vartheta_{min}}^{\frac{\pi}{2}} = L - L_z, \tag{A.9}$$

(e.g. Goldstein 1980), so

$$L = J_\vartheta + J_\varphi. \tag{A.10}$$

Next we want to evaluate from (A.6)

$$S_r = \int_{r_{min}}^{r} p_r(r)\,\mathrm{d}r. \tag{A.11}$$

This is best calculated using the substitution

$$\sqrt{b^2+r^2} - b \equiv u = a(1 - e\cos\psi), \tag{A.12}$$

where

$$a = -\frac{k}{2H} - b; \quad e^2 = 1 - \frac{L^2(a+b)}{ka^2}, \tag{A.13}$$

and $\psi = 0$ corresponds to r_{min} and $\psi = \pi$ to r_{max}. For bounded orbits, $a \geq 0$ and $0 \leq e^2 < 1$.

Equation (A.11) reads now

$$\begin{aligned} S_r &= \frac{1}{2} \int \left(\frac{1}{u} + \frac{1}{u+2b} \right) (2H\, u(u+2b) - L^2 + 2ku)^{\frac{1}{2}}\, \mathrm{d}u \\ &= \frac{1}{2} \int \left(\frac{1}{u} + \frac{1}{u+2b} \right) \left[\frac{k}{a+b} (ae \sin\psi)^2 \right]^{\frac{1}{2}} ae \sin\psi\, \mathrm{d}\psi, \end{aligned} \tag{A.14}$$

where the integrand is expanded in partial fractions for evaluation. Using the integral

$$\frac{1}{\pi} \int_0^\pi \frac{\sin^2 x\, \mathrm{d}x}{p + q\cos x} = q^{-2}(p - \sqrt{p^2 - q^2}), \tag{A.15}$$

we have

$$J_r = \frac{1}{\pi} S_r \Big|_{r_{\text{min}}}^{r_{\text{max}}} = \sqrt{k(a+b)} - \frac{1}{2}(L + K), \tag{A.16}$$

where

$$K = \sqrt{4bk + L^2}. \tag{A.17}$$

Expressed in actions, H is thus, from (A.16) and (A.13),

$$H(\boldsymbol{J}) = -2k^2(2J_r + L + K)^{-2}. \tag{A.18}$$

In order to obtain the relations $\boldsymbol{\theta} = \partial S / \partial \boldsymbol{J}$ the frequencies $\boldsymbol{\omega} = \partial H / \partial \boldsymbol{J}$ are useful. From (A.10) and (A.18) it follows that $\omega_\varphi = \omega_\vartheta \equiv \omega_r = \partial H / \partial L$. Now

$$\theta_r = \frac{\partial S}{\partial J_r} = \frac{\partial S_r}{\partial H} \omega_r. \tag{A.19}$$

Differentiating (A.14) under the integral sign before integrating and using the variable ψ, we obtain the (generalized) Kepler equation

$$\theta_r = \psi - \frac{ae}{a+b} \sin\psi. \tag{A.20}$$

The angle variables $\theta_\tau, \tau = \vartheta, \varphi$, are

$$\theta_\tau = \frac{\partial S}{\partial J_\tau} = \sum_{\tau'} \frac{\partial S_{\tau'}}{\partial J_\tau} + \frac{\partial S_r}{\partial J_\tau} = \sum_{\tau'} \frac{\partial S_{\tau'}}{\partial J_\tau} + \frac{\partial S_r}{\partial H}\omega_\tau - \Lambda, \qquad (A.21)$$

where

$$\Lambda = -\left(\frac{\partial S_r}{\partial J_\tau}\right)_H = -\left(\frac{\partial S_r}{\partial L}\right)_H. \qquad (A.22)$$

Using (A.19), this can be written as

$$\theta_\tau = \sum_{\tau'} \frac{\partial S_{\tau'}}{\partial J_\tau} + \gamma\,\theta_r - \Lambda, \qquad (A.23)$$

where the ratio γ of the frequencies ω_τ, ω_r is

$$\gamma = \frac{1}{2}(1 + \frac{L}{K}). \qquad (A.24)$$

For φ,

$$\sum_{\tau'} \frac{\partial S_{\tau'}}{\partial J_\varphi} = \varphi + \chi + \left(\frac{\partial S_\vartheta}{\partial J_\varphi}\right)_L, \qquad (A.25)$$

