

DOE/ET-53088-382

IFSR #382

**Chaotic Transport in Hamiltonian Dynamical Systems
with Several Degrees of Freedom**

Hyungtae Kook
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

July 1989

**CHAOTIC TRANSPORT IN HAMILTONIAN DYNAMICAL
SYSTEMS WITH SEVERAL DEGREES OF FREEDOM**

Publication No. _____

Hyungtae Kook, Ph.D.

The University of Texas at Austin, 1989

Supervising Professor: Philip J. Morrison

The dynamics and transport in the phase space of Hamiltonian systems with $N+1$ degrees of freedom are studied. Such systems can be reduced to $2N$ -dimensional symplectic maps using Poincaré's surface of section method. A $2N$ -dimensional symplectic map which satisfies the twist condition can be derived from a Lagrangian generating function, which enables the action formulation for orbits. Reversible maps of this form have 2^{N+1} invariant symmetry sets. We show that orbits of reversible, symplectic,

twist maps can be classified by frequency, symmetry, and Morse index. The properties of orbits in each class are studied, including various bifurcations. The structure of phase space is viewed via resonances and chaotic layers using a four-dimensional example. A fast and stable Newton's algorithm for finding periodic orbits, based on the variational method, is devised for a class of Lagrangian mappings. Its computation time and required storage space are shown to be linear in the period. Partial barriers in the chaotic regions of phase space and Markov models for transport in area preserving maps, as well as the existence of co-dimension one manifolds for higher dimensional maps and their roles in transport are discussed. The characteristic function method is used to obtain the diffusion tensor for $2N$ -dimensional symplectic maps. At lowest order this method gives the quasilinear result, and a series in higher order correlations is developed. Comparison of the theory to numerical experiments, using a four-dimensional example, shows good agreement for moderately large parameters. A study of Arnol'd diffusion for the thick layer case shows that the short time correlations in one canonical plane affect the diffusion in the other plane even in the limit of zero coupling. Accelerator modes exist for doubly periodic maps, and cause enhancements in the diffusion. The long-time behavior of the correlation function is also discussed.

**CHAOTIC TRANSPORT IN HAMILTONIAN DYNAMICAL
SYSTEMS WITH SEVERAL DEGREES OF FREEDOM**

APPROVED BY

SUPERVISORY COMMITTEE:

Philip A. Morrison

James S. Meir

R. D. Fetters

Wendell Horton

R. Brueckner

Harry L. Furman

Copyright
by
Hyungtae Kook
1989

**CHAOTIC TRANSPORT IN HAMILTONIAN DYNAMICAL
SYSTEMS WITH SEVERAL DEGREES OF FREEDOM**

by

HYUNGTAE KOOK, B.S., M.S.

DISSERTATION

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

THE UNIVERSITY OF TEXAS AT AUSTIN

August, 1989

Dedicated to my parents

Acknowledgements

I gratefully thank my advisor, Dr. James D. Meiss. Without his kind guidance and endless encouragement, this work would not have been possible. Whenever I have been faced with difficult problems, he has always been with me to help resolve the problems. He has helped me not to lose sight of the larger picture in this project and has always treated me as a collaborator, which has given me pride in our work.

I also thank professors Philip Morrison, Roger Broucke, Richard Hazeltine, Wendell Horton and Harry Swinney for their careful reading of my dissertation and their helpful suggestions. I am especially indebted to professors Morrison, Horton and Hazeltine for their interest and valuable suggestions on applications of this work to fusion problems.

I wish to express my gratitude to Dr. Robert MacKay for his hospitality during my visit to Warwick University and for many other valuable communications with him. I would also like to thank professors Doochul Kim, Koo-Chul Lee and Duk-In Choi for their interest and valuable discussions on my work during my visit to Seoul National University.

It is my pleasure to thank Drs. Seunghwan Kim, Jonghoon Oh, Sangyoon Kim, Jaewan Kim, Jangjin Chae, Bongguen Hong and Yilbong Kim for their friendship and fruitful discussions on various problems of

dynamics and physics. I have also enjoyed discussions and friendship with other IFS fellows, Qi Chen, Chris Kueny, Raul Acevedo, Ohjin Kwon, Junghoon Han and Yongjin Kim.

I am grateful to the Institute for Fusion Studies and the Department of Physics of the University of Texas for their financial support throughout my study in Austin.

I treasure the support and love of my parents and brothers. I am especially indebted to my brother, Hyung Joon Kook, and his family for giving me a warm environment during my stay in Austin.

H. Kook

Austin, Texas

July 1, 1989

Table of Contents

| | | |
|---|-------|-----|
| Acknowledgements | | v |
| Abstract | | vii |
| Table of Contents | | ix |
| List of Figures | | xii |
| 1. Introduction | | 1 |
| 1.1. Overview | | 2 |
| 1.2. Preliminaries from Hamiltonian Mechanics | | 8 |
| 1.3. Poincaré Surface of Section and Symplectic Maps | | 12 |
| 1.4. Examples | | 15 |
| 1.5. Number Theory and Dynamical Systems | | 25 |
| 2. Periodic Orbits for Reversible, Symplectic Maps | | 34 |
| 2.1. Introduction | | 35 |
| 2.2. Symplectic Twist Maps | | 39 |

| | |
|---|------------|
| 2.3. Bifurcations and Loss of Stability | 45 |
| 2.4. Reversibility | 55 |
| 2.5. Minimizing Orbits | 66 |
| 2.6. Resonance Zones, Channels and Chaotic Layers of Four-dimensional Maps | 71 |
| 2.7. Conclusions | 84 |
| 3. Lagrangian Formulation and Newton's Methods | 86 |
| 3.1. Introduction | 87 |
| 3.2. Lagrangian Dynamical Systems | 91 |
| 3.3. Direct Map Iteration using Broyden's Method | 95 |
| 3.4. Linear Newton's Algorithm for Variational Method | 99 |
| 3.5. Discussion | 106 |
| 4. Geometric Structure and Transport in Phase Space | 111 |
| 4.1. Introduction | 112 |
| 4.2. Dynamics in Area preserving Maps | 114 |
| 4.3. Markov Models for Transport in Area Preserving Maps | 125 |
| 4.4. Invariant Manifolds of $N+1$ d.o.f. Hamiltonian Systems | 139 |
| 4.5. Discussion | 147 |
| 5. Statistical Method for Diffusion in Symplectic Maps | 149 |
| 5.1. Introduction | 150 |

| | |
|--|------------|
| 5.2. Diffusion Tensor | 153 |
| 5.3. Characteristic Function Method | 156 |
| 5.4. Principal Terms | 160 |
| 5.5. Quasilinear Diffusion and Corrections | 163 |
| 5.6. Accelerator Modes | 166 |
| 5.7. Numerical Experiments | 169 |
| 5.8. Conclusions | 177 |
| | |
| 6. Summary and Conclusions | 179 |
| | |
| Appendix A. Constant Twist Case | 185 |
| Appendix B. Existence of Minimizing Periodic Orbits | 187 |
| Appendix C. Subroutines for Finding Periodic Orbits | 190 |
| | |
| List of References | 202 |

Vita

List of Figures

| | | |
|--------|---|----|
| 1.1 | Particle motion in storage ring. | 18 |
| 1.2 | Betatron oscillation around the beam center. | 18 |
| 1.3 | Binary Farey tree construction. | 30 |
| 1.4 | Generalized Farey tree construction. | 33 |
| 2.1 | Stability regions in the A-B plane. | 47 |
| 2.2(a) | Stability diagram for the fixed point at $\mathbf{p} = 0$ and $\mathbf{q} = 0$ as a function of k_1 and h for $k_2 = 0$ | 48 |
| 2.2(b) | Stability diagram of the orbit with frequency (1,1,2) on the Fix(S) plane for $k_2=0$ | 49 |
| 2.3 | Collision of multiplier pairs on complex plane. | 52 |
| 2.4 | Symmetry planes projected onto the canonical p_1-q_1 and p_2-q_2 planes. | 60 |
| 2.5 | Configuration space projection of the four symmetric orbits with frequency (12,21,49) when $k_1=0.4$, $k_2=0.3$, $h=0.65$ | 62 |
| 2.6 | Sketch of the evolution of the stability of two symmetric orbits with $\omega = (2,3,4)$ (a) on the Fix(S) and (b) on the Fix(SR _(1,0)) plane as the parameters vary on the path $k_1=10h$, $k_2=0$ | 64 |
| 2.7 | Symmetry breaking bifurcations. | 65 |

| | | |
|---------|--|-----|
| 2.8 | Inflection of the action function upon a symmetry breaking bifurcation. | 69 |
| 2.9 | Pairs of rationals (ω_1, ω_2) generated by Kim-Ostlund tree up to level 13. Equivalently, each point represents the intersection point \mathbf{p} of a periodic orbit with the frequency $\mathbf{p} = (\omega_1, \omega_2)$ on a plane $\mathbf{q} = \text{constant}$ for the integrable system. | 72 |
| 2.10(a) | Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5, k_2=0.3, h=0.2$. (a) on $\mathbf{q} = (0,0)$ plane. | 73 |
| 2.10(b) | Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5, k_2=0.3, h=0.2$. (b) on $\mathbf{q} = (1/2, 1/2)$ plane. | 74 |
| 2.10(c) | Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5, k_2=0.3, h=0.2$. (c) on $\mathbf{q} = (0, 1/2)$ plane. | 75 |
| 2.11 | Sectional view of two stochastic orbits on the $\mathbf{q} = 0$ plane. | 78 |
| 2.12 | Projections of one of the irrational tori surrounding resonance zones with frequency $(0,0)$ | 80 |
| 2.13 | Staircase of rotation numbers for initial conditions on the line $(p_1, -p_1, 0,0)$ | 82 |
| 2.14 | Resonance zones for rotation numbers up to level 5 on the Farey tree (where $(1,1,2)$ is level 0). | 83 |
| 3.1. | Evolution of the points on a periodic orbit for the mapping Eq. (3.32). | 110 |
| 4.1 | Cantorus partial barrier. | 119 |

| | | |
|-----------|--|---------|
| 4.2 | Resonance of frequency 1/3 for the standard map in symmetry coordinates. | 123 |
| 4.3 | Markov chain near a boundary circle | 130 |
| 4.4 | Schematic of the connected chaotic region of an area preserving map near a boundary circle. | 132 |
| 4.5 | Overlap of turnstiles. | 135 |
| 5.1 | Principal terms represented by paths in (n,k) space (a) for $\chi_t[(1,1),\pm(1,1)]$ and (b) $\chi_t[(1,0),\pm(1,0)]$ | 162 |
| 5.2 | Accelerator mode stability regions. | 167 |
| 5.3 | Comparison of numerical diffusion coefficients with theoretical results. | 170-172 |
| 5.4 | Enhancement of diffusion coefficient D_{aa} due to the presence of the stable accelerator mode with $\mathbf{j}=(1,0)$ | 174 |
| 5.5 | Oscillatory dependence of $\lim_{c \rightarrow 0} D_{cc}/D_{ql}$ on a. | 176 |
| Table 2.1 | Symmetry planes for the symmetric orbits with frequency (m_1, m_2, n) of a four-dimensional reversible map. | 59 |

CHAPTER 1

INTRODUCTION

1.1. Overview

Many systems are known to be Hamiltonian, including the systems of celestial bodies, magnetic field lines, geodesic flows, geometric optics, hydrodynamics of ideal fluids, etc. A state of a Hamiltonian system is represented by a point in even-dimensional phase space and the dynamics of the system can be depicted by flows in phase space.

For an integrable Hamiltonian system of $N+1$ degrees of freedom (d.o.f.), all motions are regular and can be described as uniform rotations on invariant $(N+1)$ -tori; the reason for using '+1' is because the number of the degrees of freedom can be reduced by one using Poincaré's surface of section method to give a $2N$ -dimensional symplectic map (see section 1.3). Thus, for an integrable case, the $2N+2$ dimensional phase space is foliated by an $N+1$ parameter family of $(N+1)$ -tori.

Upon small perturbations, all resonant tori are typically broken and chaotic (stochastic) motions start to appear in some regions of phase space; e.g., near the unstable equilibrium points of a perturbed one-dimensional pendulum. Roughly speaking, chaotic motion implies that the trajectory wanders on a manifold of dimension greater than $N+1$ in phase space. Chaotic motion in phase space is characterized by a positive Lyapunov exponent which is an average divergence of nearby orbits in the chaotic region. Namely, in a chaotic region, there exists a sensitive dependence of the motion on its initial conditions. This sensitive dependence implies that for a practical purpose knowledge of the history of the motion is rapidly lost

and, hence, the motion is irreversible. Consequently, it is quite plausible that a statistical description may apply to the evolution of phase volume in chaotic regions. A natural question is to determine to what extent the phase space is chaotic as described above.

The Kolmogorov-Arnol'd-Moser (KAM) theorem [Arnol'd 1978] shows that under sufficiently small perturbations most nonresonant invariant tori (KAM tori) survive, on which motions are quasiperiodic and, hence, still regular. The theorem also implies the existence of 'regular regions' of finite measure in phase space that retain the smoothness associated with motions on the tori. Therefore, chaos is restricted to the outside of the regular regions even though the regular regions are not perfectly bounded for the systems of more than two degrees of freedom. Furthermore, it is seen that within every regular region there are also chaotic regions, and inside these chaotic regions there are again regular regions. This structure continues ad infinitum. This determines the generically complicated structure of Hamiltonian phase space as a mixture of regular and chaotic regions on all scales.

For an area preserving map, which can be obtained as a Poincaré return map for two d.o.f. Hamiltonian flows, the KAM tori are invariant circles in two-dimensional surface of section and, therefore, they can divide phase space into distinct regions. Thus chaotic motions are limited within each region bounded by these invariant circles and, hence, global transport does not occur. As nonlinearity increases, more invariant circles break up and chaos in phase space increases. When the last invariant circle breaks up,

chaotic motions can occur across the whole phase space. Within each chaotic region, however, the motions are still limited by the presence of partial barriers. Two kinds of them are known to exist: cantorus partial barriers and partial separatrices of resonances, which are the remnants of invariant circles and separatrices, respectively. These also exist on every scale.

When there are more than two degrees of freedom, the situation is quite different from the above. In this case, the energy surface has $2N+1$ dimensions and the dimensionality of invariant tori, $N+1$, is no longer large enough for them to divide the energy surface. In other words, there is no inside or outside of an invariant torus. Consequently, the chaotic regions on all scales are connected, in what is called the Arnol'd web. This implies that, under any virtually non-zero perturbations, a universal instability exists over almost the entire phase space, resulting in Arnol'd diffusion along the chaotic web. Arnol'd diffusion typically occurs on a very long time scale [Nekhoroshev 1965]. The structure of phase space for these higher d.o.f. cases is not well known yet; however, a recent work shows the possibility of existence of partial barriers which play roles similar to those for area preserving maps (see chapter 4).

Despite rather less understanding of the phase space for Hamiltonian systems of more than two degrees of freedom, it seems quite clear that the geometric structure within chaotic regions acts as a barrier to chaotic motions in phase space. In fact, this is closely related to the fundamental problem of validation of the ergodic hypothesis in Hamiltonian dynamical

systems. A system is ergodic if a time average taken along a motion is equivalent to the average taken over an ensemble of states with the same energy. Generally, this hypothesis does not hold for finite d.o.f. Hamiltonian systems since a chaotic orbit does not fill densely and uniformly the energy surface. However, there still remains the unanswered question of whether the hypothesis might hold as the number of degrees of freedom goes to infinity (i.e., the thermodynamic limit).

In this work we study the dynamics of several d.o.f. Hamiltonian systems, and the geometric structure of phase space and transport thereby. We concentrate on symplectic maps, which allow an efficient way of studying Hamiltonian systems in numerical calculations as well as in theoretical analysis. The presentation is organized as follows.

In chapter 1, some preliminaries from Hamiltonian mechanics are briefly reviewed. It is shown that the Hamiltonian flows in $2N+1$ dimensional energy surface can be studied by a Poincaré surface of section with dimension $2N$. The induced maps on the surface of section are $2N$ -dimensional symplectic maps. Several examples of Hamiltonian systems are shown emphasizing three d.o.f. systems. Some results from number theory, which are relevant to studies of dynamical systems, are reviewed; continued fraction expansion, binary Farey tree construction for the reals, and simultaneous rational approximations for higher dimensional cases.

In chapter 2, we consider $2N$ -dimensional symplectic maps with special properties: positive definite twist and reversibility. The twist condition guarantees the existence of a Lagrangian type generating function. A

symmetric orbit has a reflection symmetry and has points on symmetry planes. A classification of symmetric orbits is obtained using commutation relations between relevant operators. Symmetric periodic orbits are useful since they are easily found and quasiperiodic orbits can be approximated by symmetric periodic orbits. Orbits of special interests occur at minima and minimax points of the action. Typically, these orbits coincide with symmetric orbits for systems only slightly perturbed from the integrable case. Various bifurcations of orbits are discussed. We also show that the generalized Farey tree construction for three-frequency dynamical systems has a very close contact with the dynamics of four-dimensional symplectic maps.

In chapter 3, a formal method to obtain a map by discretizing a continuous time flow is presented. Two schemes of Newton's method for finding periodic orbits in Lagrangian dynamical systems are also presented. A direct map iteration method is shown using Broyden's quasi-Newton algorithm. A fast and stable Newton algorithm based on the variational method is devised for unstable periodic orbits. It is shown that the computation time and the size of the storage space are linear in the period of the orbit.

In chapter 4, the well-known dynamics in area preserving maps are reviewed emphasizing the ideas of flux and partial barriers. These ideas play a central role in constructing Markov models for transport in area preserving maps. Three different models are discussed: Markov chain via cantori, Markov tree via cantori and Markov chain of resonances. We discuss the existence of normally hyperbolic invariant manifolds with their

stable and unstable manifolds of co-dimension one. For higher d.o.f. cases, these co-dimension one manifolds play a role similar to the partial barriers in area preserving maps.

In chapter 5, we present a statistical method to obtain an analytical diffusion tensor for symplectic maps using the characteristic functions. In the lowest approximation, the quasilinear result is recovered. It is shown that the low order corrections explain the oscillation of the diffusion coefficient around the quasilinear limit. The theoretical predictions are compared with numerical results, and show excellent agreement when the parameters are moderately large. We show that accelerator modes exist in doubly periodic maps. The parameter regions of first order stable accelerators are obtained analytically. Their effects on enhancement of diffusion are discussed using several kinds of statistics, and their role in determining long time behavior of correlations is discussed.

In chapter 6, we conclude.

1.2. Preliminaries from Hamiltonian Mechanics

A Hamiltonian system is the one describable by a Hamiltonian function $H(\mathbf{p}, \mathbf{q}, t)$. A state of the system is represented by a point (\mathbf{p}, \mathbf{q}) in $2N+2$ dimensional phase space, where \mathbf{q} is the $N+1$ dimensional configuration coordinate vector and \mathbf{p} is the canonical momentum vector conjugate to \mathbf{q} . The dynamics in phase space is governed by Hamilton's equations

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}}$$

$$\frac{d\mathbf{q}}{dt} = \frac{\partial H}{\partial \mathbf{p}}$$

for given initial conditions $(\mathbf{p}(0), \mathbf{q}(0))$. These equations can be also derived from a variation principle,

$$\delta \int \mathbf{p} \cdot d\mathbf{q} - H dt = 0$$

The variation is subject to $\delta \mathbf{q} = 0$ at the endpoints. For now, we consider autonomous systems for convenience, whose Hamiltonians do not have explicit time dependence. The nonautonomous case can be considered as autonomous in an extended phase space of $2N+4$ dimensions.

A completely integrable system is defined as the one which has $N+1$ constants of the motion (integrals), $I_i(\mathbf{p}, \mathbf{q})$, such that

- (1) they are pairwise in involution; by definition, the Poisson bracket of any pair vanishes;

$$[I_i, I_j] \equiv \sum_k \frac{\partial I_i}{\partial q_k} \frac{\partial I_j}{\partial p_k} - \frac{\partial I_j}{\partial q_k} \frac{\partial I_i}{\partial p_k} = 0$$

(2) The vectors of the gradients $\nabla I_i(\mathbf{p}, \mathbf{q})$ are independent almost everywhere.

Then, it can be shown that for a bounded system the motion for given values of the I_i 's is a uniform rotation on an $(N+1)$ -dimensional torus [Berry 1978, Arnol'd 1978]; these are called invariant tori. Thus phase space is foliated in an $N+1$ parameter family of $N+1$ tori. The condition (2) includes integrable systems which has separatrix motions on which the integrals cease to be independent.