where

$$\chi \equiv \frac{\partial S_\vartheta}{\partial L}. \qquad (A.26)$$

The quantity χ can be calculated from (A.8) differentiating under the integral sign with the substitution $x = l\,\cos\vartheta$ (note the choice of integration limits in (A.8), corresponding to an additional constant $\frac{\pi}{2}$):

$$\chi = \int \frac{l\,\sin\vartheta\,\mathrm{d}\vartheta}{\sqrt{1 - l^2\cos^2\vartheta}} = \arccos(l\,\cos\vartheta), \qquad (A.27)$$

where the auxiliary quantity l is

$$l \equiv \left(\sqrt{1 - \frac{L_z^2}{L^2}}\right)^{-1} = \cos^{-1}\vartheta_{\mathrm{min}}. \qquad (A.28)$$

Similarly we find, using the substitution $x = l'\cot\vartheta; l' \equiv lL_z/L$ in integration,

$$\left(\frac{\partial S_\vartheta}{\partial J_\varphi}\right)_L = \int \frac{l'\,\mathrm{d}\vartheta}{\sin^2\vartheta\sqrt{1 - l'^2\cot^2\vartheta}} = \arccos(l'\cot\vartheta). \qquad (A.29)$$

For motion in the xy-plane, (A.27) and (A.29) vanish.

For ϑ,

$$\sum_{\tau'} \frac{\partial S_{\tau'}}{\partial J_\vartheta} = \chi. \tag{A.30}$$

To complete (A.23), we need the quantity Λ (A.22). Again, differentiating (A.14) under the integral sign and using the integral

$$\int (q + \cos x)^{-1}\, \mathrm{d}x = \frac{2}{\sqrt{q^2-1}} \arctan\left(\sqrt{\frac{q+1}{q-1}} \tan\frac{x}{2}\right), \tag{A.31}$$

we obtain

$$\Lambda(\psi) = \arctan\left(\sqrt{\frac{1+e}{1-e}} \tan\frac{\psi}{2}\right) + \frac{L}{K} \arctan\left(\sqrt{\frac{a(1+e)+2b}{a(1-e)+2b}} \tan\frac{\psi}{2}\right). \tag{A.32}$$

With $r^2 = u(u+2b)$ and $\cos\psi = -\frac{1}{ae}(\frac{k}{2H} + \sqrt{b^2+r^2})$ from (A.12), the above equations establish the relation $(\boldsymbol{J}, \boldsymbol{\theta}) \leftrightarrow (\boldsymbol{p}, \boldsymbol{x})$; the radial momentum (A.6) can also be written as (cf. (A.14))

$$p_r = \frac{\sqrt{-2H}}{r} ae \sin\psi. \tag{A.33}$$

Appendix B

Useful formulae

For convenience, we give here a short collection of useful formulae, comments, and other odds and ends that do not quite fit in the main text and would look clumsy in footnotes. Because this material can essentially be found elsewhere, detailed discussion is omitted.

Stäckel potentials and elliptic coordinates

Here are listed a few basic properties of confocal elliptic coordinates, and Stäckel potentials as functions of these; for details, see de Zeeuw (1985) and Binney & Tremaine (1987). The transformation from the elliptic coordinates (u, v) to Cartesian ones reads

$$x = \Delta \sinh u \, \cos v, \quad y = \Delta \cosh u \, \sin v, \tag{B.1}$$

where Δ is the focal length of the coordinate system, i.e., the foci of the ellipses u =const. and the hyperbolae v =const. are at $x = 0, y = \pm\Delta$. The inverse transformation is

$$2\sinh^2 u = (r/\Delta)^2 - 1 + \sqrt{C(x,y)}, \quad 2\sin^2 v = (r/\Delta)^2 + 1 - \sqrt{C(x,y)}, \tag{B.2}$$

where $r^2 = x^2 + y^2$ and

$$C(x,y) \equiv [(r/\Delta)^2 - 1]^2 + 4(x/\Delta)^2 = [(r/\Delta)^2 + 1]^2 - 4(y/\Delta)^2. \tag{B.3}$$

The correct sign for u (if negative u are employed, as for box orbits) and quadrant for v are obvious from (x, y).