There is a standard way to coordinatize phase space which leads to the so-called action-angle variables as follows:

(1) Consider a torus parametrized by given values of the I_i 's. There are many ways to construct a set of independent J_i 's from a combinations of I_i 's. We fix one set by

$$J_i = \oint_{\gamma_i} \mathbf{p} \cdot d\mathbf{q}$$

where γ_i 's are irreducible cycles on the torus: there are $N+1$ independent irreducible cycles on $N+1$ torus. The J_i 's are called the 'action' variables.

(2) The J_i 's are constants of motion. Therefore, the Hamiltonian is now a function only of J_i 's, since by Hamilton's equations, $\partial H / \partial \theta_i = 0$. Thus the conjugate coordinates, θ_i , are the 'angle' variables for the rotational motion on the torus, whose frequencies are given as

$$\omega_i(\mathbf{J}) \equiv \frac{d\theta_i}{dt} = \frac{\partial H(\mathbf{J})}{\partial J_i}$$

The motion on a torus of $\omega(\mathbf{J})$ is periodic if the ω_i 's are commensurate; i.e., $\omega \cdot \mathbf{m} = 0$ for some integer vector $\mathbf{m} \neq 0$. We call such tori resonant tori. Similarly, when the ω_i 's are incommensurate, the motion is quasiperiodic and densely covers the torus. We call such tori nonresonant tori. The resonant tori form a countably infinite set of measure zero in phase space. The commensurability of the frequency is very important in determining the survival of such motions upon the introduction of small perturbations.

The KAM theorem [Arnol'd 1978, Lichtenberg 1983] deals with how the motions are altered upon the addition of perturbations which destroy the integrability. Roughly speaking, it states that most nonresonant tori survive under sufficiently small perturbations, while all resonant tori are destroyed leaving only a finite number of commensurate motions.

The surviving invariant tori play an important role in transport for two d.o.f. Hamiltonian systems. In this case the phase space is four dimensional and the motion is confined to a constant energy hypersurface of three dimensions. A surviving invariant torus is two dimensional and, hence, it can divide the three dimensional energy surface into two distinct regions. Thus the excursion of motion on the energy surface is bounded by these invariant tori: invariant tori are perfect barriers to transport in two d.o.f. Hamiltonian systems. However, the $(N+1)$ -dimensional torus does not divide the $(2N+1)$ -dimensional energy surface for $N > 1$. Clearly, the

invariant tori are not perfect barriers any more in the systems of more than two degrees of freedom.

Another feature of Hamiltonian systems is that Poincaré's integral, $\oint_c \mathbf{p} \cdot d\mathbf{q}$, is conserved around any loop following the flow. The infinitesimal version of this is the preservation of the symplectic form

$$\omega((\delta\mathbf{p}, \delta\mathbf{q}), (\delta\mathbf{p}', \delta\mathbf{q}')) = \delta\mathbf{p} \cdot \delta\mathbf{q}' - \delta\mathbf{p}' \cdot \delta\mathbf{q}$$

A sequence of integral invariants can be also represented by high order exterior products ω^k , $k = 1, \dots, N+1$. The last invariant, ω^{N+1} , gives the conservation of phase space volume (Liouville's theorem).

1.3. Poincaré Surface of Section and Symplectic Maps

Poincaré surface of section is a technique for obtaining a description of the motion on a phase space of reduced dimension and it is suited best for two d.o.f. systems since the energy surface in this case is three dimensional. The concept, however, is easily generalized to higher d.o.f. systems even though a visualization of flows on the surface section is still difficult. A general procedure is as follows.

Consider a Hamiltonian flow on a $2N+2$ dimensional phase space with coordinates $z = (q_0, p_0, z_1, z_2, \dots, z_N)$. The flow lies on a $2N+1$ dimensional energy surface of a constant energy h . We assume that this surface is compact so that the motion is bounded. Suppose there is a $2N+1$ dimensional surface Σ on which $dq_0/dt = \partial H/\partial p_0 \neq 0$; e.g., a plane of constant q_0 . This surface intersects the energy surface transversely as well as the flow. A transversal intersection of two $(2N+1)$ -dimensional surfaces is $2N$ dimensional, we denote it by Σ_h . On Σ_h , one can solve $H = H(q_0, p_0, z_1, \dots, z_N) = h$ for $p_0(h, q_0, z_1, \dots, z_N)$ (Implicit function theorem). Then, the equations of motion can be written in the following form:

$$\frac{dp_i}{dq_0} = \frac{\frac{dp_i}{dt}}{\frac{dq_0}{dt}} = \frac{-\frac{\partial H}{\partial q_i}}{\frac{\partial H}{\partial p_0}} = -\frac{\partial p_0}{\partial q_i}$$

$$\frac{dq_i}{dq_0} = \frac{\frac{dq_i}{dt}}{\frac{dq_0}{dt}} = \frac{\frac{\partial H}{\partial p_i}}{\frac{\partial H}{\partial p_0}} = -\frac{\partial p_0}{\partial p_i} \quad \text{for } i = 1, \dots, N$$

Thus p_0 is the reduced Hamiltonian on Σ_h , and q_0 takes a role of the time.

Now consider the map T induced by the flows on Σ_h ; $T: \Sigma_h \rightarrow \Sigma_h$. T is obtained by following the flow starting from a particular time q_0 to the next return of the flow to Σ_h , i.e., when q_0 returns to its initial value. In particular, the return map is symplectic since a Hamiltonian flow preserves the symplectic form. Let DT represent the Jacobian matrix, or the tangent map of T . Then T is symplectic providing

$$\widetilde{DT} J DT = J = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix}$$

where “ \sim ” denotes transpose of a matrix and J is the $2N \times 2N$ anti-symplectic matrix shown. Another important feature of the Poincaré return map is that it is invertible since the flow is reversed under the time reversal. For $N = 1$ (two degrees of freedom), Σ_h is two-dimensional and the preservation of the symplectic form implies that the return map T is area preserving ($\det(DT) = 1$) and orientation preserving ($\det(DT) > 0$).

We have seen that a Hamiltonian flow induces a symplectic map. Despite its formal existence, it is often difficult in practice to derive a

symplectic map from Hamiltonian flows since it involves an integration of the flow for the first return to the surface of section; this is essentially equivalent to solving Hamilton's equations. In most cases a reasonable approximation is obtained for the integration of the flow for the first return; however, care must be exercised to maintain the symplecticity of the map. The resulting symplectic map is expected to show qualitatively the same behavior as the original Hamiltonian flow. One of the most important merits of maps is that iterating a map is much faster and more reliable than numerically integrating the Hamilton's equations.

1.4. Examples

1.4.1. Restricted Three-Body Problem

The problem of three bodies is the simplest nontrivial problem in celestial mechanics and has been a long-standing problem stimulating many research over a hundred years, while a complete solution is out of reach. The problem can be stated as follows [Whittaker 1964]:

“Three particles attract each other according to Newton’s law, so that between each pair of particles there is an attractive force which is proportional to the product of the masses of the particles and the inverse square of their distance apart. They are free to move in space, and are initially supposed to be moving in any given manner; the problem is to determine their subsequent motion”.

Originally, the system has nine degrees of freedom, each three degrees-of-freedom of which are due to the motion of each body. However, the motion of the system is more restricted since there exist several integrals of motion. Namely, it is known that the effective number of degrees of freedom can be reduced by the number of integrals which are pairwise in involution (Lie’s Theorem [Hagihara 1970]). For a system of n bodies there exist five such integrals beside the energy integral: the total linear momentum vector, and the magnitude of the total angular momentum together with one of its components (three components of angular momentum are not all in involution). Therefore, for the system of three

bodies, the number of degrees of freedom is reduced to four. This can be further reduced by one using the energy integral, which would give a six-dimensional symplectic map (see section 1.3).

Now, let us suppose that the motion of three bodies takes place in a plane; this is the case when three initial velocity vectors are in the plane of the bodies. The system has six degrees of freedom. Now, beside the energy integral, there are three integrals pairwise in involution; the total linear momentum vector in the plane of the motion, and the total angular momentum perpendicular to the plane. Consequently, the system of three bodies on a plane is equivalent to a system of three degrees of freedom.

The resulting reduced Hamiltonian has a form too complicated to be analyzed completely [Whittaker 1964]. With a further assumption that three bodies are of equal masses, Davoust and Broucke [Davoust 1982] have studied the stability and bifurcations of several families of periodic orbits.

We remark that several versions of restricted motions of three bodies have been devised and studied. The most famous one might be the 'planar, circular, restricted three-body problem', which has two degrees of freedom and has been used to study the Kirkwood gaps in the asteroid belt between Mars and Jupiter [Berry 1978]. A more developed version due to Wisdom includes the effect of the eccentricity of the Jupiter's motion, which results in a four-dimensional Poincaré map [Wisdom 1982, Dermott 1983]. Another is the problem of Sitnikov in which the third body of zero mass oscillates on the line perpendicular to the plane of motions of the first two and going through the center of mass [Moser 1973].

Hamiltonian systems with lower degrees of freedom is often useful to study many body systems. One of the interesting problems in Astrophysics is the 'galactic dynamics' which studies the motions and the stabilities of stars in galaxies [Binney 1987]. Although a galaxy is composed of many stars, it can be considered that a star feels the potential of the background mass distribution in the galaxy rather than the interactions from other individual stars. For instance, the potential at the inner parts of distorted triaxial elliptical galaxies can be assumed to be the one of a three-dimensional harmonic oscillator, hence of three degrees of freedom, with small anharmonic perturbations [Contopoulos 1986].

1.4.2. Beam-Beam interaction in Particle Accelerator

In a particle accelerator, using two opposing beams contained in two storage rings is accepted as a better way to achieve larger center-of-mass energy than bombarding a target at rest with one beam. In this case, it is necessary to have large beam densities to increase collision events. However, it is known that the so-called 'beam-beam interaction' becomes significant as beam-densities increase, which often leads to a beam expansion or loss. To explain such phenomena, there have been attempts from many disciplines [Month 1979].

Since the collisions between the opposing beams occur in a limited region, the motion of a particle can be considered to be integrable along the storage ring until the particle enters the collision region (see Fig. 1.1).

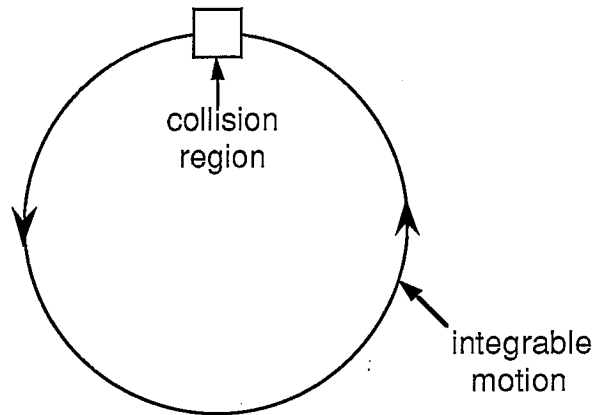


Fig. 1.1 Particle motion in storage ring

As the particles travel around the storage ring, they oscillate transversely about the beam center; these are called 'betatron oscillation' (see Fig. 1.2).

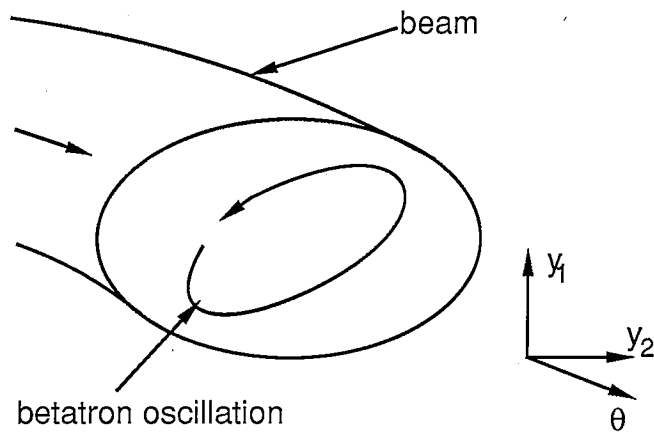


Fig. 1.2 Betatron oscillation around the beam center

In the lowest approximation, the betatron oscillation can be considered as a composition of two linear harmonic oscillations in the vertical and horizontal directions, respectively. This oscillation is affected by nonlinear beam-beam interaction at the collision region of two beams and the particle loss results from the growth of betatron oscillation amplitude. A simple description for this mechanism is as follows. At the collision region, the effect of beam-beam interaction is modeled by periodic forcings, given by a delta function in θ , to the harmonic oscillations in the vertical and horizontal positions y_1 and y_2 [Helleman 1979]:

$$\frac{d^2 y_1}{d\theta} = -Q_1^2 y_1 + \sum_t \delta(\theta - 2\pi t) f_1(y_1, y_2)$$

$$\frac{d^2 y_2}{d\theta} = -Q_2^2 y_2 + \sum_t \delta(\theta - 2\pi t) f_2(y_1, y_2)$$

Here Q_1 (Q_2) is the machine tune given by the ratio of vertical (horizontal) oscillation frequency to the revolution frequency around the ring. The functions f_1 and f_2 represent nonlinear interactions at the collision region. The corresponding four-dimensional map equations can be obtained exactly by integrating over the one revolution around the ring, using momenta $p_i = dy_i/d\theta$. Variables are discretized by choosing their values just before the collisions:

$$y_1' = ay_1 + \frac{b}{Q_1}p_1$$

$$p_1' = -bQ_1y_1 + ap_1 + f_1(y_1', y_2')$$

$$y_2' = cy_2 + \frac{d}{Q_2}p_2$$

$$p_2' = -dQ_2y_2 + cp_2 + f_2(y_1', y_2')$$

where $a = \cos(2\pi Q_1)$, $b = \sin(2\pi Q_1)$, $c = \cos(2\pi Q_2 + \beta)$, $d = \sin(2\pi Q_2 + \beta)$, and the phase factor β is the phase difference between the horizontal and vertical tunes.

Mostly, area preserving maps have been used to model this problem, assuming the degree of freedom associated with the horizontal oscillation is negligible. However, considering the fact that the required revolutions of particles are in practice quite large ($\geq 10^8$), it may turn out that couplings between these degrees of freedom will play a significant role.

1.4.3. Magnetic Field Line Flow and Guiding Center Motion of Charged Particle in Time-Varying Field

The line of force described by a magnetic field can be thought of as a conservative dynamical system due to the conservation of flux ($\nabla \cdot \mathbf{B} = 0$). In fact, a variational principle for the field line flow is given by

$$\delta \int d\lambda \mathbf{A}(\mathbf{x}) \cdot \frac{d\mathbf{x}}{d\lambda} = 0$$

where $\mathbf{A}(\mathbf{x})$ is the vector potential and $\mathbf{x} = \mathbf{x}(\lambda)$ is the parametrization of the field line with any arbitrary parameter λ [Cary 1983]. The variation is subject to the boundary condition $\mathbf{A} \cdot \delta \mathbf{x} = 0$ at the endpoints. The resulting Euler-Lagrange equation is simply

$$(\nabla \times \mathbf{A}) \times \frac{d\mathbf{x}}{d\lambda} = 0$$

which shows that $d\mathbf{x}/d\lambda$ is parallel to $\mathbf{B} (= \nabla \times \mathbf{A})$. Furthermore, using the gauge transformation of \mathbf{A} , it can be shown that a time-independent magnetic field line gives a Hamiltonian flow with one and a half degrees of freedom. Suppose

$$\mathbf{A} \cdot d\mathbf{x} = A_1(x_1, x_2, x_3) dx_1 + A_2(x_1, x_2, x_3) dx_2 + A_3(x_1, x_2, x_3) dx_3.$$

Choose a gauge transformation, $\mathbf{A}' = \mathbf{A} + \nabla S$ such that $\partial S / \partial x_1 = -A_1$, giving the new vector potential

$$\mathbf{A}' \cdot d\mathbf{x} = A_2'(x_1, x_2, x_3) dx_2 + A_3'(x_1, x_2, x_3) dx_3.$$

By introducing a new variable $p = A_2'(x_1, x_2, x_3)$ and solving it for x_1 , we obtain the canonical form for $\mathbf{A}' \cdot d\mathbf{x}$.

$$\mathbf{A}' \cdot d\mathbf{x} = p dx_2 - H(p, x_2, x_3) dx_3$$

Here, $H(p, x_2, x_3) = -A_3'(x_1(p, x_2, x_3), x_2, x_3)$ plays the role of Hamiltonian. Thus the field line flow is a one and a half degrees-of-freedom Hamiltonian flow, parametrized by x_3 . Equivalently, it is two degrees-of-freedom flow in an extended phase space. An example is given by a vector potential such that

$$H = -A_3 = \frac{1}{2} \iota' x^2 + k \cos y \sum_n \delta(z - 2\pi n)$$

where ι' is the constant shear of magnetic field in the x direction. One can easily obtain the standard map by discretizing x and y variables by choosing their values just before the value of z takes an integer multiple of 2π (see section 1.4.2).

A similar conversion to a canonical Hamiltonian form does not work for a time-dependent field case. In this case, to define the flow, we consider the motion of a charged particle in a time-dependent magnetic field. Its motion has three and a half degrees of freedom; gyration around the field line, motion along the field line, motion across the field line, and time-dependence. By averaging out the fast gyro-motion, one can expect to obtain a description of the guiding center motion of a charged particle which, then, has two and a half degrees-of-freedom. This is done using Littlejohn's Lagrangian [Littlejohn 1983].

Littlejohn's Lagrangian is derived from the usual Lagrangian for particle motion by taking the gyroaverage. It is given by

$$L(Z, \dot{Z}, t) = \frac{1}{\varepsilon} \mathbf{A}^* \cdot \dot{\mathbf{X}} + \varepsilon \mu \dot{\Theta} - H$$

Here the units are chosen so that $e = m = c = 1$, ε is the adiabatic ordering parameter which represents the ratio of gyroradius to the scale length. \mathbf{X} represents the guiding center position, μ magnetic moment, and Θ the gyrophase. H is the Hamiltonian given by $H = \Phi + \mu B + 1/2 U^2 + o(\varepsilon)$, where Φ is the electric potential and $U = \mathbf{b} \cdot \dot{\mathbf{X}}$ is the parallel velocity of the guiding center. The modified vector potential \mathbf{A}^* is defined by $\mathbf{A}^* = \mathbf{A} + \varepsilon U \mathbf{b} + o(\varepsilon^2)$,

where \mathbf{b} is the unit vector along magnetic field. \mathbf{A}^* is the gyroaverage of the canonical momentum. Finally, Z represents $(\mathbf{X}, U, \Theta, \mu)$. The terms of the higher order in ε can be obtained systematically. The resulting Euler-Lagrange equations give the drift equations of the guiding center up to a desired order in ε .

We now consider the Lagrangian up to the zeroth order in ε , which corresponds to taking gyroaverage. For the simplest case, we neglect the terms of the electric potential and the magnetic moment. Namely, we have

$$L = (\mathbf{A} + U\mathbf{b}) \cdot \dot{\mathbf{X}} - \frac{1}{2}U^2$$

We further assume that the magnetic field points predominantly along the z -axis,

$$\mathbf{A} = \Psi(x, y, z, t) \mathbf{k} + x \mathbf{j} \quad \rightarrow \quad \mathbf{B} = -\mathbf{k} \times \nabla \Psi + \mathbf{k}$$

$$\mathbf{b} \approx \mathbf{k}$$

where \mathbf{j} and \mathbf{k} are unit vectors in the y - and the z -axis directions, respectively. Thus we obtain

$$L = (\Psi + U) \dot{z} + x \dot{y} - \frac{1}{2}U^2$$

The canonical momenta are obtained using the usual definitions as

$$p_z = \frac{\partial L}{\partial \dot{z}} = \Psi + U \quad \text{and} \quad p_y = \frac{\partial L}{\partial \dot{y}} = x$$

Finally, the Hamiltonian is obtained as

$$H(p_z, z, x, y, t) = \frac{1}{2}U^2 = \frac{1}{2}(p_z - \Psi(x, y, z, t))^2$$

Note that x plays a role as a canonical momentum conjugate to y . This Hamiltonian gives the following equations of motion.

$$\dot{x} = -\frac{\partial H}{\partial y} = U \frac{\partial \Psi}{\partial y}$$

$$\dot{y} = -\frac{\partial H}{\partial x} = -U \frac{\partial \Psi}{\partial x}$$

$$\dot{p}_z = -\frac{\partial H}{\partial z} = U \frac{\partial \Psi}{\partial z}$$

$$\dot{z} = -\frac{\partial H}{\partial p_z} = U$$

Note that U is not constant, whose equation of motion is given by the p_z equation as

$$\dot{p}_z = U \frac{\partial \Psi}{\partial z} = \frac{\partial \Psi}{\partial x} \dot{x} + \frac{\partial \Psi}{\partial y} \dot{y} + \frac{\partial \Psi}{\partial z} \dot{z} + \frac{\partial \Psi}{\partial t} + \dot{U}$$

Using the equations of motion for x , y and z , one can see $\dot{U} = -\frac{\partial \Psi}{\partial t}$.