Stäckel potentials are of the form

$$\Phi(u, v) = \frac{U(u) - V(v)}{\cosh^2 u - \sin^2 v}, \tag{B.4}$$

where we have used

$$\begin{aligned} U(u) &= -w \sinh u \, \arctan(\Delta \sinh u \, / a), \\ V(v) &= w \cos v \, \operatorname{artanh}(\Delta \cos v \, / a), \end{aligned} \tag{B.5}$$

where w, a are parameters. Stäckel potentials are Hamilton-Jacobi separable, and have thus as many isolating integrals as there are dimensions: for two dimensions, one is the energy E and the other we call I_2. When $I_2 > 0$, the orbit is a loop orbit; when $I_2 < 0$, it is a box one. The conjugate momenta to (u, v) are

$$\begin{aligned} p_u &= \pm\Delta\sqrt{2[E \sinh^2 u - I_2 - U(u)]}, \\ p_v &= \pm\Delta\sqrt{2[E \cos^2 v + I_2 + V(v)]}. \end{aligned} \tag{B.6}$$

These are related to the Cartesian momenta by

$$\begin{pmatrix} p_u \\ p_v \end{pmatrix} = \Delta \begin{pmatrix} \cosh u \, \cos v & \sinh u \, \sin v \\ -\sinh u \, \sin v & \cosh u \, \cos v \end{pmatrix} \begin{pmatrix} p_x \\ p_y \end{pmatrix}. \tag{B.7}$$

Surfaces of section and orbit integration

The surfaces of section $(x, \dot{x})$ or $(y, \dot{y})$ of the constructed tori are easily computed. In the case of the harmonic oscillator mapping, one just plots the values by varying one angle θ_i while holding the other $\theta_j = 0$. In the isochrone mapping, one can solve numerically for the angle θ_φ while varying θ_r and holding φ constant.

For numerical orbit integration in a given Hamiltonian, we used symplectic integrators (see e.g. Candy & Rozmus 1991, Forest & Ruth 1990); the second-order 'leapfrog' is usually sufficient, but a fourth-order one is easy to use as well. For a leapfrog-type explicit symplectic integrator to work, the Hamiltonian must be separable in the form $H(\boldsymbol{p}, \boldsymbol{x}) = T(\boldsymbol{p}) + V(\boldsymbol{x})$. Therefore orbit integration in a rotating frame of reference is easiest to do by integrating the orbit in inertial space and then transforming it into the rotating frame. For this, the 'static' leapfrog can easily be modified to integrate a time-dependent case $H(\boldsymbol{p}, \boldsymbol{x}, t) = T(\boldsymbol{p}) + V(\boldsymbol{x}, t)$ by extending the phase-space to include t as a coordinate.

Finally, we note (just in case you have not heard of this) that a fast and accurate way of obtaining surfaces of section from numerically integrated orbits is described by Hénon (1982). Instead of solving, by root-finding in a time interval, for the point for which a coordinate (or other function) is zero on the desired surface, one can define a new set of functions for which the coordinate is an independent variable. One knows the magnitude of a step that brings that variable to zero: using one Runge-Kutta-iteration step, the corresponding time and the other quantities are accurately obtained.

Melnikov-Arnold integral

The Melnikov-Arnold (MA) integral can be employed in estimating the width of the local stochastic layer associated with a separatrix in a near-integrable Hamiltonian. We give here a form (following Lichtenberg & Lieberman 1983 and Chirikov 1979) that can be used with perturbations of the Fourier-type form employed in chapters 4 and 5. For simplicity, the Fourier coefficients are taken to be constants (in the region of interest) as in chapter 5.

As explained in chapter 4, averaging over the fast angles yields a Hamiltonian of the form

$$H(J, p, \varphi) = H_0(J) + \frac{1}{2} G p^2 - F \cos\varphi, \tag{B.8}$$

where J is the constant of motion J_ξ of (4.13), and the corresponding fast angle θ is taken to increase linearly; p and φ describe the integrable pendulum motion. Here $F = 2\,|\,H_{r,-s}\,|$, where $H_{r,-s}$ is the Fourier coefficient of the resonance $r : s$ in question, and G (taken to be positive) is as defined in chapter 4 with (4.15). To obtain a more accurate picture, we acknowledge the presence of other terms than this: resonances between the fast angle and the slow motion of φ lead to a layer of stochasticity and secondary islands. (The higher-order island formation can be studied *ad infinitum* by renormalizing the system of the current order such that it resembles that of the previous order; this is explained in e.g. Lichtenberg & Lieberman 1983).