The above result shows that the motion of a charged particle in time-dependent magnetic field has two and a half degrees of freedom when one degree of freedom is removed by taking the gyroaverage.

1.5 Number Theory and Dynamical Systems

As will be shown in section 2.2, the frequency, or the rotation number, of an orbit is defined through a lift of the mapping over the covering space \mathcal{R}^N as

$$\omega = \lim_{t \rightarrow \infty} \frac{\mathbf{q}t}{t}$$

Thus, by definition of periodic orbits, Eq.(2.14), every periodic orbit has a frequency $\omega = \mathbf{m}/n$. There also exist orbits which have rational frequencies which are, however, not periodic. These are heteroclinic orbits lying at the intersections of stable and unstable manifolds of hyperbolic periodic points. An orbit is quasiperiodic if the components of ω are incommensurate, and if it is recurrent; an orbit homoclinic to a quasiperiodic orbit has the frequency of the quasiperiodic orbit, but it is not a quasiperiodic orbit because it is not recurrent.

The KAM theorem states that most tori on which orbits are quasiperiodic survive small perturbations [Arnol'd 1978]. In particular, numerical results in area preserving maps strongly imply that the last KAM circle is typically the one with a noble frequency; the last KAM circle has a frequency of the golden mean for the standard map [Greene 1979]. Thus it appears that persistence of an orbit is determined to some extent by number theoretical properties of its frequency. However, the details of persistence

as well as a possible classification of such irrational tori will depend on the particular dynamics.

In studies of dynamical systems, many results from number theory have been used. In this section we review briefly some of relevant results. The continued fraction expansion and the Farey tree construction are well known and have established themselves as important tools in studying dynamical systems with two relevant frequencies. There exists no such a well-established number theory for dynamical systems with more than two frequencies. However, there are a few candidates and, among them, the generalized Farey tree construction devised by Kim et al [Kim 1987] is shown. This construction is used to find the structure of channels of resonances in phase space for four-dimensional symplectic maps (see section 2.6).

1.5.1 Continued Fraction Expansion

The continued fraction algorithm [Khinchin 1964] is one of the most frequently used tools in the study of two frequency dynamical systems, e.g. area preserving maps. It is a scheme for systematically generating a sequence of rational approximations, m_i/n_i to a number ω ; the measure of approximation is given by the distance $\| \omega - m_i/n_i \|$, where $\| \cdot \|$ denotes a norm which will be defined below.

The continued fraction expansion ('simple' continued fraction in Khinchin's terminology) of a positive number ω is the sequence $[a_0, a_1, a_2, \dots]$ of positive integers such that

$$\omega = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots}}$$

We call a_j the j -th element. For irrationals, such a sequence is unique and infinite. For rationals, the sequence is finite and there exist two equivalent sequences for each rational. That is, for rationals,

$$[a_0, \dots, a_n + 1] = [a_0, \dots, a_n, 1]$$

The rationals obtained by truncations of the sequence at some level are called convergents, or approximants of the number ω ;

$$m_j/n_j = [a_0, a_1, \dots, a_j]$$

The continued fraction expansion has many important properties some of which we mention below.

(1) It is strongly convergent: for any $\varepsilon > 0$, there exists an integer j such that

$$|m_i - \omega n_i| < \varepsilon \quad \forall i \geq j$$

where $||$ indicates the ordinary Euclidean distance (Weak convergence uses the norm $|\omega - m_i/n_i|$).

(2) All the rational approximants to an irrational number are best approximants in the strong sense: given an approximant m/n to ω , there are no m'/n' with $n' < n$ satisfying $|m' - \omega n'| < |m - \omega n|$.

(3) There is a sequence of theorems which gives a classification of irrationals in the sense of their approximability by rationals. For example, the irrationals whose elements include large numbers admit good approximation by rational approximants:

$$\frac{1}{n_j^2(a_{j+2})} < \left| \omega - \frac{m_j}{n_j} \right| \leq \frac{1}{n_j^2 a_{j+1}}$$

Conversely, an irrational like $\gamma = [1, 1, 1, \dots] = \frac{1+\sqrt{5}}{2}$ is one of the most poorly approximated by rational approximants. The number γ is called the golden mean. The irrationals which have $a_j = 1$ for all j larger than some integer i are called noble numbers.

(4) Another merit of the algorithm is that the quadratic numbers (roots of a 2nd degree monic polynomial) are characterized by the periodicity of their continued fraction (Lagrange's theorem). The noble numbers are the simplest case of quadratic irrationals in that the period of the continued fraction expansion is just one.

1.5.2 Farey Tree

The Farey tree [Hardy 1954] is well known as an alternative scheme for approximating irrational numbers. Even though it converges more slowly than the continued fraction, it is often more valuable due to its elegance in organizing the real numbers. The tree generates all rational numbers together with the irrationals as limits of paths, and every number is associated with a unique path.

The tree construction is basically the successive bisections of a given interval. An interval is defined by its two end points, a pair of neighboring rationals, m_1/n_1 and m_2/n_2 ; neighboring means $m_1n_2 - m_2n_1 = \pm 1$. The bisection is the mediant operation, defined by adding the numerators and denominators of the pair, respectively:

$$\frac{m_3}{n_3} = \frac{m_1 + m_2}{n_1 + n_2}$$

The level of daughter, m_3/n_3 , on the tree is defined to be higher by one than its parents, m_1/n_1 and m_2/n_2 . One can easily show that a daughter is again a neighbor to both of its parents. Thus successive bisections can be iterated ad infinitum, resulting in a self-similar structure on any scale. Figure 1.3 shows the Farey tree construction over the interval $[0/1, 1/1]$ up to a few levels.

The Farey tree construction has a close connection with the continued fraction expansion described above. Consider the interval $[0,1]$ and let the level of $1/2$ be 2 on the tree, then the sum of elements of the continued fraction expansion gives the level of the corresponding rational on the tree. Conversely, to obtain the continued fraction expansion of a rational on the tree, first find a path from $1/1$ to the rational. Starting from $1/1 \equiv [0,1]$, go down along the path, adding 1 to the last element of the previous continued fraction expansion unless the direction of the path changes. When the direction changes, the rule for the last element is

$$[a_0, a_1, \dots, a_j] \rightarrow [a_0, a_1, \dots, a_j - 1, 2]$$

From its construction, it is clear that the Farey tree provides a binary coding scheme for real numbers; starting from $1/2$, assign the left going direction to '+' and the right going direction to '-'. For each path we obtain a sequence of +'s and -'s.

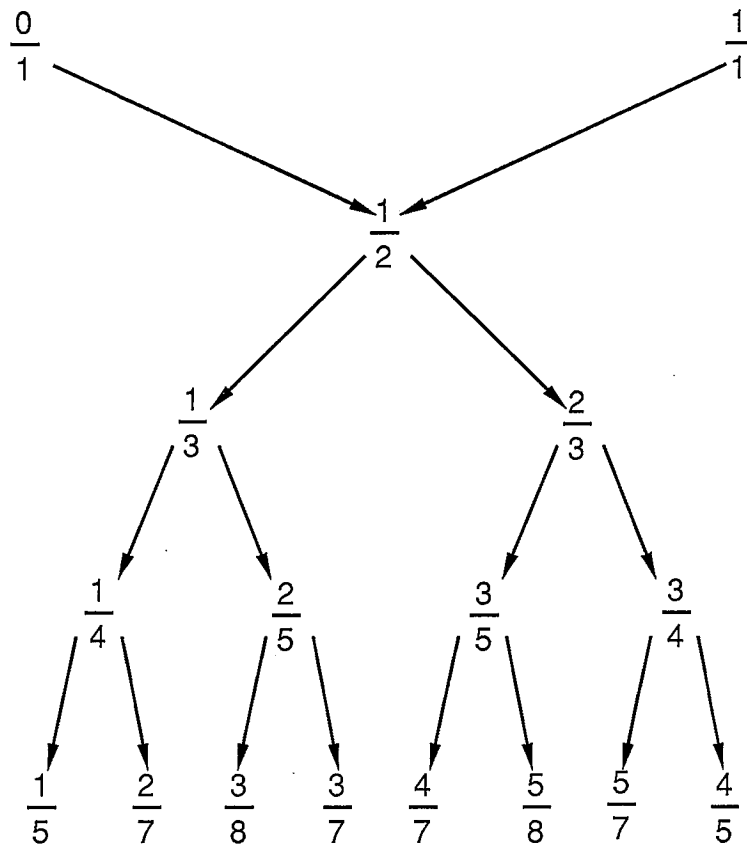


Fig. 1.3 Binary Farey tree construction.

1.5.3 Simultaneous Rational Approximations

For the higher dimensional case, one would like to approximate an N dimensional vector ω by a set of rationals with common denominator, \mathbf{m}/n . For the case $N=2$, such a continued fraction algorithm was first given by Jacobi (1868). His idea has been extended for general N by Perron (1907) and the algorithm is known as the "Jacobi-Perron" algorithm [Bernstein 1971]. It is weakly convergent for any N ; however, it does not generate all the best approximants. Furthermore, the characterization of numbers for which the Jacobi-Perron algorithm is periodic is unknown.

Many other algorithms have been devised to approximate irrationals with simultaneous n -tuple rationals. Among these, the work of Brentjes may be valuable for application to dynamical systems [Brentjes 1981]. His algorithm produces all the best rational approximants for $N=2$. However the sequence of approximants also contains other rationals besides the best and, consequently, converges more slowly than optimal. The implementation of this algorithm on the computer is straightforward except that one must do inordinately long calculations to determine the direction of successive approximation at each level. No algorithm has yet been shown to be suited for the characterization of numbers analogous to Lagrange's theorem.

Recently, the Farey tree construction method has been successfully extended to the $N=2$ case by Kim and Ostlund [Kim 1987]. Even though it

does not generally produce the sequence of best approximants, it presents a very elegant way of organizing simultaneous rationals.

The tree is obtained from a set of three integer vectors $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ in \mathcal{R}^3 , which satisfy $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = 1$, for example $(1,0,1)$, $(0,1,1)$, and $(0,0,1)$. Each vector represents a pair of rationals with common denominator, $\mathbf{a} \rightarrow (a_1/a_3, a_2/a_3)$. The vectors can be represented symbolically as the vertices of a right isosceles triangle, as shown in Fig. 1.4. The level zero rational is obtained by vector (Farey) sum of the hypotenuse vertices: $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$, and is represented by the midpoint of the hypotenuse. The original triangle is now divided into two right isosceles triangles, each of which can be used to construct a new vertex on level one, giving four triangles for the next level. The vertices of each Farey triangle satisfy $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = 1$ by construction. One can show that every rational pair (in the convex hull of $(\mathbf{a}, \mathbf{b}, \mathbf{c})$) is obtained by a finite path on the tree; however, each pair in the interior of the triangle $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is obtained by two paths. This multiplicity is inconvenient, but not crucial in the manipulation of the algorithm. Furthermore every rational pair is in lowest common terms. It has been also shown that the numbers can be addressed by binary codes so that the algorithm can be handled with the shift operation on the corresponding address. This coding scheme makes the computational implementation of the algorithm very convenient.

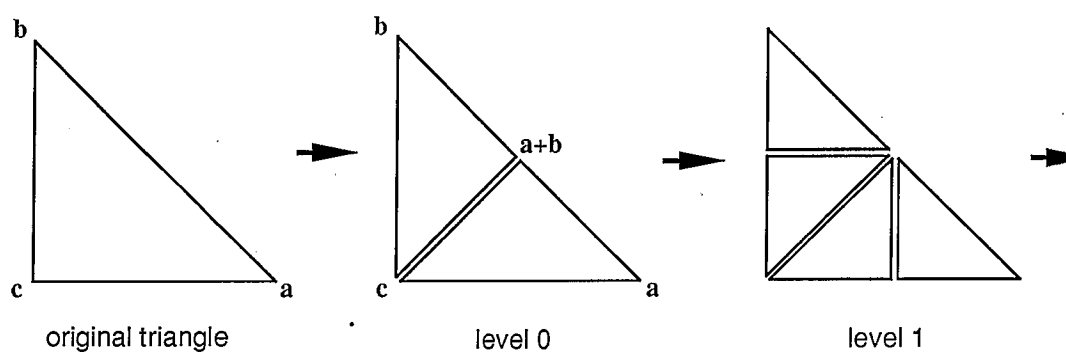


Fig. 1.4 Generalized Farey tree construction. At each level a new vertex is obtained from the Farey sum of vertices of the hypotenuse of the preceding triangle. This gives two isosceles triangles for which the tree can be continued.

CHAPTER 2

PERIODIC ORBITS FOR REVERSIBLE, SYMPLECTIC MAPS

2.1. Introduction

Symplectic maps are useful for the description of many physical problems, including the motion of the asteroids, particle accelerators, plasma wave heating, and the restricted three-body problem. In particular any system described by a time independent Hamiltonian with $N+1$ degrees of freedom generates a $2N$ dimensional symplectic map by the Poincaré surface of section. Similarly an N degree of freedom, periodically time dependent Hamiltonian flow generates a symplectic map of the phase space by strobing at the forcing period.

A great deal is known about the periodic orbits of two-dimensional symplectic maps (area preserving maps). Poincaré proposed, in his “last geometric theorem” that a map on the annulus which rotates the upper and lower boundaries of the annulus in opposite directions must have a fixed point. This was later proved by Birkhoff, and implies the existence of at least two periodic orbits for each period. For the case of twist maps on a cylinder, Mather and Aubry proved the existence of periodic orbits for any rotation frequency [Mather 1982, Aubry 1983]. The mapping preserves the cyclic order of points on these orbits; this implies that a limit of frequencies approaching an irrational gives a quasiperiodic orbit.

The quasiperiodic orbits either densely cover invariant circles, or invariant Cantor sets. If the frequency is sufficiently irrational, the Kolmogorov-Arnold-Moser (KAM) theorem [Arnold 1978] guarantees the

former occurs for small enough perturbations from the integrable case. The invariant Cantor sets, called cantori [Percival 1979], are the remnants of the invariant circles after they have been destroyed. Knowledge of the properties of such orbits is extremely useful in understanding the break-up of invariant circles [Greene 1979, MacKay 1983], and the transport of orbits in the chaotic region of phase space [MacKay 1984]. In particular, cantori are the dominant barriers for the flow of trajectories through phase space.

For higher dimensional mappings, much less is known. Recently, Bernstein and Katok [Bernstein 1987] have shown that for C^1 perturbations from the integrable case there exist at least $N+1$ periodic orbits in a neighborhood of the unperturbed torus for any given commensurate frequency. However, the properties of incommensurate frequency orbits are less clear. For many systems it can be shown that the KAM tori are destroyed for strong enough perturbations [MacKay 1988], and it seems reasonable that, by analogy with the area preserving case, they will be replaced by remnant quasiperiodic orbits. In some cases these do exist and can be shown to be cantori [Chen 1989]; however, whether this typically occurs and how the transition from the KAM regime occurs remain to be understood.

In this chapter we consider symplectic maps with special properties: positive definite twist, and reversibility. In section 2.2 we introduce our notation and the twist condition, which guarantees the existence of a Lagrangian type generating function. Positive definite twist is the analogue of the Legendre condition for continuous time systems. Orbital stability is

discussed in section 2.3, and in section 2.4 we review the notion of reversibility: a system is reversible if each orbit is related to its time reverse by a symmetry transformation. Symmetric orbits must have points on N dimensional symmetry planes, and can be classified by which of these planes they intersect. This property is especially useful to reduce computational effort. An invariant torus must be invariant under the time reversal symmetry, and can be studied by considering nearby periodic orbits which are also symmetric.

Orbits of special interests occur at minima and minimax points of the action. For the area preserving case Aubry and Mather have shown the existence of global minima for each frequency. In our computations we easily find orbits which are locally minimizing for any rational frequency. Minimizing orbits are not always symmetric: we discuss symmetry breaking bifurcations in section 2.5.

As an illustration we use a four dimensional generalization of the standard map due to Froeschlé [Froeschlé 1971]. The configuration space is the two dimensional torus, and we use coordinates (q_1, q_2) with unit periods. The momenta are designated (p_1, p_2) , and the map is given by

$$T: \begin{cases} \mathbf{p}' = \mathbf{p} - \nabla V(\mathbf{q}) \\ \mathbf{q}' = \mathbf{q} + \mathbf{p}' \end{cases} \quad (2.1)$$

$$V(q_1, q_2) = \frac{-1}{(2\pi)^2} \{k_1 \cos(2\pi q_1) + k_2 \cos(2\pi q_2) + h \cos[2\pi(q_1, q_2)]\}$$

When only one of the three parameters (k_1, k_2, h) is non-zero this map has a global invariant: for example, if only h is non-zero then $p_1 - p_2$ is conserved. If $h = 0$ then the map is separable into uncoupled two dimensional standard maps for (q_1, p_1) and (q_2, p_2) . When h and at least one of k_1 or k_2 is non-zero it appears neither to have global invariants, nor to be separable.

For the four dimensional case, the symmetry planes are two dimensional, and thus in section 2.6 we construct the orbital skeleton by plotting symmetric orbits which intersect a given symmetry plane. We find two distinct structures: resonances and channels. Corresponding to each periodic orbit with frequency $\omega = (m_1/n, m_2/n)$ is a resonance; its intersection with the symmetry plane is a two dimensional neighborhood in which there are no other rotational periodic orbits. Points initially in this region oscillate about the resonance for long periods. Corresponding to each commensurability, $\omega_1 j_1 + \omega_2 j_2 + j_3 = 0$, for integers j_1, j_2, j_3 , we find a channel on the symmetry plane. These channels are the directions along which orbits "diffuse" by Arnold diffusion. Resonances occur at the intersection of an infinite number of channels. Outside the resonances and channels are closely packed periodic orbits, as well as the invariant tori or their remnants.

2.2. Symplectic Twist Maps

Consider a dynamical system represented by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$. Here \mathbf{q} is a point in some N -dimensional configuration space and $\dot{\mathbf{q}}$ is the velocity. The action of a path $\mathbf{q}(t)$, $t_0 \leq t \leq t_1$ is defined as

$$W[\mathbf{q}(t)] = \int_{t_0}^{t_1} L(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (2.2)$$

Stationary paths, determined by setting the first variation, δW to zero for $\mathbf{q}(t_0)$ and $\mathbf{q}(t_1)$ fixed, satisfy the Euler-Lagrange equations of motion. If L is periodic in t with period one, then these equations can be formally converted into a mapping of the configuration space onto itself. A discrete Lagrangian can be defined by integrating L along a stationary orbit segment, $\mathbf{q}(t)$, which begins at the point \mathbf{q} at $t=0$, and ends at \mathbf{q}' at $t=1$

$$F(\mathbf{q}, \mathbf{q}') \equiv \int_0^1 L(\mathbf{q}, \dot{\mathbf{q}}, t) dt \quad (2.3)$$

$\mathbf{q}(t)$ stationary

The action of an orbit segment from $t=i$ to $t=j$ can now be written

$$W(\mathbf{q}_i, \mathbf{q}_{i+1}, \dots, \mathbf{q}_j) = \sum_{t=i}^{j-1} F(\mathbf{q}_t, \mathbf{q}_{t+1}) \quad (2.4)$$

Variation of W with respect to the intermediate points yields the mapping equations

$$\frac{\partial}{\partial \mathbf{q}_t} [F(\mathbf{q}_{t-1}, \mathbf{q}_t) + F(\mathbf{q}_t, \mathbf{q}_{t+1})] = 0 \quad (2.5)$$

This equation locally defines a unique mapping, T , from $(\mathbf{q}_{t-1}, \mathbf{q}_t)$ to $(\mathbf{q}_t, \mathbf{q}_{t+1})$ providing $\det(\partial^2 F / \partial \mathbf{q} \partial \mathbf{q}')$ is never zero. We make a stronger assumption, the twist condition: the mixed partial derivative matrix is assumed to be uniformly negative definite, i.e. there exists a $B > 0$ such that for any $\delta \mathbf{q}, \mathbf{q}, \mathbf{q}'$

$$\delta \mathbf{q} \cdot \mathbf{b}(\mathbf{q}, \mathbf{q}') \cdot \delta \mathbf{q} \geq B |\delta \mathbf{q}|^2$$

$$\mathbf{b}(\mathbf{q}, \mathbf{q}') \equiv - \partial^2 F(\mathbf{q}, \mathbf{q}') / \partial \mathbf{q} \partial \mathbf{q}' = F_{12}(\mathbf{q}, \mathbf{q}') \quad (2.6)$$

It can be seen that this globally implies the existence of T [MacKay 1988]. The twist condition is the analogue of the Legendre condition in continuous time.

Equation (2.5) can be converted to Hamiltonian form by noting that $F(\mathbf{q}, \mathbf{q}')$ can be taken to be a canonical generating function where the momenta are defined by

$$\begin{aligned} \mathbf{p}' &= \frac{\partial F(\mathbf{q}, \mathbf{q}')}{\partial \mathbf{q}'} = F_2(\mathbf{q}, \mathbf{q}') \\ \mathbf{p} &= - \frac{\partial F(\mathbf{q}, \mathbf{q}')}{\partial \mathbf{q}} = - F_1(\mathbf{q}, \mathbf{q}') \end{aligned} \quad (2.7)$$

where the subscripts F_1 and F_2 represent the partial derivatives of F with respect to its first and second arguments, respectively. These equations are implicit, but Eq. (2.6) implies they can be inverted to obtain an explicit mapping $\mathbf{z}' = T(\mathbf{z})$ where $\mathbf{z} \equiv (\mathbf{p}, \mathbf{q})$ represents a phase space point. In phase

space, the twist condition has a simple geometrical interpretation: an increment, $\delta\mathbf{p}$, in momentum results in a position change in the same half space:

$$\delta\mathbf{p} \cdot \delta\mathbf{q}' \Big|_{\delta\mathbf{q}=0} > 0$$

A mapping is symplectic if it preserves the symplectic form ω , which is defined for two tangent vectors $\delta\mathbf{z} = (\delta\mathbf{p}, \delta\mathbf{q})$ and $\delta\widehat{\mathbf{z}}$ as

$$\omega(\delta\mathbf{z}, \delta\widehat{\mathbf{z}}) = \delta z^i \omega_{ij} \delta\widehat{z}^j = \delta\mathbf{p} \cdot \delta\widehat{\mathbf{q}} - \delta\mathbf{q} \cdot \delta\widehat{\mathbf{p}} \quad (2.8)$$

so that in canonical coordinates ω_{ij} is the $2N \times 2N$ matrix

$$\omega = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \quad (2.9)$$

Thus if $(DT)^i_j = \partial z^i / \partial \widehat{z}^j$ represents the Jacobian matrix, or tangent map; then T is symplectic providing

$$\widetilde{DT} \omega DT = \omega \quad (2.10)$$

where ' \sim ' designates transpose. If T is obtained from a canonical generating function, F , then it is automatically symplectic.

The symplectic form can also be expressed in terms of configuration space tangent vectors $(\delta\mathbf{q}, \delta\mathbf{q}')$, upon noting that $\delta\mathbf{p} = -F_{11}\delta\mathbf{q} - F_{12}\delta\mathbf{q}'$:

$$\omega(\delta\mathbf{z}, \delta\widehat{\mathbf{z}}) = \delta\widehat{\mathbf{q}} \cdot \mathbf{b} \cdot \delta\mathbf{q}' - \delta\mathbf{q} \cdot \mathbf{b} \cdot \delta\widehat{\mathbf{q}}' \quad (2.11)$$

where \mathbf{b} is $\mathbf{b}(\mathbf{q}, \mathbf{q}')$ defined by Eq. (2.6)

We will assume that the configuration variables represent physical angles with period one ($\mathbf{q} + \mathbf{m} \rightarrow \mathbf{q}$, where \mathbf{m} is an integer vector); thus the configuration space is the N-torus, \mathcal{T}^N , and the phase space is $\mathcal{R}^N \times \mathcal{T}^N$. This implies that the map commutes with translations by any integer vector \mathbf{m} . Defining the translation operator $R_{\mathbf{m}}$ by

$$R_{\mathbf{m}}\mathbf{q} = \mathbf{q} - \mathbf{m}, \quad (2.12)$$

then $TR_{\mathbf{m}} = R_{\mathbf{m}}T$. This implies that the generating function satisfies

$$F(\mathbf{q}+\mathbf{m}, \mathbf{q}'+\mathbf{m}) = F(\mathbf{q}, \mathbf{q}') + \mathcal{F}_{\mathbf{m}}$$

where $\mathcal{F}_{\mathbf{m}}$ is a constant. To interpret $\mathcal{F}_{\mathbf{m}}$ let C be a curve in the phase space which connects (\mathbf{p}, \mathbf{q}) to $(\mathbf{p}, \mathbf{q}+\mathbf{m})$, and $C' = TC$ be its iterate. Parameterize these curves by $\lambda \in [0, 1]$, so that $\mathbf{q}(0)=\mathbf{q}$, $\mathbf{q}(1) = \mathbf{q}+\mathbf{m}$, etc. Using Eq. (2.7) we can show that the Poincaré invariant between C' and C gives $\mathcal{F}_{\mathbf{m}}$:

$$\int_{C'} \mathbf{p}' \cdot d\mathbf{q}' - \int_C \mathbf{p} \cdot d\mathbf{q} = \int_0^1 \frac{dF(\mathbf{q}(\lambda), \mathbf{q}'(\lambda))}{d\lambda} d\lambda = \mathcal{F}_{\mathbf{m}}$$

or that $\mathcal{F}_{\mathbf{m}}$ is the sum of the areas which move through the projections of the curve C onto the canonical planes; i.e. the net flux of symplectic area. The net flux clearly depends only on the homotopy class, \mathbf{m} , of C ; since there are N independent loops on \mathcal{T}^N the net flux has N independent components. We will assume, as in example (2.1), that the net flux is zero.

An orbit is a doubly infinite sequence of configurations $\{ \dots, \mathbf{q}_t, \mathbf{q}_{t+1}, \mathbf{q}_{t+2}, \dots \}$ such that every finite segment $\{ \mathbf{q}_t, \mathbf{q}_{t+1}, \dots, \mathbf{q}_s \}$ is a stationary point

of the action $W(\mathbf{q}_t, \mathbf{q}_{t+1}, \dots, \mathbf{q}_s)$ for given \mathbf{q}_t and \mathbf{q}_s . It is convenient to think of these orbits as occurring on the covering space of the N -torus, \mathcal{R}^N . Then the frequency (if it exists) is defined through a lift of the mapping as

$$\omega = \lim_{t-t' \rightarrow \infty} \frac{\mathbf{q}(t) - \mathbf{q}(t')}{t - t'} \quad (2.13)$$

(and is unique up to an integer vector).

An orbit is periodic with period n if

$$\mathbf{q}_{t+n} = R_{-\mathbf{m}} \mathbf{q}_t, \quad \mathbf{p}_{t+n} = \mathbf{p}_t \quad (2.14)$$

for some integer vector \mathbf{m} . Thus, every periodic orbit has a frequency $\omega = \mathbf{m}/n$, which we also write as $\omega = (\mathbf{m}, n)$. The action for a periodic orbit is

$$W_\omega(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_n) \dots W(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}, \mathbf{q}_0 + \mathbf{m}) \quad (2.15)$$

Equation (2.15) gives a variational principle for periodic orbits: a periodic orbit of frequency ω is a critical point of W_ω where all points $\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}$ are varied freely. This is true because criticality of W_ω implies Eq.(2.5) for $0 < t < n$, and variation with respect to \mathbf{q}_0 implies $F_1(\mathbf{q}_0, \mathbf{q}_1) + F_2(\mathbf{q}_{n-1}, \mathbf{q}_0 + \mathbf{m}) = 0$, or by Eq.(2.7) $\mathbf{p}_0 = \mathbf{p}_n$.

An important class of symplectic maps are those with constant twist, i.e. \mathbf{b} is a constant matrix. Such maps are of physical relevance because nonlinear maps typically have twist at least locally, and furthermore because an arbitrary twist map can locally be approximated by a map with constant twist. This is discussed extensively by Chirikov for the area preserving case [Chirikov 1979].

For the special case of constant twist, and zero net flux we show in Appendix A that the most general form for the generating function is

$$F = \frac{1}{2} (\mathbf{q}' - \mathbf{q}) \cdot \mathbf{b} \cdot (\mathbf{q}' - \mathbf{q}) - V(\mathbf{q}) \quad (2.16)$$

where \mathbf{b} is a symmetric matrix. The mapping generated by Eq.(2.16) is

$$T: \begin{cases} \mathbf{p}' = \mathbf{p} - \nabla V(\mathbf{q}) \\ \mathbf{q}' = \mathbf{q} + \mathbf{b}^{-1} \cdot \mathbf{p}' \end{cases} \quad (2.17)$$

For $N=1$ the twist, \mathbf{b} , can be scaled to unity by a (non-canonical) change of coordinates; however, this cannot be done in general for $N>1$. The example (2.1) is of the form (2.17) with \mathbf{b} chosen to be the identity matrix.

2.3. Bifurcations and Loss of Stability

The stability of an orbit is determined by the tangent map $DT^n = \partial \mathbf{z}_n / \partial \mathbf{z}_0$, which is a symplectic matrix. As is well known [Arnol'd 1978], if λ is an eigenvalue of a symplectic matrix then so are λ^{-1} and λ^* , where “*” denotes complex conjugate. This implies that the λ 's come either in reciprocal pairs which are real or of modulus unity, or in a quadruplet with $\lambda_1 = 1/\lambda_2 = \lambda_3^* = 1/\lambda_4^*$. We will call the eigenvalues of DT^n the multipliers, to distinguish them from a later use of “eigenvalue”. Let $\lambda_1, 1/\lambda_1, \dots, \lambda_N, 1/\lambda_N$ represent the complete set of multipliers and define the traces, ρ_i , and residues, R_i , as

$$\rho_i = \lambda_i + 1/\lambda_i$$

$$R_i = 1/4 (2 - \rho_i) \tag{2.18}$$

The orbit is stable only when all the traces are real and fall in the interval $[-2, 2]$, or equivalently the residues are real and in $[0, 1]$. An orbit with a quadruplet of complex multipliers has a pair of complex conjugate residues; the residues are otherwise real.

In terms of the traces, characteristic polynomial is

$$\lambda^{-n} \det (DT^n - \lambda I) = \prod_{i=1}^N (\rho - \rho_i) = 0 \tag{2.19}$$

In general the coefficients of ρ_j , $j=0,1,\dots, N-1$ in this polynomial give N parameters which determine the stability of the orbit. For $N=2$, the polynomial is quadratic

$$\rho^2 - A\rho + B - 2 = 0$$

$$A \equiv \text{tr}(DT^n) = \rho_1 + \rho_2$$

$$B \equiv \frac{1}{2} \{ [\text{tr}(DT^n)]^2 - \text{tr}[(DT^n)^2] \} = \rho_1 \rho_2 + 2 \quad (2.20)$$

Consequently, the two independent real quantities A and B determine the stability of the periodic orbit [Broucke 1969]. There are seven different regions in the orbit stability diagram in the A - B plane, as shown in Fig. 2.1. The only stable region is the central area labelled " \mathcal{EE} "; it is bounded by the period-doubling bifurcation line ($B = -2A - 2$), the saddle-node bifurcation line ($B = 2A - 2$) and the complex instability parabola ($B = A^2/4 + 2$). All the λ 's have modulus one in the stable region, thus the orbit is doubly elliptic. The other stability regions are labelled using \mathcal{E} for an elliptic multiplier pair, \mathcal{H} for regular hyperbolic, \mathcal{I} for inversion hyperbolic (a pair of negative multipliers). The final region, \mathcal{CQ} , corresponds to a complex quadruplet of multipliers. These seven regions can be mapped onto the parameter space for a given orbit, as shown in Fig. 2.2 for Eq. (2.1).

An alternative construction of the multipliers can be given in terms of the configuration space representation of the linearized mapping DT , which is obtained by linearizing Eq. (2.5):

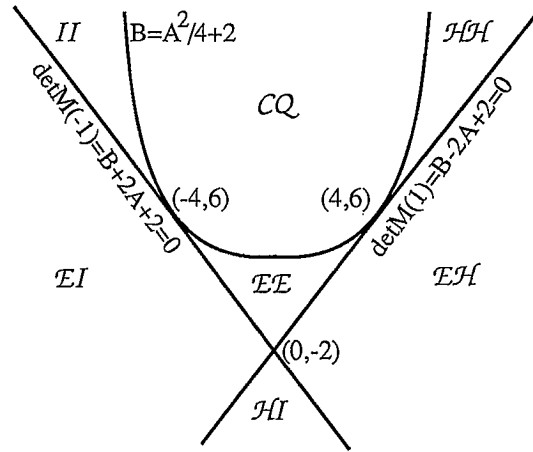


Fig. 2.1 Stability regions in the A-B plane. Stability type for each region is determined by two multiplier pairs whose natures are represented with symbols; \mathcal{E} for elliptic, \mathcal{H} for regular hyperbolic, \mathcal{I} for inversion hyperbolic and \mathcal{CQ} for complex quadruplet.

$$-\mathbf{b}_{t+1} \delta \mathbf{q}_{t+1} + \mathbf{a}_t \delta \mathbf{q}_t - \tilde{\mathbf{b}}_t \delta \mathbf{q}_{t-1} = 0$$

$$\mathbf{b}_t \equiv -F_{12}(\mathbf{q}_{t-1}, \mathbf{q}_t)$$

$$\mathbf{a}_t \equiv F_{11}(\mathbf{q}_t, \mathbf{q}_{t+1}) + F_{22}(\mathbf{q}_{t-1}, \mathbf{q}_t) \quad (2.21)$$

giving a linear second difference equation. The multipliers of a period n orbit are determined by the eigenvalue problem

$$\delta \mathbf{q}_{t+n} = \lambda \delta \mathbf{q}_t$$

which closes the recurrence relation (2.21).

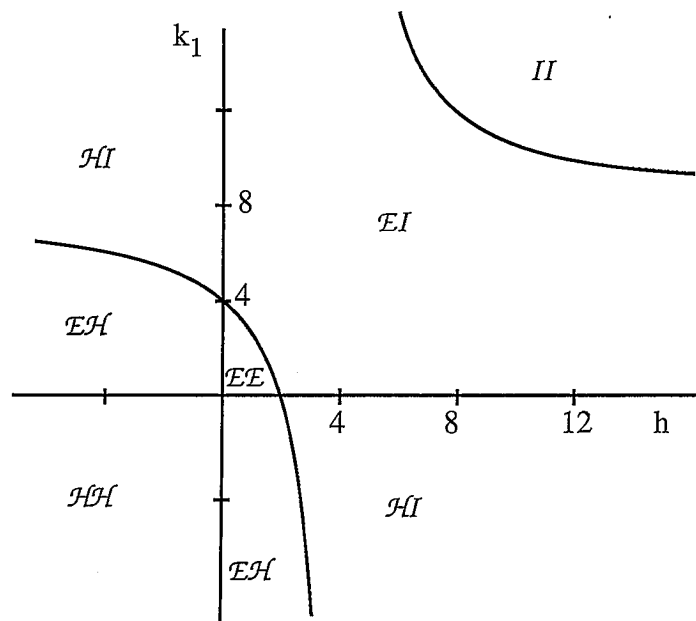


Fig. 2.2 (a) Stability diagram for the fixed point at $\mathbf{p} = 0$ and $\mathbf{q} = 0$ as a function of k_1 and h for $k_2 = 0$. Diagrams for the three other fixed points are obtained by flipping this vertically and/or horizontally. No region of complex instability (CQ) occurs. The Krein signature in the region EE is negative definite.

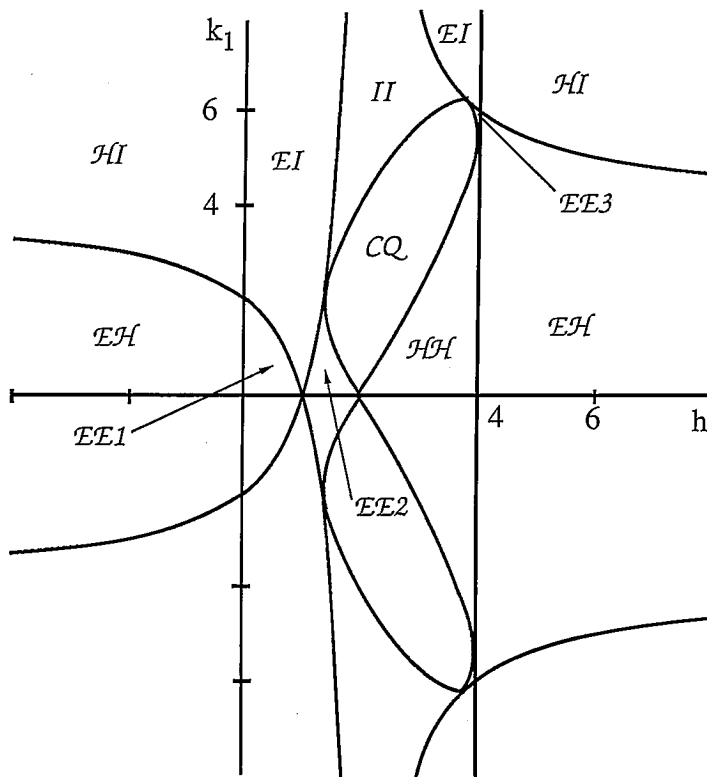


Fig. 2.2 (b) Stability of the orbit with frequency $(1,1,2)$ on the $\text{Fix}(S)$ plane for $k_2=0$. The diagram is symmetric under $k_1 \rightarrow -k_1$. To obtain the corresponding diagram for the orbit on $\text{Fix}(SR_{(1,0)})$ let $h \rightarrow -h$. There are three elliptic regions; the Krein signature in the $EE1$ region is negative definite and in $EE2$ and $EE3$ is mixed.

In matrix form, this can be rewritten as $\mathbf{M} \delta \mathbf{Q} = 0$ where $\delta \mathbf{Q} = (\delta \mathbf{q}_0, \dots, \delta \mathbf{q}_{n-1})$ and the $(Nn) \times (Nn)$ dimensional matrix $\mathbf{M}(\lambda)$ is

$$\mathbf{M}(\lambda) = \begin{pmatrix} a_0 & -\mathbf{b}_1 & 0 & \dots & \frac{-1}{\lambda} \tilde{\mathbf{b}}_0 \\ \tilde{\mathbf{b}}_1 & a_1 & -\mathbf{b}_2 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & a_{n-2} & -\mathbf{b}_{n-1} \\ \lambda \mathbf{b}_0 & \dots & \dots & \tilde{\mathbf{b}}_{n-1} & a_{n-1} \end{pmatrix} \quad (2.22)$$

Thus the multipliers are obtained by solving $\det \mathbf{M} = 0$. This equation is an N^{th} order polynomial in λ and in $1/\lambda$; and since the mapping is symplectic we know this polynomial must be reflexive. Its roots are the multipliers, which are given by Eq. (2.19); thus it must have the form

$$\det \mathbf{M}(\lambda) = C \prod_{i=1}^N (\rho_i - \rho)$$

where C is a constant independent of λ . In fact, C is precisely the coefficient of $(-\lambda)^N$ which is easily obtained from Eq. (2.22) by inspection, thus we obtain

$$\mathcal{M}(\lambda) \equiv \det \mathbf{M}(\lambda) = \prod_{t=0}^{n-1} \det(\mathbf{b}_t) \prod_{i=1}^N (\rho_i - \rho) \quad (2.23)$$

We will discuss these relationships further in section 2.5.

When $\mathcal{M}(1) = 0$, the orbit has a pair of multipliers at +1. If, upon variation of parameters, a periodic orbit reaches a point where a multiplier is unity, then one of two things can happen. The orbit may collide with another orbit of the same period in which case the two orbits could be annihilated in a saddle-node bifurcation or they could pass through each other. Secondly, there may be a new pair of orbits created (or annihilated); this is the Rimmer bifurcation (see section 2.4). It is commonly observed (i.e. co-dimension one) in the reversible case.