Briefly reintroducing the previously dropped terms, we have

$$H(J,p,\varphi) = H_0(J) + \frac{1}{2}Gp^2 - F\cos\varphi + \sum_{k,l} \Lambda_{k,l} \cos\left(\frac{k}{r}\varphi - \frac{l}{r}\theta\right), \tag{B.9}$$

where the $\Lambda_{k,l}$-coefficient corresponding to $2H_{i,-j}$ is $\Lambda_{i,jr-is}$. Our purpose is to calculate how much J maximally varies from the idealized constant value: this leads to the stochastic layer. Including only the leading $\Lambda_{k,l}$-term and using the same principle as with creating a so-called perturbed twist mapping from a perturbed Hamiltonian, we arrive at the form describing the half-width of the layer in terms of the action J (see Lichtenberg & Lieberman 1983; they have a minor error in this final form):

$$\Delta J = \frac{\Lambda_{k,l}}{r\omega_0}\mathcal{A}_{2k/r}(l\beta), \tag{B.10}$$

where $\omega_0 = \sqrt{FG}$ (the frequency for small librations of the pendulum),

$$\beta = \frac{\omega}{r\omega_0}, \tag{B.11}$$

ω the frequency associated with the fast variable θ, and $\mathcal{A}_m(x)$ is the MA integral

$$\mathcal{A}_m(x) = \int_{-\infty}^{\infty} \cos\left[\frac{m}{2}\phi(s) - xs\right] \mathrm{d}s, \tag{B.12}$$

where $\phi(s) = 4\arctan(e^s) - \pi$, the unperturbed separatrix motion for φ when $s = \omega_0 t$. The MA integral is actually improper, but it can be divided into an oscillating part that can be neglected, and a shift; this shift is taken to be ΔJ. For m integer, it can be evaluated analytically. Usually $m \ll x$, and then one can use the asymptotic expansion

$$\mathcal{A}_m(x) = \frac{1}{(m-1)!} 4\pi(2x)^{m-1} \exp(-x\pi/2). \tag{B.13}$$

Thus the layer is exponentially small in the frequency ratio $l\omega/r\omega_0$. For $x < 0$ the exponential curve is even sharper: $\mathcal{A}_m(x) = (-1)^m \mathcal{A}_m(-x)\exp(x\pi)$.

Taking into account the shift $\omega\Delta J$ in $H_0(J)$ caused by ΔJ, we can estimate the maximum 'kick' to p, producing the new $p'_{\max}$ at $\varphi = 0$, i.e. the island half-width with the stochastic layer added:

$$p'_{\max} = \sqrt{\frac{2(2F + \omega\Delta J)}{G}} = 2\sqrt{\frac{F}{G}}\sqrt{1 + \frac{\omega\Delta J}{2F}} \simeq p_{\max}\left(1 + \frac{\omega\Delta J}{4F}\right), \tag{B.14}$$

where $p_{\max}$ is the old island size.

For example, the widths of the stochastic layers for the 3:5 and 3:4 islands of (5.16) in chapter 5 can be directly computed from (B.10). Even with the largest possible islands before actually overlapping, they will never be larger than about one hundredth part of the island widths, and the layers shrink exponentially as $F \to 0$.

One might claim that in some places there is enough handwaving to blow the notes off the desk; I leave the justification to Chirikov and Lichtenberg & Lieberman. From comparisons with lots of numerical integrations of the complete equations of motion, the above seems to give a decent estimate for the size of the local stochastic layer. The twist mapping constructed with ΔJ, often call the separatrix mapping as well, has often been used in studying chaotic behaviour. It is interesting to note that important phenomena local to an island (chaos and higher-order islands) are caused by frequency terms other than the one responsible for the main island. The global or 'inter-island' phenomena (such as the transition to global stochasticity) caused by many terms have so far not been modelled as well as those describing the local phenomena.

Appendix C

The modified pendulum

For simplicity, the canonical momentum is here denoted by p; in the main text it is δJ_ψ. The pendulum is described by (4.19):

$$H \equiv \frac{1}{2}(G + \beta \cos m\psi)p^2 + (\alpha p + F)\cos m\psi. \tag{C.1}$$