When $\mathcal{M}(-1) = 0$ the orbit has a pair of multipliers at -1. Upon variation of parameters such an orbit typically undergoes a period doubling bifurcation. These bifurcations have been studied for the four dimensional case by Mao and co-workers [Mao 1986, 1987].

The set $\mathcal{M}(\lambda) = 0$ for real p is a one parameter family of co-dimension one surfaces in the N dimensional stability parameter space. The envelope of this family, given by $\mathcal{M} = \mathcal{M}_p = 0$, is the co-dimension one complex instability surface, since these equations imply that \mathcal{M} has a double root. For the four dimensional case this surface is the parabola $B = A^2/4 + 2$.

On one side of the complex instability surface there is a complex quadruplet of λ 's. This boundary contains both the case of collision of two pairs of unit modulus multipliers (Krein collision), and of real multipliers (Fig. 2.3). Loss of stability by passing through this surface can occur only when the "Krein signature" of the two multiplier pairs on the unit circle is mixed [Moser 1958]. Conversely, when the signature is definite, such a collision

cannot lead to the complex instability: such an orbit is strongly stable under small perturbations. In the case of mixed signature, collision leading to the loss of stability typically occurs in one parameter families of maps, while for definite signature, collisions are less common (co-dimension three) [Howard 1986]. However, for the case of the “in-phase” orbits of a coupled quadratic map, definite signature collisions without splitting off the unit circle naturally occur in the successive stable regions of the period doubling bifurcation sequence [Mao 1987].

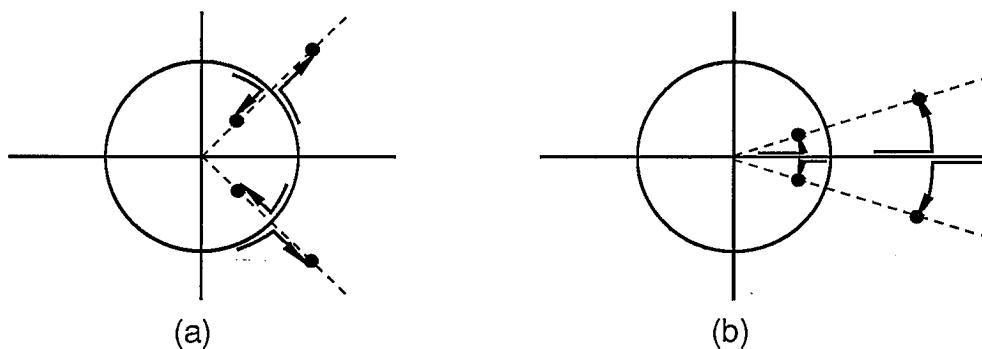


Fig. 2.3 Collision of multiplier pairs on complex plane. The multipliers of mixed Krein signatures split off to form a quadruplet after collision. (a) Collision on the unit circle causes the stability to change from EE to CQ . (b) Collision on the real axis causes a transition from HH to CQ for the positive axis and from II to CQ for the negative axis.

For example (2.1), there are four fixed points, all at $\mathbf{p} = 0$, and corresponding to $\mathbf{q} = \mathbf{m}/2$. Stability is easily determined analytically; for the

fixed point at $\mathbf{p} = \mathbf{q} = \mathbf{0}$, when $k_2 = 0$, the three curves in the A-B plane (see Figure 2.1) are mapped to the k_1 - h plane as follows:

$$\det M(-1) = (k_1 - 8)(h - 4) - 16 = 0$$

$$\det M(+1) = k_1 h = 0$$

$$B - A^2/4 - 2 = -(k_1^2 + 4h^2)/4 = 0$$

Figure 2.2(a) is the stability diagram when $k_2 = 0$ for the point $\mathbf{q} = \mathbf{0}$. The stability diagrams for the other three points are obtained from this by flipping the signs of k_1 and/or h :

| | | | | |
|--------------|------------|-------------|--------------|-------------|
| \mathbf{q} | $(0,0)$ | $(0,1/2)$ | $(1/2,0)$ | $(1/2,1/2)$ |
| signs | (k_1, h) | $(k_1, -h)$ | $(-k_1, -h)$ | $(-k_1, h)$ |

The stable region is labelled \mathcal{EE} . For any given parameters only one fixed point can be \mathcal{EE} ; when k_1 and h are positive, this is the point $\mathbf{q}=(0,0)$. In this case the point $(1/2, 0)$ is always hyperbolic. The other two fixed points have one hyperbolic direction (negative residue) and one direction which is either elliptic or inversion hyperbolic (positive residue). Complex instability does not occur since the signature of the stable region is definite.

There are also four period two orbits with frequency $(1,1,2)$. The stability of the period two orbits can be analytically determined for the two

orbits at $\mathbf{q} = (0,0) \rightarrow (1/2, 1/2)$ and at $(1/2, 0) \rightarrow (0, 1/2)$; for the first orbit, the three curves in the A-B plane are mapped to the k_1 -h plane as follows:

$$\det M(-1) = 16 (1-h)^2 - k_1^2 (2-h)^2 = 0$$

$$\det M(+1) = -k_1^2 h (h-4) = 0$$

$$B - A^2/4 - 2 = -k_1^4/4 + k_1^2 h^2 - 4 \{(h-1)^2 - 1\}^2 = 0$$

The diagram for the orbit at $\mathbf{q} = (0,0) \rightarrow (1/2, 1/2)$ is given by Figure 2.2 (b); for the orbit at $\mathbf{q} = (1/2, 0) \rightarrow (0, 1/2)$, reverse the signs of both k_1 and h . In region \mathcal{EE}_1 the multipliers have definite signatures, and there are no collisions except at $\lambda = \pm 1$, hence there is no complex instability boundary. However there are two more stable regions for the period 2 case (region \mathcal{EE}_2 and \mathcal{EE}_3 of (b)) for which the signatures are mixed, and for which there is a boundary leading to region \mathcal{CQ} . Note that a bifurcation to \mathcal{CQ} can not only occur from an \mathcal{EE} region, as sketched in Fig. 2.3 (a), but also from \mathcal{HH} , as in Fig. 2.3 (b), or \mathcal{II} .

In numerical computations for higher periods, we commonly observe (upon the variation of a single parameter) the collision of two multiplier pairs either on the unit circle or the real axis leading to the \mathcal{CQ} region. We will discuss this further below.

2.4. Reversibility

The flow generated by the simple Hamiltonian $H = 1/2p^2 + V(\mathbf{q})$ is time reversible: for every orbit $[\mathbf{p}(t), \mathbf{q}(t)]$ there is another orbit $[-\mathbf{p}(-t), \mathbf{q}(-t)]$, which is its time reverse. Time reversal is accomplished by reversing the momentum. In general a map is reversible if there is an anti-symplectic involution, S , which reverses time [DeVogelaere 1958, MacKay 1982]:

$$STS = T^{-1} \quad (2.24)$$

S is anti-symplectic if $\widetilde{DS} \omega DS = -\omega$, and an involution if $S^2 = I$. Equation (2.24) implies that the map can be factored as

$$T = (TS) S$$

showing that TS is also an anti-symplectic involution. Since it satisfies (2.24), it is also a reversor for T , called the complementary symmetry to S . In fact, each operator $T^k S$ is also a symmetry of T , with complementary symmetry $T^{k+1} S$.

A set invariant under the symmetry operator is called a symmetry set. Near a fixed point of S , symplectic coordinates (ξ, η) can always be found such that $S(\xi, \eta) = (-\xi, \eta)$ [Ginzberg 1988]. Thus the symmetry sets are N -dimensional submanifolds, which we will call symmetry planes. In fact they are Lagrangian manifolds, since if $\delta \mathbf{z}$ and $\delta \widehat{\mathbf{z}}$ are tangent vectors at a point \mathbf{z}_0 in $\text{Fix}(S)$, then $\delta \mathbf{z} = DS \delta \mathbf{z}$, and therefore $\widetilde{\delta \mathbf{z}} \widetilde{DS} \omega DS \delta \widehat{\mathbf{z}} = \widetilde{\delta \mathbf{z}} \omega \delta \widehat{\mathbf{z}}$; the anti-symplectic condition implies this is equal to its negative and therefore $\widetilde{\delta \mathbf{z}} \omega \delta \widehat{\mathbf{z}}$

= 0. We will assume that these manifolds are graphs over the momentum plane.

As an example, the map generated by eq. (2.16) is easily shown to be reversible if $V(\mathbf{q}) = V(-\mathbf{q})$; a pair of symmetry operators is

$$\begin{aligned} S: & \begin{cases} \mathbf{p}' = \mathbf{p} - \nabla V(\mathbf{q}) \\ \mathbf{q}' = -\mathbf{q} \end{cases} \\ TS: & \begin{cases} \mathbf{p}' = \mathbf{p} \\ \mathbf{q}' = -\mathbf{q} + \mathbf{b}^{-1} \cdot \mathbf{p} \end{cases} \end{aligned} \quad (2.25)$$

The fixed sets of these symmetries are the N-planes, $\text{Fix}(S) = \{\mathbf{q} = \mathbf{0}\}$, and $\text{Fix}(TS) = \{\mathbf{p} = 2\mathbf{b} \cdot \mathbf{q}\}$. This applies to the Froeshle example (2.1).

Due to the periodicity of T, there are other symmetries obtained from the above by translation: the translation of a symmetry by an integer vector \mathbf{m} , the operator $SR_{\mathbf{m}}$, is also a symmetry. It is easy to see that upon translation the fixed sets are translated by half integral vectors; for the example (2.25) this gives

$$\begin{aligned} \text{Fix}(SR_{\mathbf{m}}) &= \{ (\mathbf{p}, \mathbf{q}) \mid \mathbf{q} = 1/2\mathbf{m} \} \\ \text{Fix}(TSR_{\mathbf{m}}) &= \{ (\mathbf{p}, \mathbf{q}) \mid \mathbf{p} = 2\mathbf{b} \cdot (\mathbf{q} - 1/2\mathbf{m}) \} \end{aligned} \quad (2.26)$$

These fixed sets are indeed Lagrangian manifolds since \mathbf{b} is symmetric.

On the torus, there are 2^N symmetry planes for S, obtained by letting the elements of \mathbf{m} be either 0 or 1, and the same number generated by TS.

Of course each symmetry family $(T_i S, T_{i+1} S)$ has such a set of symmetry planes.

There exist very useful commutation relations between operations T , S and R (translation) as follows:

$$RT = TR$$

$$SR = R^{-1}S$$

$$ST = T^{-1}S$$

Symmetric orbits are those invariant with respect to time reversal, which means that for the symmetry S and an orbit $\{z_t\}$ each point z_i has a symmetric counterpart: there exists a j such that $z_j = Sz_i$.

The importance of symmetric orbits arises from three facts. Firstly, it is easy to see that every symmetric orbit must have a point on a symmetry plane [MacKay 1982]. If $z_j = Sz_i$ and $k = j - i$ is even, then Eq. (2.24) implies

$$z_{i+k/2} = T^{-k/2} z_j = T^{-k/2} S z_i = S T^{k/2} z_i$$

or that $z_{i+k/2}$ is a fixed point of S . On the other hand, if k is odd, then $z_{i+(k+1)/2}$ is a fixed point of TS . This greatly simplifies finding symmetric orbits by limiting the search to the symmetry planes.

Secondly, each symmetric periodic orbit of frequency (m, n) has a second point on another symmetry plane. Suppose z_i is a fixed point of some symmetry, S' . Eq. (2.24) implies

$$z_{i+j} = S' z_{i-j} \tag{2.27}$$

Applying the periodicity condition, Eq. (2.14), to this yields

$$\begin{aligned}
\mathbf{z}_{i+j} &= R_{-m} T^{-n} S' \mathbf{z}_{i-j} = R_{-m} S' \mathbf{z}_{n+i-j} \\
&= S'R_m \mathbf{z}_{n+i-j}
\end{aligned} \tag{2.28}$$

Now pick $j = k \equiv [(n+1)/2]$, where $[]$ represents integer part. Equation (2.28) implies that \mathbf{z}_{i+k} is a fixed point of $S'R_m$ for n even and of $TS'R_m$ for n odd.

Conversely, suppose we have an orbit segment $\{\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k\}$ such that $\mathbf{z}_0 \in \text{Fix}(S')$, and $\mathbf{z}_k \in \text{Fix}(S'R_m)$ for some symmetry S' . A periodic orbit of period $n=2k$ is obtained by using Eq. (2.28) with $i=0$ to define the \mathbf{z}_j for $j \in [k+1, n]$. This implies that $\mathbf{z}_n = S'R_m \mathbf{z}_0 = R_{-m} \mathbf{z}_0$, so that the orbit is indeed periodic and has frequency (m, n) . A similar construction applies to the odd period case, for which $\mathbf{z}_k \in \text{Fix}(TS'R_m)$, and $n = 2k+1$. Thus a symmetric orbit of period n can be obtained by searching for a segment of length k which has end points on two symmetry planes (see section 3.3).

The final reason for the studying symmetric orbits is that every invariant torus which is homotopic to $\mathbf{p}=\mathbf{0}$ and which has a dense orbit, must be symmetric, since it must intersect the symmetry planes. Thus it seems reasonable to use symmetric periodic orbits to approximate quasiperiodic orbits.

Classification of the symmetric orbits is attained by considering the the phase space $\mathcal{R}^N \times \mathcal{T}^N$ where there are 2^{N+1} distinct symmetry planes generated by S and TS . Each symmetric orbit as a point on two of these planes, and so there are 2^N types (with one exception, see below). Using the above relation, the classes can be easily obtained for each (m, n) ; they

Table 2.1 Symmetry planes for the four symmetric orbits with frequency (m_1, m_2, n) of a 4D reversible map. There are eight classes of frequencies depending on whether the frequency components are even "e" or odd "o". Each orbit has its first point on the plane in the column $(\mathbf{q}_0, \mathbf{p}_0)$ and its k^{th} point on the plane in the column $(\mathbf{q}_k, \mathbf{p}_k)$. Here $k = \lfloor (n+1)/2 \rfloor$.

| (m_1, m_2, n) | $(\mathbf{q}_0, \mathbf{p}_0)$ | $(\mathbf{q}_k, \mathbf{p}_k)$ | | | |
|-----------------|--------------------------------|--------------------------------|---------|----------------------|----------------------|
| (o,o,e) | S | SR _(1,1) | (o,e,o) | S | TSR _(1,0) |
| | SR _(1,0) | SR _(0,1) | | SR _(1,0) | TS |
| | TS | TSR _(1,1) | | SR _(0,1) | TSR _(1,1) |
| | TSR _(1,0) | TSR _(0,1) | | SR _(1,1) | TSR _(0,1) |
| (e,o,e) | S | SR _(0,1) | (e,e,o) | S | TS |
| | SR _(1,0) | SR _(1,1) | | SR _(1,0) | TSR _(1,0) |
| | TS | TSR _(0,1) | | SR _(0,1) | TSR _(0,1) |
| | TSR _(1,0) | TSR _(1,1) | | SR _(1,1) | TSR _(1,1) |
| (o,e,e) | S | SR _(1,0) | (e,e,e) | S | S |
| | SR _(0,1) | SR _(1,1) | | SR _(1,0) | SR _(1,0) |
| | TS | TSR _(1,0) | | SR _(0,1) | SR _(0,1) |
| | TSR _(0,1) | TSR _(1,1) | | SR _(1,1) | SR _(1,1) |
| (o,o,o) | S | TSR _(1,1) | | TS | TS |
| | SR _(1,0) | TSR _(0,1) | | TSR _(1,0) | TSR _(1,0) |
| | SR _(0,1) | TSR _(1,0) | | TSR _(0,1) | TSR _(0,1) |
| | SR _(1,1) | TS | | TSR _(1,1) | TSR _(1,1) |
| (e,o,o) | S | TSR _(0,1) | | | |
| | SR _(1,0) | TSR _(1,1) | | | |
| | SR _(0,1) | TS | | | |
| | SR _(1,1) | TSR _(1,0) | | | |

depend only on the evenness or oddness of the components of the frequency. In Table 2.1 the classification is given for $N=2$ where there are eight frequency types. For example when the frequency is of the type (even,odd,odd), then the symmetric orbit with a point on the $\text{Fix}(S)$ plane, also has a point on the $\text{Fix}(\text{TSR}_{(0,1)})$ plane. These two symmetric points are

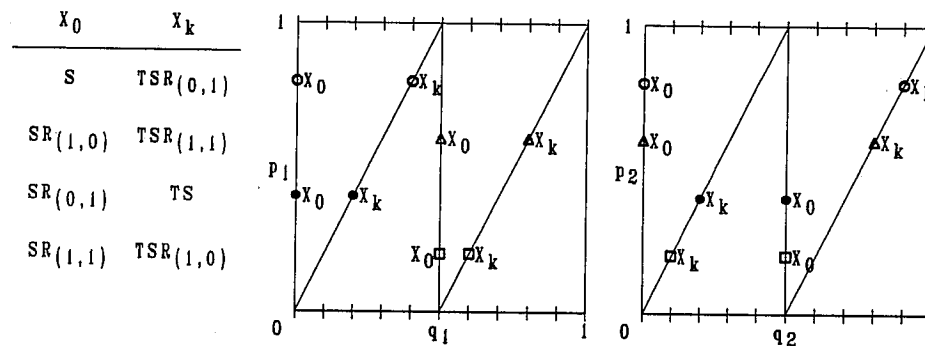


Fig. 2.4 Symmetry planes projected onto the canonical p_1 - q_1 and p_2 - q_2 planes. For example, $\text{Fix}(\text{TSR}_{(1,0)})$ is projected onto $p_1=2q_1-1$ on the p_1 - q_1 plane and $p_2=2q_2$ on the p_2 - q_2 plane. The symmetric points for the four cases of an orbit with frequency type (e,o,o) are shown. Each orbit intersects two symmetric planes. The orbit on $\text{Fix}(S)$ has one point (z_0) on $\text{Fix}(S)$ and its k^{th} iterate (z_k) on $\text{Fix}(\text{TSR}_{(0,1)})$.

indicated by open circles on each canonical plane in Fig. 2.4. The other three symmetric orbits are also shown in the figure. The case (even,even,even) in Table 2.1 refers to an orbit which has a minimal period which is even, and yet the elements of \mathbf{m} are also even; such orbits arise, for example, from period doubling bifurcations.

In Appendix B we show that there will be at least one orbit of each symmetry class for each frequency. Therefore a reversible, symplectic twist map has at least 2^N symmetric periodic orbits for each frequency (\mathbf{m},n) .

For the four dimensional example (2.1) we routinely find all four symmetric periodic orbits for any given frequency. The simplest technique is to begin with a point in the first symmetry plane in Table 2.1 and vary \mathbf{p} until the orbit returns to the second symmetry plane in Table 2.1 after k iterations. This requires a 2D root finding routine; we use Broyden's method (see section 3.3), which is a generalization of the secant method. This method fails to converge when the parameters (k_1, k_2, h) are large because the orbits are strongly unstable, but reasonably large parameters can be obtained by using extrapolation from small parameter values to give good initial guesses.

A more sophisticated technique is based on Newton's method for extremizing the action W_0 . We use this method when the orbits are strongly unstable; however, it also requires good initial guesses. The details of this method are discussed in section 3.4.

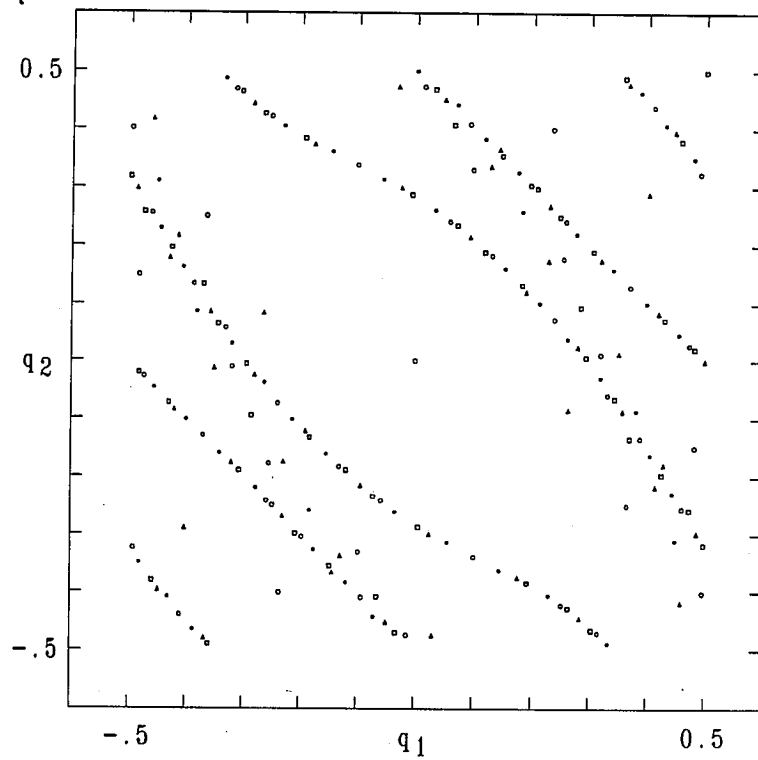


Fig. 2.5 Configuration space projection of the four symmetric orbits with frequency (12,21,49) when $k_1=0.4$, $k_2=0.3$, $h=0.65$. Each orbit is represented with symbols; “o” for the orbit on $\text{Fix}(S)$, “•” on $\text{Fix}(SR_{(0,1)})$, “Δ” on $\text{Fix}(SR_{(1,0)})$ and “□” on $\text{Fix}(SR_{(1,1)})$.

Figure 2.5 displays a projection of the four symmetric orbits of frequency (12,21,49) onto the \mathbf{q} -plane. Note that, except for the point at $\mathbf{q} = (0,0)$, the orbits tend to avoid the region near the origin. This phenomena will be explained in terms of the action function in the next section.

Typically, when the parameters (k_1, k_2, h) are not too large, one of the symmetric orbits is of type EE , another is HH , and the remaining two are EH (we observed this already for period one and two in section 2.3). This is the analogue of the familiar island chain formed by an elliptic-hyperbolic pair for the area preserving map. As the parameters increase various bifurcations can occur as shown in Fig. 2.6. For example in Fig. 2.6(a), the (2,3,4) orbit on $\text{Fix}(S)$ first undergoes a period doubling bifurcation, then regains stability by its inverse (thus the period eight orbit only exists for a limited range, forming a “bubble”), only to lose it again by a complex bifurcation. The multipliers coalesce on the real axis in an inverse Krein collision. The resulting HH orbit then undergoes a symmetry breaking bifurcation to EH when a pair of multipliers passes through $+1$ [Rimmer 1978]. When this occurs a pair of non-symmetric orbits with the the same period is created, as shown in Fig. 2.7 (b), and the residue corresponding to the original orbit changes sign. There is a corresponding change in sign of $\mathcal{M}(1)$. Finally the orbit reaches the HI region. Similarly the orbit on $\text{Fix}(SR_{(1,0)})$ undergoes the normal symmetry breaking bifurcation shown in Fig. 7(a). Spiraling in the A-B plane like that shown in Fig. 2.6 is common, and becomes more rapid for longer period orbits.

Symmetry breaking bifurcations of the ordered periodic orbits do not appear to occur for the standard map (the minimizing orbit always appears to be symmetric), but can occur for a more generally, for example, when the potential is $V(q) = a_1 \cos(2\pi q) + a_2 \cos(4\pi q)$ and a_2 is large enough.

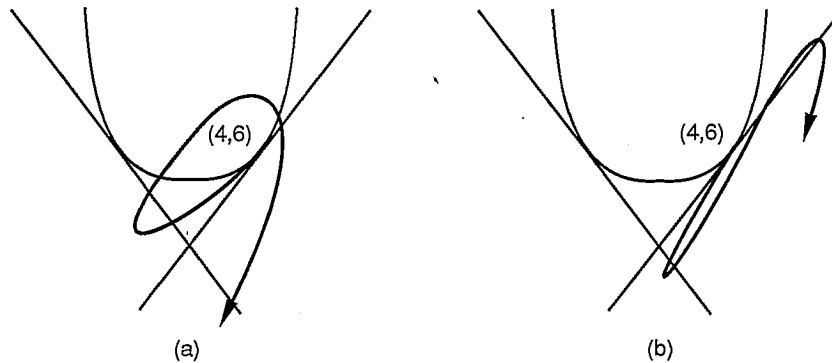


Fig. 2.6 Sketch of the evolution of the stability of two symmetric orbits with $\omega = (2,3,4)$ (a) on the $\text{Fix}(S)$ and (b) on the $\text{Fix}(SR_{(1,0)})$ plane as the parameters vary on the path $k_1=10h$, $k_2=0$. Bifurcations occur each time a stability curve is crossed.

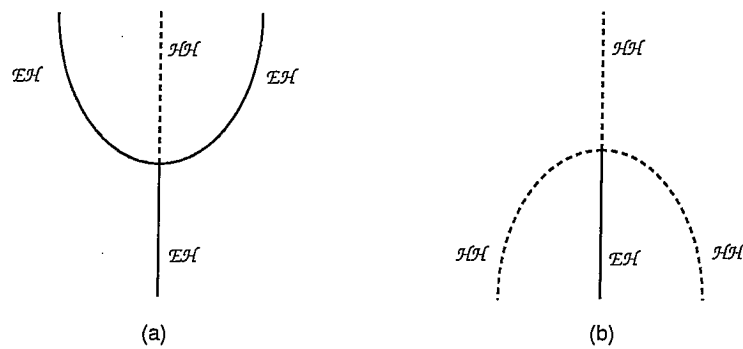


Fig. 2.7 Symmetry breaking bifurcations. Bifurcating orbits are of the same frequency as their progenitor. The example shown is for the transition between $\mathcal{E}\mathcal{H}$ and $\mathcal{H}\mathcal{H}$. (a) direct process. (b) inverse process.

2.5. Minimizing Orbits

For area preserving twist maps, existence theorems for periodic and quasiperiodic orbits are based to a large extent on the properties of the action. Aubry and Mather prove, for example, that an area preserving twist map has a minimizing orbit for each (m,n) [Mather 1982, Aubry 1983]. An orbit is defined to be minimizing if the action of every finite segment of the orbit is not larger than that of any variation with fixed end points:

$$W(q_j, \dots, q_k) \leq W(q_j, q_{j+1} + \Delta q_{j+1}, \dots, q_{k-1} + \Delta q_{k-1}, q_k)$$

for all $j < k$. Aubry shows that an (m,n) minimizing configuration is ordered on the circle in the same way as a simple rotation with frequency m/n . This implies that there also exists a second periodic orbit of frequency (m,n) because if the orbit of q_0 is minimizing, then so is that of $T^t q_0$. If we choose t so that $T^t q_0$ is the nearest point on the orbit to the right of q_0 , then between these neighbors there occurs a point which is minimum in all directions except one; this is the minimax configuration.

For $N > 1$ the straightforward generalization of Aubry's minimizing condition does not appear to be fruitful. We use a weaker definition: a periodic orbit $\{ \mathbf{q}_t \}$ is a minimizing if its periodic action (2.15) is not larger than any other periodic configuration

$$W_\omega(\mathbf{q}_t) \leq W_\omega(\mathbf{q}_t + \Delta \mathbf{q}_t) \tag{2.29}$$

This does not in general imply that the orbit is minimizing in Aubry's sense: a counter-example is given by Hedlund for geodesics on a three torus

[Hedlund 1932]. He shows that a geodesic of given winding number ω which minimizes the analogue of W_ω , may not be a minimum of the periodic action for some multiple of the winding number: $W_{j\omega}$. In this case minimizing orbits (in Aubry's sense) do not exist for each ω .

A periodic orbit is locally minimizing if the second variation of the periodic action about the orbit is positive definite. This latter concept is all that we can check numerically.

The Morse index of the action is defined as the number of convex directions at its extremum, i.e. the number of the negative eigenvalues of the quadratic form $\delta^2 W_\omega$ [Milnor 1963]. A minimizing orbit has index zero.

In Appendix B we show that there exist minimizing periodic orbits for each frequency. This result follows from a bound on the growth of the generating function with $|\mathbf{q} - \mathbf{q}'|$ entirely analogous to the area preserving case. Furthermore, there exist orbits with non-zero indices as well (also in Appendix B). The minimum of the periodic action for a given value of \mathbf{q}_0 exists, and gives a function on the N -torus. Each critical point of this function gives a periodic orbit. Generically the critical points of such a function are non-degenerate, and in this case Morse theory implies that there are at least 2^N . Thus there are typically at least 2^N orbits for each frequency (\mathbf{m}, n) . The indices of these orbits range from 0 to N , and there will typically be at least $N! / i! (N-i)!$ orbits with index i .

There is a close relationship between the stability matrix (2.22) and the second variation of the action (2.15) about the periodic orbit:

$$\delta^2 W_\omega \Big|_{\delta \mathbf{Q}} = \frac{1}{2} \delta \tilde{\mathbf{Q}} \mathbf{M}(1) \delta \mathbf{Q} \quad (2.30)$$

where $\delta\mathbf{Q}$ represents an arbitrary variation. Thus $\mathbf{M}(1)$ is the Hessian matrix corresponding to the quadratic form δ^2W_ω . For the area preserving case the stability type of a periodic orbit is characterized solely by the determinant of $\mathbf{M}(1)$ [MacKay 1983], however when $N>1$, this is no longer true. We can combine (2.23) and (2.30) to obtain a relation which generalizes the area preserving case:

$$\prod_{i=1}^N R_i = \left(-\frac{1}{4}\right)^N \frac{\mathcal{M}(1)}{\prod_t \det \mathbf{b}_t} \quad (2.31)$$

where the R_i are the residues (2.18). Since \mathbf{b} is assumed positive definite, Eq. (2.31) gives a relationship between the sign of the product of the residues, and of the determinant of the second variation of the action:

If the index is odd (even) then there must be an odd (even) number of residues which are real and positive.

For the area preserving case, a non-degenerate minimizing orbit must be regular hyperbolic. For four dimensions a minimizing orbit falls in the region $B>2A-2$, i.e. above the saddle node line in Fig. 2.1; this does not completely determine its stability type, excluding only the cases \mathcal{EH} and \mathcal{HI} .

We often find that as the parameters (k_1, k_2, h) are increased from zero one of the symmetric orbits is minimizing and of type \mathcal{HH} . This orbit can undergo a symmetry breaking bifurcation, however, and its index will become one. The action as a function of a configuration variable will

undergo an evolution like that shown in Fig. 2.8. The two newly created non-symmetric orbits are now minimizing, and of type \mathcal{HH} , as indicated in Fig. 2.7(b). Similarly, the non-zero index, symmetric orbits can also spawn non-symmetric orbits.

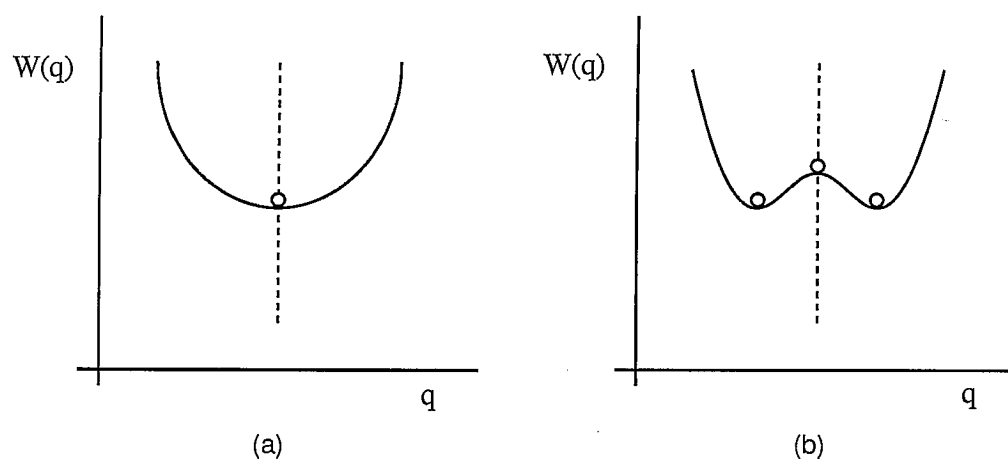


Fig. 2.8 Inflection of the action function upon a symmetry breaking bifurcation. Here “ q ” represents one of the configuration coordinates of the orbit, and the dotted line represents the symmetry plane. Prior to bifurcation (a), the orbit minimizes the action and its index is zero. After the bifurcation (b), the action acquires one convex direction at the extremum point on the symmetry plane. The index of the symmetric orbit becomes one, and the two new minimizing orbits are not symmetric.

The symmetry plane which goes through the maximum of $F(\mathbf{q}, \mathbf{q})$ plays a special role, and we will call this plane the dominant symmetry plane [MacKay 1982]. In particular the index N fixed point occurs on the dominant symmetry plane. Furthermore $F(\mathbf{q}, \mathbf{q})$ has critical points on each of the other symmetry planes $\text{Fix}(\text{SR}_{\mathbf{m}})$, and so there are fixed points on each. In

general there may also be other critical points of F , which give non-symmetric fixed points.

Now consider periodic orbits of other frequencies ω . For large enough potential $V(\mathbf{q})$, a minimizing orbit should have most of its points near the minimum of $-V$, since the potential dominates the action, Eq. (2.16). An orbit with dominant symmetry, however, must have a point on the plane where $-V$ is maximum, and thus ought to have an action larger than those without points on this plane. It certainly seems reasonable that the minimizing orbit will not have a point on the dominant symmetry plane, and we conjecture that the minimizing orbit never has the dominant symmetry. For the case $N=1$ this conjecture is supported by extensive numerical evidence [MacKay 1982, MacKay 1984].

Our scans of parameter space for the example (2.1) have never given a minimizing orbit on the dominant plane. For most cases the index two orbit has the dominant symmetry for small parameter values; however, as the parameters increase the orbit on the dominant plane can undergo a symmetry breaking bifurcation, and then will no longer have index 2. We also see, as in Fig. 2.5, that the minimizing and index one orbits tend to avoid the dominant plane, leaving a hole inside of which sits a single point of the index two orbit. As the parameters are increased this hole grows in size until it encircles the torus producing a stripe which is seen in Fig. 2.5. We suspect that a similar phenomena for quasiperiodic orbits is responsible for the destruction of an invariant torus.

2.6. Resonance Zones, Channels, and Chaotic Layers of Four-dimensional Maps

The symmetry and the self-similar structure inherited from the binary construction of the generalized Farey tree, introduced in section 1.5, is shown in Fig. 2.9. The figure displays the rationals $\omega=(\omega_1,\omega_2)$ in the range $[-1/2,1/2] \times [-1/2,1/2]$ produced by the generalized Farey tree algorithm up to level 13. On the ω plane each triangle is not actually right isosceles; for example, one of the level two triangles has vertices $(0/1,0/1)$, $(1/2,0/2)$ and $(1/3,1/3)$. Notice that the sampling is extremely non-uniform, with sparse regions around every rational pair. These of course will be densely filled by higher level rationals, and in the limit of infinite level, by incommensurate pairs.

Our numerical results suggest the close contact of the generalized Farey tree construction to the dynamics of the maps given by Eq (2.1). The points in Fig. 2.9 are equivalent to the intersections of periodic orbits with frequency ω on a constant angle plane for the integrable map: for example (1) with $k_1= k_2= h= 0$, a periodic orbit with frequency ω occurs at $\mathbf{p}=\omega$. This symmetry is broken as soon as the parameters are non-zero, as shown in Fig. 2.10 for the same set of periodic orbits. The orbits are obtained by using Broyden's method (see section 3.3), and are all symmetric since they are on the symmetry planes: (a) on the $\text{Fix}(S)$ plane, (b) on the $\text{Fix}(SR_{(1,1)})$ plane, and (c) on the $\text{Fix}(SR_{(0,1)})$ plane. Now the sparse regions are larger, and are not filled in by higher level orbits.

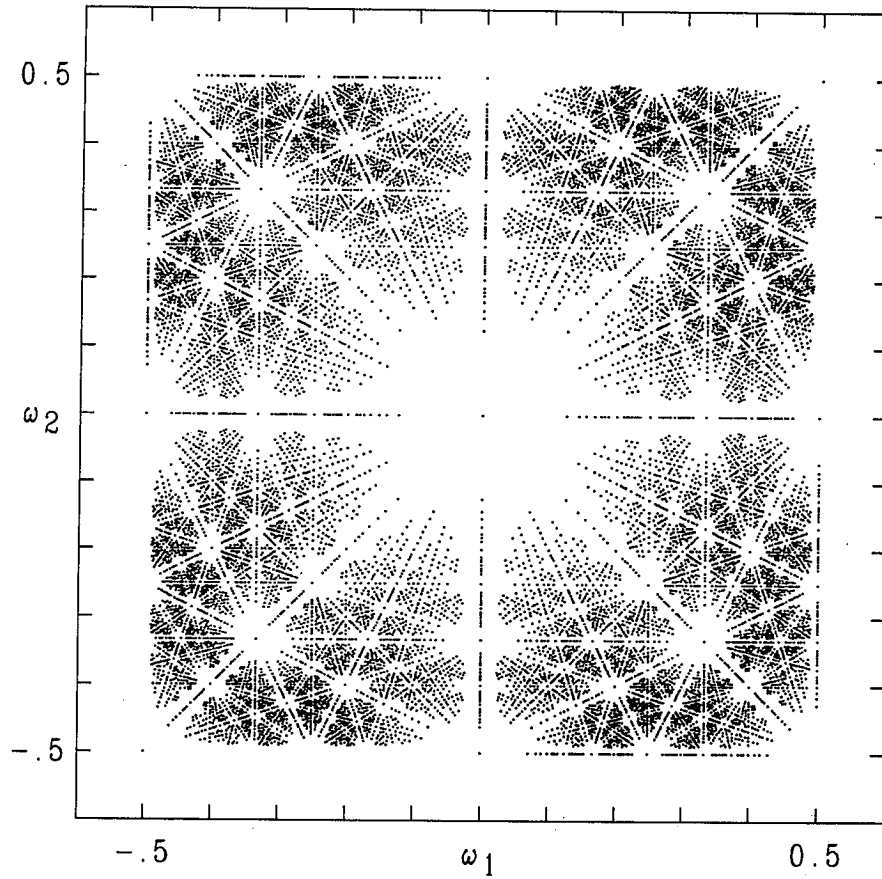


Fig. 2.9 Pairs of rationals (ω_1, ω_2) generated by the Kim-Ostlund tree up to level 13. Equivalently, each point represents the intersection point \mathbf{p} of a periodic orbit with frequency $\mathbf{p} = (\omega_1, \omega_2)$ on a plane $\mathbf{q} = \text{constant}$ for the integrable system.

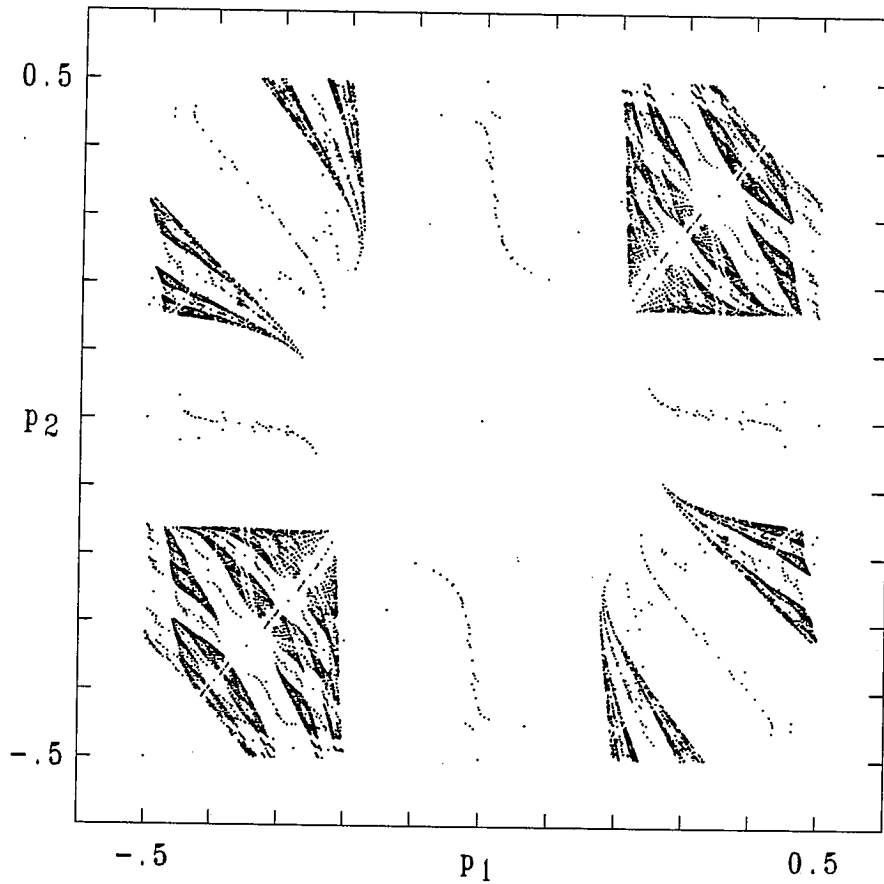


Fig. 2.10(a) Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5$, $k_2=0.3$, $h=0.2$, on $\mathbf{q} = (0,0)$ plane. Higher period orbits are expelled from the vicinity of shorter period orbits creating resonance regions. Furthermore, commensurability lines are expanded into channels which connect the resonances.