We denote the numerical value of H by h. Only the cases where $|\beta| < |G|$ and $F + \alpha p + \frac{1}{2}\beta p^2$ has the same sign as F are of interest; $FG < 0$ as in the main text. From the Hamiltonian equations of motion we find that in rotation $(0 \leq m\psi \leq 2\pi)$ $\dot{p} = 0$ at $m\psi = n\pi$ and thus p varies between

$$\begin{aligned} p_{\max}^{\pm} &= \frac{1}{G+\beta}[-\alpha \pm \sqrt{\alpha^2 - 2(G+\beta)(F-h)}], \\ p_{\min}^{\pm} &= \frac{1}{G-\beta}[\alpha \pm \sqrt{\alpha^2 + 2(G-\beta)(F+h)}], \end{aligned} \tag{C.2}$$

where + and − correspond to the two regions outside the two separatrix branches, and max and min to $\psi = 0$ and $m\psi = \pm\pi$, respectively. In libration, $\dot{\psi} = 0$ at ψ_0, at which

$$p_0 = -\frac{\alpha \cos m\psi_0}{G + \beta \cos m\psi_0}, \tag{C.3}$$

so that

$$\cos m\psi_0 = \frac{FG - h\beta + \sqrt{(h\beta + FG)^2 - 2\alpha^2 hG}}{\alpha^2 - 2F\beta}, \tag{C.4}$$

and $|\psi| \leq |\psi_0|$. (When $\alpha = 0$, $\cos m\psi_0 = h/F$.)

Writing the Hamiltonian equations of motion in separated coordinates and factorizing $\dot{p}$ using $p^{\pm}_{\mathrm{max}}$ and $p^{\pm}_{\mathrm{min}}$, we can integrate the expression for $\dot{p}$ to obtain time t as a function of p in terms of the elliptic integral of the second kind (see Gradshteyn & Ryzhik 1980); this is then readily inverted using Jacobian elliptic functions (for their numerical evaluation, see e.g. Press et al. 1986) to give $p(t)$.

For rotation, we have

$$\frac{m}{2}\sqrt{G^2 - \beta^2}t = \int_b^p \frac{\mathrm{d}p}{\sqrt{(a-p)(p-b)(p-c)(p-d)}}, \tag{C.5}$$

where a, b, c, d denote $p^{\pm}_{\mathrm{max}}, p^{\pm}_{\mathrm{min}}$ in ascending or descending order depending on which of the two rotation cases is chosen. On integration we obtain

$$t = \frac{4F(\arcsin\lambda(p), r)}{m\sqrt{G^2 - \beta^2}\sqrt{(a-c)(b-d)}}, \tag{C.6}$$

where $F(,)$ is the elliptic integral of the second kind and

$$\lambda(p) = \sqrt{\frac{(a-c)(p-b)}{(a-b)(p-c)}}, \qquad r = \sqrt{\frac{(a-b)(c-d)}{(a-c)(b-d)}}. \tag{C.7}$$

One-half of the rotation period is obtained by setting $p = a$, i.e. using $F(\pi/2, r)$. Defining

$$s_{\mathrm{r}}(t) \equiv \mathrm{sn}[\frac{m}{4}\sqrt{G^2 - \beta^2}\sqrt{(a-c)(b-d)}\,t, r], \tag{C.8}$$

and $c_{\mathrm{r}}(t)$ similarly using cn[ ], where sn[ ], cn[ ] are Jacobian elliptic functions ($c_{\mathrm{r}}(t)^2 = 1 - s_{\mathrm{r}}(t)^2$), we obtain from (C.6)

$$p(t) = \frac{c_{\mathrm{r}}(t)^2 bc + a(s_{\mathrm{r}}(t)^2 c - b)}{c - s_{\mathrm{r}}(t)^2 b - c_{\mathrm{r}}(t)^2 a}. \tag{C.9}$$

For libration, $p^{\pm}_{\mathrm{min}}$ are complex-valued; now we have

$$\frac{m}{2}\sqrt{G^2 - \beta^2}t = \int_d^p \frac{\mathrm{d}p}{\sqrt{(a-p)(p-d)[(p-e)^2 + f^2]}}, \tag{C.10}$$

where e, f are defined by $p^{\pm}_{\rm min} \equiv e \pm {\rm i} f$. From this,

$$t = \frac{2F(\mathrm{arccot}\gamma(p), w)}{m\sqrt{G^2 - \beta^2}\sqrt{sq}}, \tag{C.11}$$

where

$$w = \frac{1}{2\sqrt{sq}}\sqrt{-(s-q)^2 + (a-d)^2}, \tag{C.12}$$

and

$$s^2 \equiv (e-a)^2 + f^2, \qquad q^2 \equiv (e-d)^2 + f^2, \tag{C.13}$$

and

$$\gamma(p) = \sqrt{\frac{q(a-p)}{s(p-d)}}. \tag{C.14}$$

Defining

$$s_{\rm l}(t) \equiv \mathrm{sn}[\frac{m}{2}\sqrt{G^2 - \beta^2}\sqrt{sq}\, t, w], \tag{C.15}$$

we obtain

$$p(t) = \frac{(c_{\rm l}(t) + 1)^2 sd + s_{\rm l}^2 qa}{(c_{\rm l}(t) + 1)^2 s + s_{\rm l}^2 q}. \tag{C.16}$$