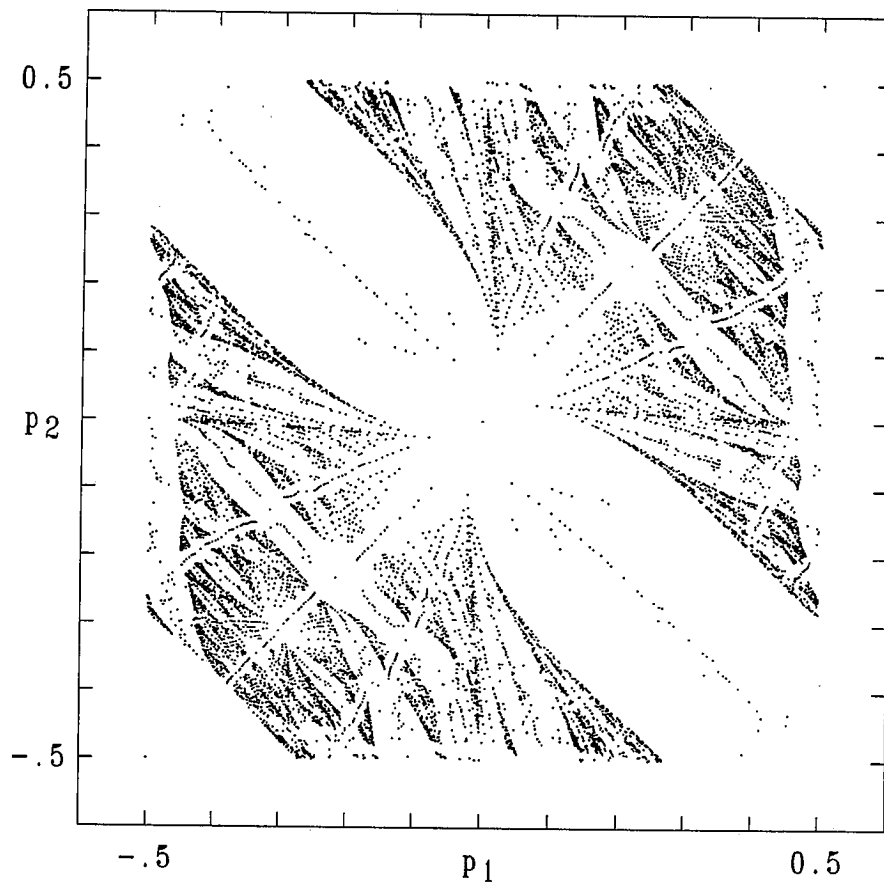


Fig. 2.10(b) Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5$, $k_2=0.3$, $h=0.2$, on $\mathbf{q} = (1/2, 1/2)$ plane.

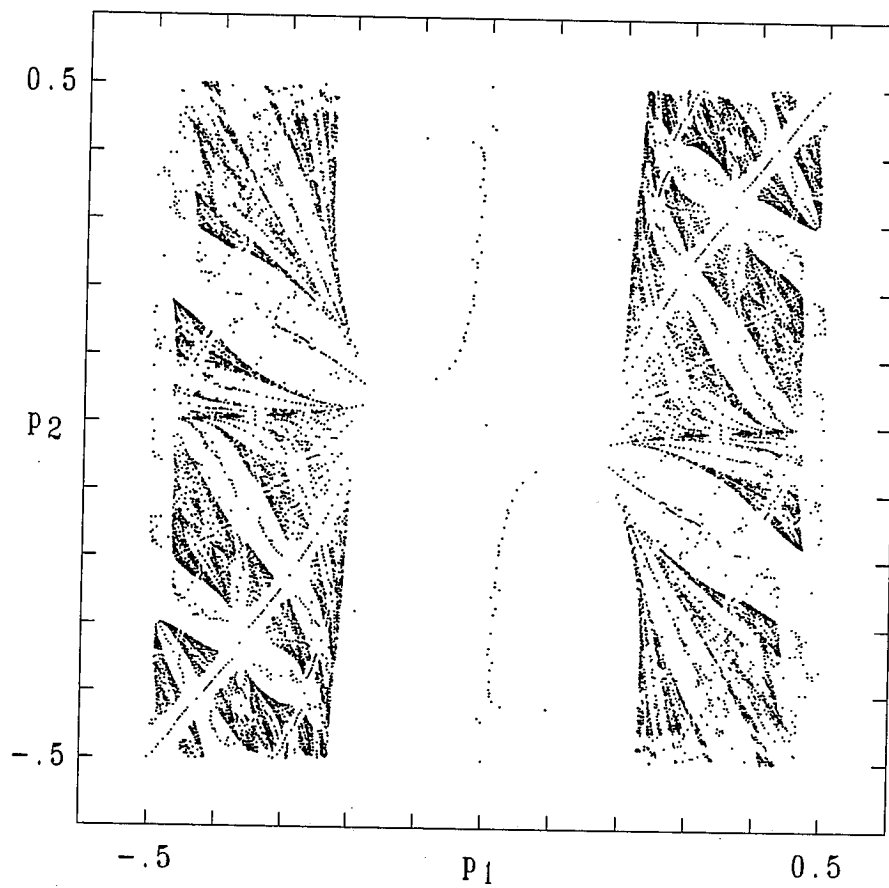


Fig. 2.10(c) Periodic orbits with the frequencies in Fig. 2.9 for $k_1=0.5$, $k_2=0.3$, $h=0.2$, on $\mathbf{q} = (0, 1/2)$ plane.

The structure is analogous to that for area preserving maps. Figure 2.10(a) corresponds to the dominant symmetry line of the standard map. Every ordered minimax orbit of the standard map seems to have a point on this line, and each corresponds to the center of a resonance [MacKay 1987]. There are no other ordered orbits for some interval around the minimax periodic orbit. These intervals are bounded by the minimax orbits which are homoclinic to the minimizing orbit of the resonance. Sequences of symmetric periodic orbits which limit to a rational, for example the sequences $\pm 1/n$, limit to homoclinic orbits as $n \rightarrow \infty$, and not to the minimax periodic orbit. In the full two dimensional phase space, we define a resonance as an area bounded by segments of the stable and unstable manifolds of the minimizing orbit (see section 4.2). Thus the structure on the symmetry line reflects this partition of phase space. If there are rotational invariant tori, they must exist outside the vacant intervals on the symmetry line.

For the four dimensional mapping, higher period orbits also tend to be expelled from the vicinity of shorter periodic orbits on the symmetry plane to form resonance zones around each periodic orbit. Orbits in the resonance zone tend to remain trapped in the zone for long periods. The resonances occupy four dimensional volumes, which cut the symmetry plane in two dimensional regions. In Fig. 2.10, we associate the empty region surrounding the $(0,0,1)$ orbit, with the $(0,0,1) \rightarrow \infty$ resonance. Many other such regions are visible in the figure. The boundaries of the resonances, which should be three dimensional, are not easy to define; they do not consist

solely of stable and unstable manifolds of the $\mathcal{H}\mathcal{H}$ orbit, since these are merely two dimensional (see section 4.4), but appear to include many other orbits, as we will discuss below.

There appear to be channels which connect the resonances, each corresponds to a commensurability relation, $\omega_1 j_1 + \omega_2 j_2 + j_3 = 0$, or equivalently $\omega \cdot \mathbf{j} = 0$ for the representation $\omega = (m_1, m_2, n)$. For example the horizontal and vertical channels corresponding to $\mathbf{j} = (0, 1, 0)$ and $(1, 0, 0)$ have widths controlled by the parameters k_2 and k_1 , respectively. Similarly, the major diagonal channel, with $\mathbf{j} = (1, 1, 0)$, corresponds to the coupling term, h , in Eq. (2.1). Other channels are generated by higher order commensurabilities, and their widths could be obtained from perturbation theory [Chirikov 1979]. Along the channels there are an infinity of resonances which satisfy the commensurability relation; in fact, as we will see below the channels seem to be composed of these resonances.

Orbits initially in a resonance zone may eventually leave the zone, even if they are started arbitrarily close to the central $\mathcal{E}\mathcal{E}$ orbit; this is the process called Arnol'd diffusion [Lichtenberg 1983, Chirikov 1979]. Similarly an orbit in the chaotic region outside a resonance can wander into resonances since invariant tori do not separate the phase space. The transitions from resonance to resonance occur along the commensurability channels, which form the so called Arnol'd web.

However, central portions of the resonances are difficult to enter. This is shown in Fig. 2.11, which is a thin slice of the phase space near the

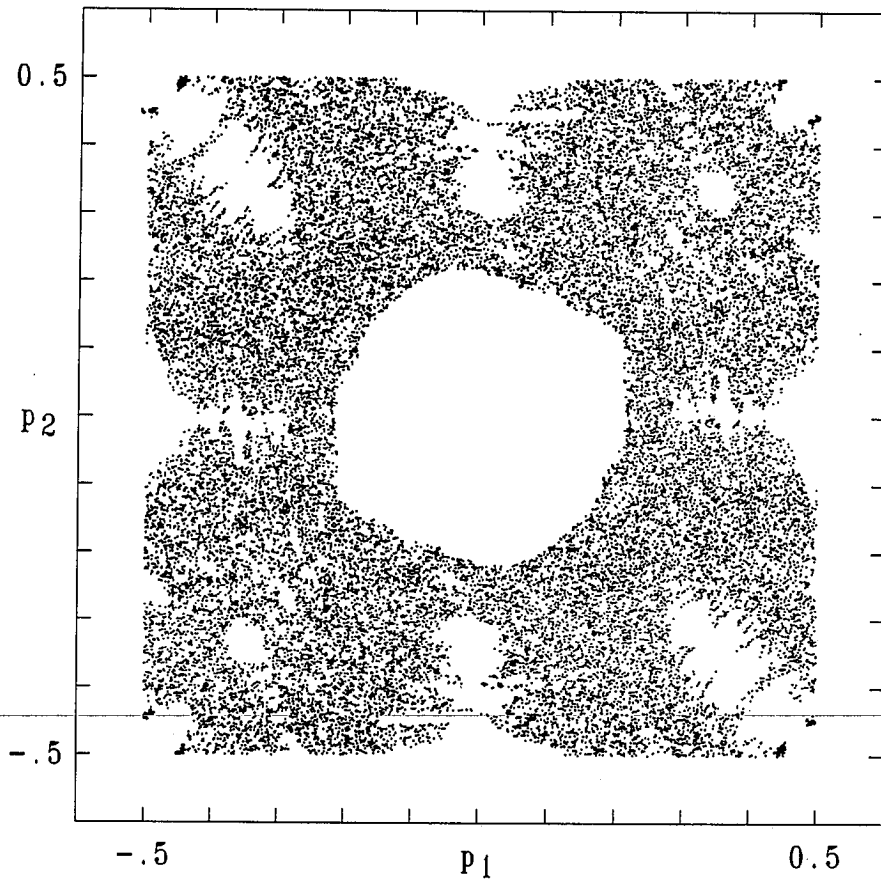


Fig. 2.11 Sectional view of two chaotic orbits on the $q = 0$ plane. Parameters are the same as in Fig. 2.10. The initial points are $(p_1, p_2, q_1, q_2) = \pm(0.10, 0.34, 0.0, 0.0)$ and the orbit is plotted whenever it falls inside the slice; $-.005 < q_1, q_2 < .005$ up to a time 2×10^8 .

dominant symmetry plane. Two initial conditions were iterated for 2×10^8 iterations, and only those points which fall in the slice are plotted (an equivalent figure is given by Kaneko [Kaneko 1985]). Many of the resonance zones shown in Fig. 2.10 are inaccessible on this time scale. The inaccessible regions appear to have fairly sharp, but irregular boundaries. On these there are many sticky points, corresponding to dark spots on the figure: these points correspond to invariant tori trapped in the resonance. Figure 2.12 shows one such torus in the primary resonance. There are many irrational tori around each resonance; however, they do not form a complete barrier: transition into a resonance occurs through the gaps between tori.

To attempt to delineate the resonance boundaries more precisely we could consider the homoclinic orbits. However, there appear to be only a few (perhaps four) homoclinic orbits on the dominant symmetry plane. For the $(0,0,1)$ resonance these correspond, to the limits of the sequences (m_1, m_2, n) with $m_i = \pm 1$ and $n \rightarrow \infty$. In Fig. 2.10 these sequences are visible as the traces down the middle of the diagonal channels. Other sequences of periodic orbits, for example those in the vertical and horizontal channels given by $(0, \pm 1, n)$ or $(\pm 1, 0, n)$, seem to limit to these four primary homoclinic orbits, or sums of these orbits.

A frequency analysis can also be used to find the resonance boundaries. All orbits trapped forever inside a resonance will have the same frequency, Eq. (2.13), as the resonance. So we can approximately delineate the resonance region by computing the frequency for some finite period of time; let ρ be this finite time frequency. Since, as Fig. 2.11 shows, Arnol'd

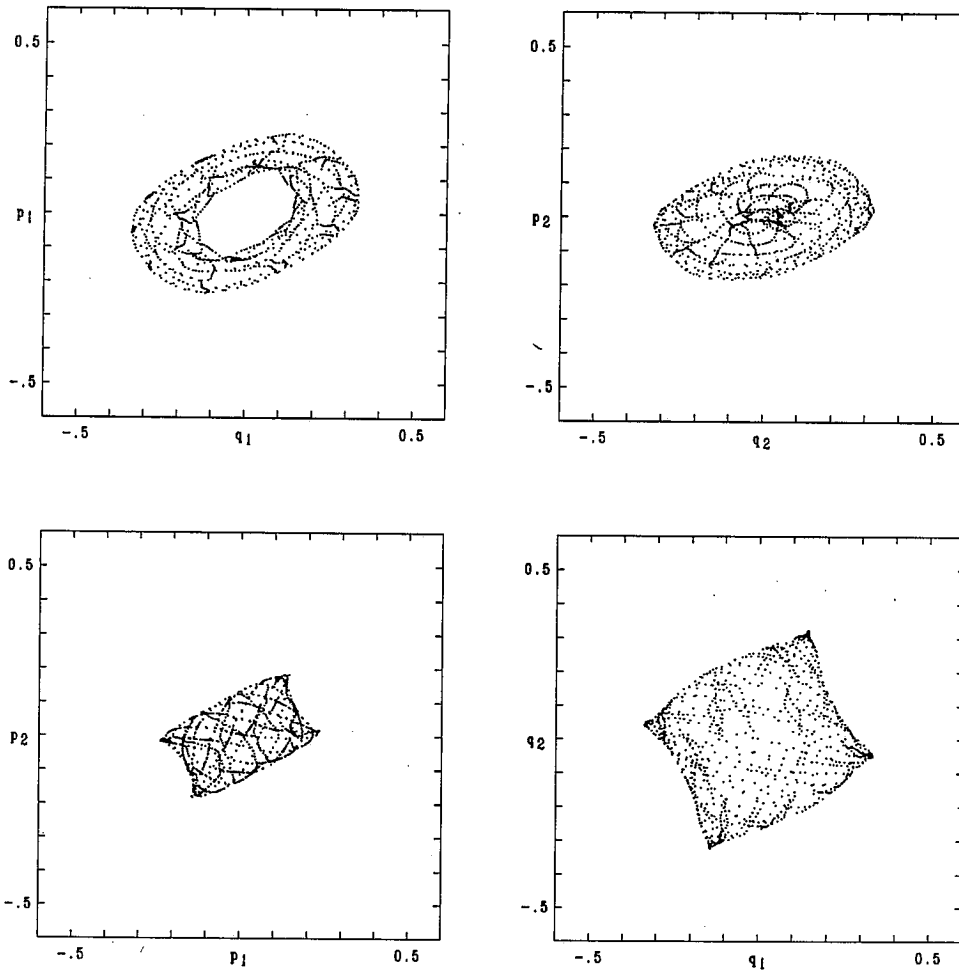


Fig. 2.12 Projections of one of the irrational tori surrounding resonance zones with frequency $(0,0)$. The initial point, $(0.1271, 0.1672, 0.0, 0.0)$ is selected from one of the stuck points at the edge of the primary resonance in Fig. 2.11.

diffusion is quite slow from the resonances, we can expect to obtain reasonable results even for modestly long times; we choose a time interval so that it would be shorter than an Arnol'd diffusion time scale but long enough to give statistical average. Typically we use $t=10^3$ with the parameter values we present. Figure 2.13 shows ρ_1 calculated for a set of initial conditions on the line $p_1=-p_2$ and $q_1=q_2=0$. A similar figure is obtained for ρ_2 , except with negative slope. These initial conditions lie in the diagonal channel corresponding to $\mathbf{j}=(1,1,0)$ which is seen in Figs. 2.10(a) and 2.11. There are flat steps in $\rho(\mathbf{p})$ corresponding to resonances of the form $(m,-m,n)$ showing that the channel is actually a collection of many resonances stacked closely together. The scattering of points near the edges of each resonance correspond to highly chaotic orbits.

Similarly, Figure 2.14 shows regions which have frequencies corresponding to low order resonances; it is a complementary view to that in Fig. 2.11. To construct Fig. 2.14 we consider initial conditions on a grid on the $\mathbf{q}=0$ plane. Each is iterated 10^3 times, and points are filled in only if $|\rho(\mathbf{p}) - \omega| < \epsilon$ for ω on the generalized Farey tree up to level five. The size of each resonance area on this plane depends weakly on the time scale we use to define the rotation number, but they correspond closely to the avoided regions in Fig. 2.11. The boundary of a resonance is not smooth, due to the chaotic orbits near its edge. The closer an orbit is to the boundary, the faster it escapes out of the resonance [Kueny 1989].

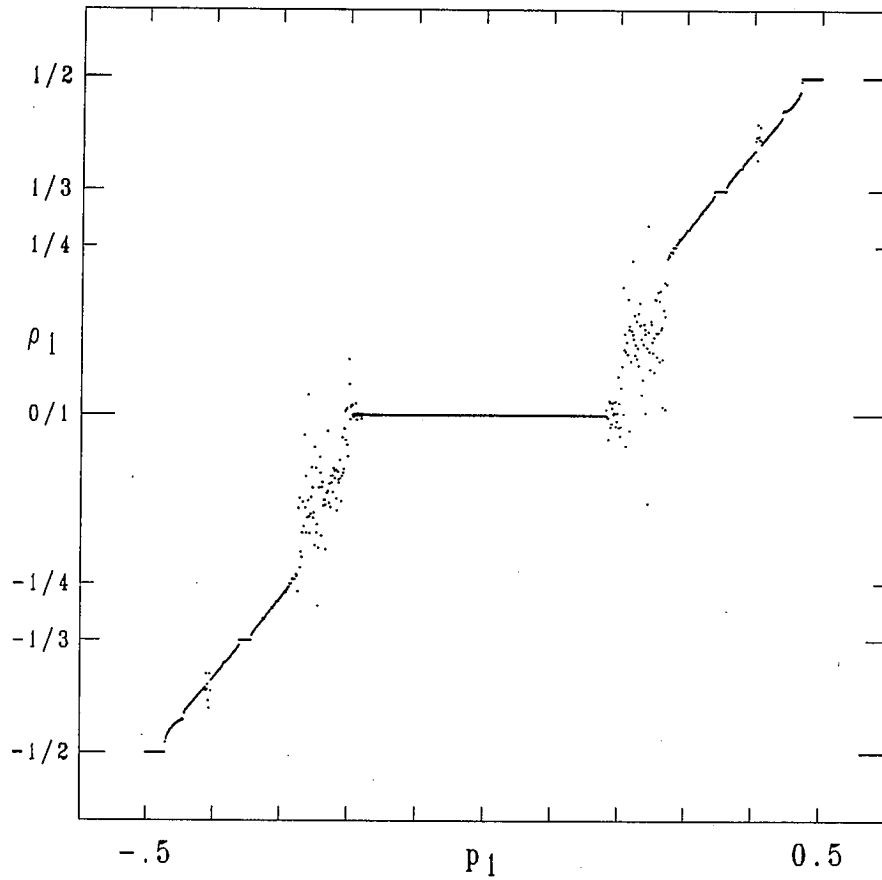


Fig. 2.13 Staircase of rotation numbers for initial conditions on the line $(p_1, -p_1, 0, 0)$. Shown is the frequency ρ_1 versus p_1 . An iteration time of 10^3 is used to define ρ .