When $p \longrightarrow a$, i.e. $t \longrightarrow T/2$, where T is the total libration period, (C.16)) must be computed in the limit of small $(T/2 - t)$. Now

$$p(t) = \frac{m^2(G^2 - \beta^2)s^2(T/2 - t)^2 d/16 + a}{m^2(G^2 - \beta^2)s^2(T/2 - t)^2/16 + 1}. \tag{C.17}$$

The above calculation for the libration case applies only when $\alpha \neq 0$. In the symmetric case $\alpha = 0$, we choose $p = 0$ at $t = 0$; now

$$t = \frac{F(\arcsin\gamma(p), w)}{m\sqrt{\beta h - GF}}, \tag{C.18}$$

where

$$\begin{aligned} \gamma(p) &= \frac{p}{p_{\rm max}}\sqrt{\frac{p_{\rm max}^2 - p_{\rm min}^2}{p^2 - p_{\rm min}^2}}, \\ w &= \frac{p_{\rm max}}{\sqrt{p_{\rm max}^2 - p_{\rm min}^2}}. \end{aligned} \tag{C.19}$$

Defining now

$$s_{\mathrm{l}}(t) \equiv \mathrm{sn}[m\sqrt{\beta h - GF}\, t, w], \tag{C.20}$$

we have

$$p(t) = \pm \frac{s_{\mathrm{l}}(t)}{\sqrt{p_{\mathrm{max}}^{-2} - \frac{c_l(t)^2}{p_{\mathrm{min}}^2}}}, \tag{C.21}$$

the sign depending on the sign of ψ_0 when $t = 0$.

Appendix D

A general point transformation for closed orbits

As explained in chapter 2, we need a coordinate system such that, under a point transformation, motion in a toy Hamiltonian naturally projects to that around a closed orbit in the target Hamiltonian. The easiest way to do this is to move a local Cartesian coordinate frame along the closed orbit, and to determine the point transformation from the local frame to the global one. Let the closed orbit be given by $\boldsymbol{x}_0 = \boldsymbol{x}_0(s)$, where s parametrizes the orbit trajectory. If the motion is circulating (loop-type), this parameter can be the time. If it is librating (e.g. the banana-orbits in the logarithmic potential), s can be the path length. In the following, prime will denote differentiation with respect to s. We define the generalized velocity, tangent to the orbit, as $\boldsymbol{v}_0 \equiv \boldsymbol{x}_0'$. The 'speed' $v_0 \equiv |\,\boldsymbol{v}_0\,|$ is always nonzero also for librating motion because of the choice of s.

The point transformation between the local and global coordinate systems is given by

$$\begin{aligned} x &= x_0(s) + \frac{y_0'}{v_0}a, \\ y &= y_0(s) - \frac{x_0'}{v_0}a, \end{aligned} \tag{D.1}$$

where a is the perpendicular distance from the orbit: when the y-axis of the local coordinate frame is aligned with $\boldsymbol{v}_0$, values on the x-axis give a (for librating motion, the orbit is traversed only in one direction for this definition). The problem with this transformation is, of course, that while $(s,a) \to (x,y)$ is unique, the inverse is not manifestly such. The inverse transformation is now defined as follows: the point (s,a) corresponding to a given (x,y) can be determined by finding where the line with slope $(y-y_0(s))/(x-x_0(s))$ coincides with the trajectory normal (slope $-x_0'(s)/y_0'(s)$):

$$(y-y_0(s))y_0'(s)+(x-x_0(s))x_0'(s)=0, \tag{D.2}$$

and the correct solution (s,a) of (D.2) is the one giving a continuous image in s of motion in H (often, but not necessarily always, the value of s is the one corresponding to the (x_0,y_0) closest to (x,y)–minimizing this distance gives equation (D.2)). However, there can be orbits for which a continuous image in s is impossible to achieve. For example, in the case of the x_1-orbit family – loops that cross themselves – in the rotating logarithmic potential (chapter 5), this transformation works only for orbits very near the closed one.

From (2.9) we have the transformation equations for the momenta; for the closed orbit, $a=0=\dot{a}$, and

$$\begin{aligned} p_s &= x_0'\dot{x}_0+y_0'\dot{y}_0=v_0^2\dot{s}, \\ p_a &= (y_0'\dot{x}_0-x_0'\dot{y}_0)/v_0=(y_0'x_0'-x_0'y_0')\dot{s}/v_0=0, \end{aligned} \tag{D.3}$$

and, along the closed orbit, $\oint p_s \mathrm{d}s = \oint \boldsymbol{p}\cdot\mathrm{d}\boldsymbol{x}$. Thus the (s,a)-system can, in its turn, be mapped by a point transformation to a toy system where the other action is zero for the closed orbit image, and the other one equals $\oint p_s \mathrm{d}s/2\pi$. When s,a,p_s,p_a are given, the momenta $\boldsymbol{p}$ in Cartesian coordinates can be computed from

$$\begin{pmatrix} p_x \\ p_y \end{pmatrix} = \mathsf{M}^{-1}\begin{pmatrix} p_a \\ p_s \end{pmatrix}, \quad \mathsf{M}=\begin{pmatrix} \frac{\partial x}{\partial a} & \frac{\partial y}{\partial a} \\ \frac{\partial x}{\partial s} & \frac{\partial y}{\partial s} \end{pmatrix}. \tag{D.4}$$

Combining (D.1) with a second point transformation, we have the chain

$$\boldsymbol{w}_{\mathrm{toy}} \leftrightarrow \boldsymbol{w}_{(\mathrm{s,a})} \leftrightarrow \boldsymbol{w}_{\mathrm{target}}. \tag{D.5}$$

For example, loop-type orbits can be mapped between the isochrone polar toy coordinates and (s, a) in the similar fashion as was done with the small J_r-loops in chapter 2. For minor-orbits of librating type, one can use the transformation

$$\begin{aligned} x_1 &= g(a)h(s), \\ x_2 &= e(a)f(s), \end{aligned} \tag{D.6}$$

where x_i are the Cartesian coordinates of a toy harmonic oscillator. Analogously with the method for the small J_r-loops, the numerically integrated values for the closed orbit in the target Hamiltonian can now be employed to map a $J_x = \oint p_s \mathrm{d}s/2\pi$, $J_y = 0$-orbit in harmonic oscillator into the target one under (D.6). Specifically, using (2.9) and considering the closed orbit $a = 0 = \dot{a}$, we obtain (choosing $g(0) = 1$)

$$\frac{\mathrm{d}h}{\mathrm{d}s} = \frac{p_s}{p_x(h)}, \tag{D.7}$$

where $p_x(x) = \sqrt{2\omega J_x - \omega_x^2 x^2}$ is the x_1-momentum of the harmonic oscillator. The function $h(s)$ can be solved for either by numerical integration or by directly integrating (D.7), which yields

$$\frac{h}{2}\sqrt{2\omega_x J_x - \omega_x^2 h^2} + J_x \arcsin h\sqrt{\frac{\omega_x}{2J_x}} = \int_0^s p_s \, \mathrm{d}s. \tag{D.8}$$

Note that our coordinate system optimizes this transformation in a natural way: the closed orbit determines only $h(s)$ (and even its job can be left to the generating function S if so wished) at $e(0) = 0 = g'(0)$ and a given $g(0)$, while $f(s)$ is arbitrary. Thus e, f, g (and h) can be given as functional series whose coefficients are to be optimized for each constructed torus. When fitting tori with $J_y > 0$, the closed target orbit $\boldsymbol{x}_0(s)$ can be extrapolated in s to cover as much of the target configuration space as needed.

When constructing tori with the orbit integration method of chapter 3, the form (D.6) is advantageous as it gives x_i explicitly when s, a are known. When using the least-squares construction technique, it is convenient to use a form where x_1 and x_2 are exchanged with, respectively, s and a in (D.6): then there is no numerical root finding anywhere, and the phase-space point $\boldsymbol{w}$ corresponding to a given trial $\boldsymbol{\theta}$ is obtained directly.

For example, the banana tori of Fig. 2.7 were constructed using (D.6). (Both torus constructing methods give the curves in Fig. 2.7 accurately.) This point transformation was indeed necessary; without it, the toy actions vary much more in the $\boldsymbol{\theta}$-space, and a good fit is harder to produce. In (s, a)-space, banana orbits look very much like bow-ties: (D.6) gives a crude initial distortion of the rectangular harmonic-oscillator box. It is not quite as complete as the elliptic coordinate distortion used for boxes. However, as a test, we tried using (D.6) instead of elliptic coordinates for boxes, and again it gave a satisfactory initial distortion, although not as good as elliptic coordinates would. To be able to use things like elliptic coordinate transformations is basically a stroke of luck; generally, this point transformation is more or less optimal for automatic use.

Bibliography

Arnold V.I., 1989, Mathematical Methods of Classical Mechanics, 2nd ed., Springer-Verlag

Binney J.J., 1978, Mon. Not. Roy. Ast. Soc., 183, 779

Binney J.J., Kumar S., 1993, Mon. Not. Roy. Ast. Soc., 261, 584

Binney J.J., Spergel D., 1984, Mon. Not. Roy. Ast. Soc., 206, 159 (B&S)

Binney J.J., Tremaine S., 1987, Galactic Dynamics. Princeton University Press, Princeton

Blitz L., Binney J.J., Lo K.Y., Bally J., Ho T.P., 1993, Nature, 361, 417

Boozer A.H., 1982, Phys. Fluids, 25, 520; 1984, Tech. report PPPL-2082 (Princeton)

Candy J., Rozmus W., 1991, J. Comp. Phys., 92, 230

Channell P.J., 1978, in Jorna S., ed., Topics in Nonlinear Dynamics, AIP Conference Proceedings 46, 248, American Institute of Physics

Chirikov B.V., 1979, Phys. Rep., 52, 265

de Zeeuw T., 1985, Mon. Not. Roy. Ast. Soc., 216, 273

Forest E., Ruth R.D., 1990, Physica 43D, 105

Gerhard O.E., 1985, Astron. Astrophys., 151, 279

Goldstein H., 1980, Classical Mechanics, 2nd ed., Addison-Wesley, Reading

Gradshteyn I.S., Ryzhik I.M., 1980, Table of Integrals, Series and Products. Academic Press, London

Hénon M., Heiles C., 1964, Astrophys. J., 69, 73

Hénon M., 1982, Physica 5D, 412

Kaasalainen M., 1994, Mon. Not. Roy. Ast. Soc., 268, 1041

Kaasalainen M., Binney J.J., 1994a, Mon. Not. Roy. Ast. Soc., 268, 1033

Kaasalainen M., Binney J.J., 1994b, to appear in the proceedings for Integration Algorithms for Classical Mechanics, Waterloo, October 1993

Kaasalainen M., Binney J.J., 1994c, Phys. Rev. Lett., submitted

Kent S.M., de Zeeuw T., 1991, Astron. J., 102, 1994

Kormendy J., 1993, in Dejonghe H., ed., Galactic Bulges, Reidel, Dordrecht

Kuijken K., Tremaine S., 1994, Astrophys. J., 421, 178

Levison H.F., Richstone D.O., 1987, Astrophys. J., 314, 476

Lichtenberg A., Lieberman M., 1983, Regular and Stochastic Motion. Springer, New York

Martens C.C, Ezra G.S., 1987, J.Chem.Phys., 86, 279

Martinet L., Udry S., 1990, Astron. Astrophys., 235, 69

McGill C., Binney J.J., 1990, Mon. Not. Roy. Ast. Soc., 244, 634 (McG&B)

Miralda J., Schwarzschild M., 1989, Astrophys. J., 339, 752

Ollongren A., 1967, Astron. J., 72, 436

Poincaré H., 1894, Les Méthodes Nouvelles de la Mécanique Céleste, Vols. I–III, Gauthier-Villars, Paris. Reprinted by Dover, 1957

Pöschel J., 1982, Commun. Pure Appl. Math., 25, 653

Press W., Flannery B., Teukolsky S., Vetterling W., 1986, Numerical Recipes. Cambridge University Press, New York

Ratcliff S.J., Chang H.M., Schwarzschild M., 1984, Astrophys. J., 279, 610

Richstone D.O., 1984, Astrophys. J., 281, 100

Saha, P., 1989, D. Phil. thesis, Oxford University

Schwarzschild M., 1979, Astrophys. J., 232, 236

Walker G.H., Ford, J., 1969, Phys. Rev. 188, 416

Warnock R., 1991, Phys. Rev. Lett., 66, 1803

Warnock R., Ruth R., 1991, Phys.Rev.Lett. 66, 990 ; 1992, Physica 56D, 188