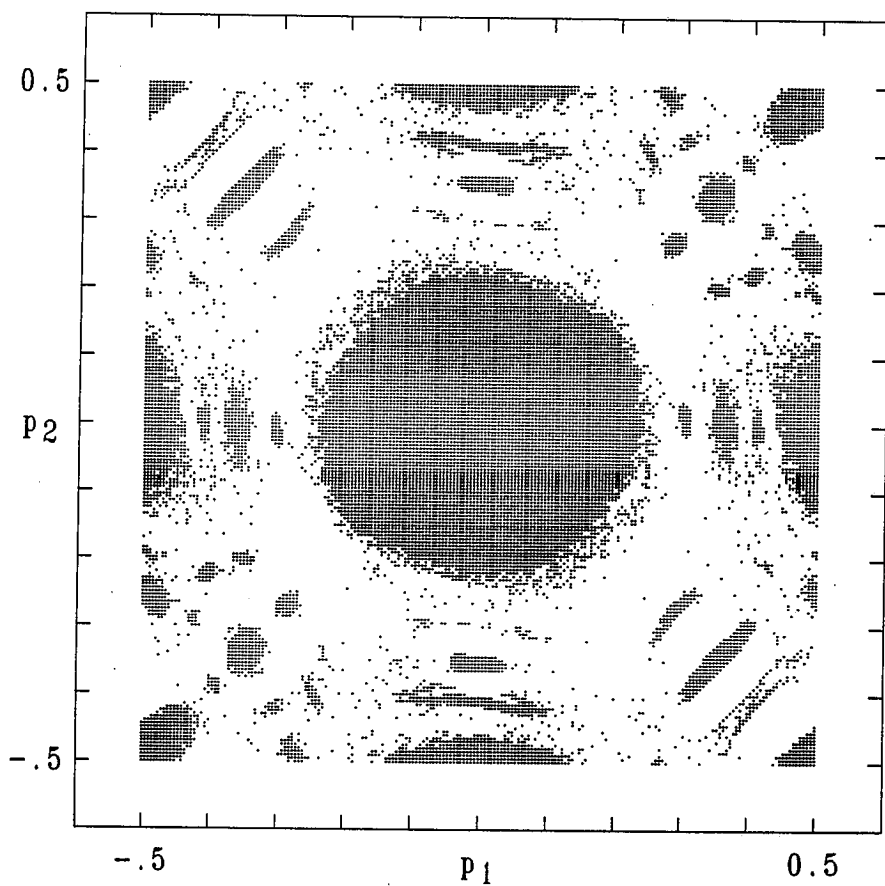


Fig. 2.14 Resonance zones for rotation numbers up to level 5 on the Farey tree (where $(1,1,2)$ is level 0). A grid of 200×200 initial conditions is taken on the $\mathbf{q} = 0$ plane. Points are filled in if they give a rotation number in one of the 81 resonances on the tree within a tolerance $\varepsilon = 0.01$.

2.7. Conclusions

We have shown that reversible, symplectic, twist mappings of a $2N$ dimensional phase space have periodic orbits which can be classified by frequency, symmetry, and Morse index. Such mappings have 2^N different types of symmetric orbits for each rational frequency vector ω . For each frequency we also find 2^N orbits for which the action function has indices ranging from 0 to N . For small deviations from the trivial twist map ($k_1=k_2=h=0$) these two sets of orbits often coincide; however the symmetric orbits commonly undergo symmetry breaking bifurcations which change their indices.

The existence of symmetric orbits is helpful for several reasons. It allows classification of orbits by symmetry properties. Furthermore symmetric orbits have points on the fixed planes of the symmetry; this reduces the numerical computations needed to find the orbits. Symmetric periodic orbits give useful information on the structure of phase space since they form the centers and edges of resonances, and can be used as approximations to quasiperiodic orbits such as the KAM tori.

Orbits which are local minima of the action appear to often be hyperbolic, though counter-examples can be constructed. Equation (2.31) implies only that the product of $-R_i$ must be positive for a non-degenerate minimum. In the area preserving case, this implies hyperbolicity.

A projection of periodic orbits onto the configuration space (e.g. Fig. 2.5) shows that as the system becomes more nonlinear (e.g. the potential

energy becomes larger), the orbits with index less than N tend to avoid the symmetry plane on which $-V$ is maximum. This is the same mechanism which gives rise to the formation of cantori in the area preserving map, so it seems reasonable to speculate that something similar happens here. When an avoided region is formed around the dominant plane, other such regions are formed by iteration. The index N orbit seems to have a single point in the center of each of these regions. The resulting picture, for an incommensurate orbit, would seem to be a set which is connected, but has an infinity of deleted holes. This is not a Cantor set, but probably a Sierpinski carpet [Sierpinski 1916]. We hope to present evidence for Sierpinski carpets in a later work.

A cross section of the phase space structure is obtained by plotting the intersection of symmetric periodic orbits with the symmetry plane, Fig. 2.10. This shows that the region around a periodic orbit is avoided by orbits with different frequencies. We interpret this region as the cross section of a resonance. Orbits in a resonance can remain trapped for long periods, and while they are trapped have the same frequency as the resonance. Collections of resonances which obey a commensurability relation, $\omega \cdot \mathbf{j} = 0$, form channels on the symmetry plane. Chaotic motion appears to take place along the edges of these channels. The collection of channels gives the Arnold web. A description of Arnold diffusion in terms of transitions from one resonance to another would require a better understanding of the resonance boundaries, and a definition for escaping flux like that for area preserving mappings [MacKay 1984, 1987].

CHAPTER 3

LAGRANGIAN FORMULATION AND NEWTON'S METHODS

3.1. Introduction

A Lagrangian system is one whose dynamics is described by a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}, \tau)$ through a variational principle. Many systems are Lagrangian including magnetic field lines, geodesic flows, geometrical optics, hydrodynamics of ideal fluids, etc. A state of the system is represented by a point \mathbf{q} in an N -dimensional configuration manifold \mathcal{M} , together with a velocity in the tangent space $T_{\mathbf{q}}\mathcal{M}$.

The Lagrangian formulation can be converted into a Hamiltonian one by Legendre transformation, introducing canonical momenta, $\mathbf{p} = \partial L(\mathbf{q}, \dot{\mathbf{q}}, \tau) / \partial \dot{\mathbf{q}}$. The dynamics is given by flow of the one parameter group of diffeomorphisms in the phase space (\mathbf{p}, \mathbf{q}) . Its behavior can be equivalently viewed by looking at successive intersections on a sectional surface in the phase space which is transverse to the flow (Poincaré surface of section). For conservative systems with a compact energy surface (i.e., the motion is bounded) the existence of such a section is guaranteed by Poincaré's recurrence theorem [Arnol'd 1978]. The induced map on the surface of section is symplectic since a Hamiltonian phase flow preserves symplectic structure. A formal conversion of a continuous time system to a discrete one by introducing a map is possible for autonomous systems and periodically time dependent systems. In the Lagrangian representation, the mapping takes a point $(\mathbf{q}, \dot{\mathbf{q}}')$ to $(\mathbf{q}', \dot{\mathbf{q}}'')$ so that its domain is a subset of $\mathcal{M} \times \mathcal{M}$.

Much of the analysis of dynamical systems relies on periodic orbits and the behavior of flows in their neighborhood. For example, quasiperiodic

orbits can be obtained by interpolating nearby periodic orbits, the frequencies of which constitute the sequence of approximants to the frequency of the quasiperiodic orbit [Greene 1979, MacKay 1984].

Finding an orbit of a map T for given period and parameters is equivalent to solving a system of nonlinear simultaneous equations. By definition, an orbit of period n is obtained by finding a point \mathbf{z}_0 such that its n -th iterate returns to itself: $\mathbf{z}_0 = (\mathbf{p}_0, \mathbf{q}_0)$ for symplectic maps or $\mathbf{z}_0 = (\mathbf{q}_0, \mathbf{q}_1)$ for a map of configuration space. Namely, the system to be solved is $\mathbf{f}(\mathbf{z}_0) \equiv T^n \mathbf{z}_0 - \mathbf{z}_0 = \mathbf{0}$. This requires a $2N$ -dimensional root finder for $2N$ -dimensional mappings. Note that the evaluation of the $\mathbf{f}(\mathbf{z}_0)$ involves iterating the map n times, the period of the orbit; this scheme of finding orbits will be referred to as 'direct map iteration method'. When the desired orbit is unstable, an initial numerical error grows exponentially upon iteration of the map: $|\delta \mathbf{z}_t| \approx \exp(\sigma t) |\delta \mathbf{z}_0|$ where $\sigma > 0$ is the Lyapunov exponent. Thus the desired precision on \mathbf{z}_t is lost along the orbit. For moderate periods the numerical error $|\delta \mathbf{z}_t|$ can easily become of order one; therefore, any scheme based on iteration becomes impractical for long enough, highly unstable, periodic orbits.

Alternatively, one can search for stationary configurations of the action function according to the variational principle for periodic orbits. This scheme is suitable for finding unstable orbits since numerical errors on the orbit do not evolve under direct mappings. In this case the set of equations to be solved is $\mathbf{g}(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}) \equiv \nabla W_n = \mathbf{0}$, where W_n is the action function of the orbit and \mathbf{g} is a $N \times n$ dimensional vector function; this scheme will be

referred to as 'variational method'. The number of equations to be simultaneously solved increases linearly with n , the period of an orbit. Accordingly, the main disadvantage of this method is that the computation time grows rapidly as n is increased. With straightforward Newton or quasi-Newton methods, the computation time grows as $O(n^3)$ [Press 1986, Polak 1971, Dennis 1977, Broyden 1965].

For the fixed point and period two orbits, either system of equations can be solved analytically as shown in section 2.3. For longer orbits, i.e. the typical case, one must rely on numerical computations. There are many methods in the literature for iterative solution of a set of nonlinear equations. However, for iterative algorithms of this kind, there is no sufficient test for existence of solutions. Convergence of an algorithm to a solution is guaranteed only when the solution exists and the initial trial solution is within the basin of attraction for the desired solution [Ortega 1970]. This basin typically has a very complicated structure; however, if a trial solution is chosen sufficiently close to the actual solution, certain algorithms which will be discussed in the following sections have been shown to exhibit excellent convergence.

In our case, the orbits are known to exist (see Appendix B). However, the equations are so complicated that it is hopeless to analyze the behavior of the function in the neighborhood of a solution. Thus the best procedure is to try several methods to find the most suitable one to the particular problem.

In this chapter we present two schemes employing Newton's method to find periodic orbits of Lagrangian systems whose configuration space is

N-dimensional. As a preliminary, the Lagrangian dynamical system is introduced in section 3.2 with a derivation of a map as discretization of a continuous time flow. Section 3.3 presents a direct map iteration scheme to find periodic orbits. The algorithm is based on Broyden's method which is one of quasi-Newton methods. We find that the direct map iteration method works well as long as the orbit is not too unstable. Section 3.4 presents the linear algorithm for Newton's method based on the variational method. It is shown that the computation time of the resulting algorithm is proportional to n , the period of the orbit. It is also shown that the algorithm occupies storage space of size $O(n)$. Section 3.5 is devoted to discussions on the variational method presented in section 3.4, which includes discussions on the problems of obtaining a good initial guess and numerical overflow stemming from the determinant of the Jacobian matrix becoming small. The relation of the index of the action function to orbital stability is also discussed. We illustrate the method for the case $N=2$. The subroutines are provided in Appendix C.

3.2. Lagrangian Dynamical Systems

Consider a dynamical system represented by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, \tau)$. Here \mathbf{q} is a point in some N -dimensional configuration manifold \mathcal{M} . The action of a path $\mathbf{q}(\tau)$, $\tau_0 \leq \tau \leq \tau_n$ is defined as

$$W[\mathbf{q}(\tau)] = \int_{\tau_0}^{\tau_n} L(\mathbf{q}, \dot{\mathbf{q}}, \tau) d\tau \quad (3.1)$$

Stationary paths, determined by setting the first variation δW to zero for $\mathbf{q}(\tau_0)$ and $\mathbf{q}(\tau_n)$ fixed, satisfy the Euler-Lagrange equations of motion.

Conversion to a discrete time mapping can be done formally for autonomous and periodically time dependent systems. To this end, introduce a partition of the time interval, $[\tau_0, \tau_1, \dots, \tau_n]$; each step corresponds to successive return of a flow to the Poincaré surface of section for autonomous systems, or the strobing period for periodically time dependent systems. The discrete Lagrangian can be defined by integrating L along an stationary orbit segment, $\mathbf{q}(\tau)$, which begins at the point \mathbf{q}_i at $\tau=\tau_i$, and ends at \mathbf{q}_{i+1} at $\tau=\tau_{i+1}$

$$F(\mathbf{q}_i, \mathbf{q}_{i+1}) \equiv \int_{\tau_i}^{\tau_{i+1}} L(\mathbf{q}, \dot{\mathbf{q}}, \tau) d\tau \quad (3.2)$$

$\mathbf{q}(\tau)$ stationary

The action of an orbit segment from $\tau=\tau_i$ to $\tau=\tau_j$ can now be written

$$W(\mathbf{q}_i, \mathbf{q}_{i+1}, \dots, \mathbf{q}_j) = \sum_{t=i}^{j-1} F(\mathbf{q}_t, \mathbf{q}_{t+1}) \quad (3.3)$$

where t is an integer index. Variation of W with respect to the intermediate points yields the mapping equations

$$\frac{\partial}{\partial \mathbf{q}_t} [F(\mathbf{q}_{t-1}, \mathbf{q}_t) + F(\mathbf{q}_t, \mathbf{q}_{t+1})] = 0 \quad (3.4)$$

This equation locally defines a unique mapping, T , from $(\mathbf{q}_{t-1}, \mathbf{q}_t)$ to $(\mathbf{q}_t, \mathbf{q}_{t+1})$ providing $\det(\partial^2 F / \partial \mathbf{q} \partial \mathbf{q}')$ is never zero. We make a stronger assumption, the twist condition: the mixed partial derivative matrix is assumed to be uniformly negative definite, i.e. there exists a $B > 0$ such that for any $\delta \mathbf{q}, \mathbf{q}, \mathbf{q}'$

$$\delta \mathbf{q} \cdot \mathbf{b}(\mathbf{q}, \mathbf{q}') \cdot \delta \mathbf{q} \geq B |\delta \mathbf{q}|^2$$

$$\mathbf{b}(\mathbf{q}, \mathbf{q}') \equiv - \partial^2 F(\mathbf{q}, \mathbf{q}') / \partial \mathbf{q} \partial \mathbf{q}' = F_{12}(\mathbf{q}, \mathbf{q}') \quad (3.5)$$

Here we use the subscripts "1" and "2" to indicate derivatives with respect to first and second arguments, respectively. It can be seen that (3.5) globally implies the existence of T [MacKay 1988]. The twist condition is the analogue of the Legendre condition in continuous time.

An orbit is a doubly infinite sequence of configurations $\{ \dots, \mathbf{q}_t, \mathbf{q}_{t+1}, \mathbf{q}_{t+2}, \dots \}$ such that every finite segment $\{ \mathbf{q}_t, \mathbf{q}_{t+1}, \dots, \mathbf{q}_s \}$ is a stationary point of the action $W(\mathbf{q}_t, \mathbf{q}_{t+1}, \dots, \mathbf{q}_s)$ for given \mathbf{q}_t and \mathbf{q}_s . An orbit is periodic with period n if

$$\mathbf{q}_{t+n} = \mathbf{q}_t \quad (3.6)$$

such that n is the least integer satisfying Eq. (3.6). The action for a periodic orbit is

$$W_n(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}) \equiv W(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}, \mathbf{q}_n = \mathbf{q}_0) \quad (3.7)$$

Equation (3.7) gives a variational principle for periodic orbits: a periodic orbit of period n is a critical point of W_n where all points $\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}$ are varied freely. This is true because criticality of W_n implies Eq. (3.4) for $0 < t < n$, and variation with respect to \mathbf{q}_0 implies $F_1(\mathbf{q}_0, \mathbf{q}_1) + F_2(\mathbf{q}_{n-1}, \mathbf{q}_0) = 0$, which implies (3.6) when (3.5) is satisfied.

The linear stability of an orbit is determined by considering its tangent map, which is obtained by linearizing Eq. (3.4):

$$\begin{aligned}
 -\mathbf{b}_{t+1} \delta \mathbf{q}_{t+1} + \mathbf{a}_t \delta \mathbf{q}_t - \tilde{\mathbf{b}}_t \delta \mathbf{q}_{t-1} &= 0 \\
 \mathbf{b}_t &\equiv -F_{12}(\mathbf{q}_{t-1}, \mathbf{q}_t) \\
 \mathbf{a}_t &\equiv F_{11}(\mathbf{q}_t, \mathbf{q}_{t+1}) + F_{22}(\mathbf{q}_{t-1}, \mathbf{q}_t)
 \end{aligned} \tag{3.8}$$

giving a linear second difference equation. Here ‘ \sim ’ designates transpose of a matrix. The multipliers of a period n orbit are determined by the eigenvalue problem

$$\delta \mathbf{q}_{t+n} = \lambda \delta \mathbf{q}_t \tag{3.9}$$

which closes the recurrence relation (3.8).

When the system represented by the Lagrangian (3.2) is periodic, that is, $F(\mathbf{q} + \mathbf{m}, \mathbf{q}' + \mathbf{m}) = F(\mathbf{q}, \mathbf{q}')$, where \mathbf{m} is an arbitrary integer vector, the configuration manifold \mathcal{M} can be thought of as an N -torus. In such a case, as shown in section 2.2, it is often convenient to consider a lift of the original map over the covering space of the N -torus, \mathcal{R}^N . Then the rotation frequency of an orbit, if it exists, can be defined as

$$\omega = \lim_{t-t' \rightarrow \infty} \frac{\mathbf{q}(t) - \mathbf{q}(t')}{t - t'}$$

and the periodicity condition for a periodic orbit, Eq.(3.6), can be generalized to

$$\mathbf{q}_{t+n} = \mathbf{q}_t + \mathbf{m} \quad \mathbf{m} \text{ is an integer vector.} \quad (3.10)$$

Therefore, every periodic orbit has a frequency $\omega = \mathbf{m}/n$.

We often adopt the definition (3.10) when we apply the Newton's methods to the periodic reversible maps since, for these maps, the symmetry properties shown in section 2.4 can be used to reduce computational effort.

3.3. Direct Map Iteration using Broyden's Method

Using the definition (3.10) for the periodicity, a periodic orbit of frequency (m,n) is obtained by finding a point \mathbf{z}_0 such that its n -th iterate is $R_m \mathbf{z}_0$, where R is the translation operator defined in Eq. (2.12). Namely, the system to be solved is $\mathbf{f}(\mathbf{z}_0) \equiv R_m T^n \mathbf{z}_0 - \mathbf{z}_0 = \mathbf{0}$, where T is the map. This requires a $2N$ -dimensional root finder for the $2N$ dimensional mapping. However, as will be shown below, it is possible to reduce the search to N dimensions, providing we consider only symmetric orbits.

Broyden's method [Broyden 1965], is a quasi-Newton or variable metric method. Briefly, the algorithm is

problem: Solve $\mathbf{f}(\mathbf{x}^*)=0$

Define $\mathbf{f}_k=\mathbf{f}(\mathbf{x}_k)$, $\mathbf{s}_k=\mathbf{x}_k-\mathbf{x}_{k-1}$ and $\mathbf{y}_k=\mathbf{f}_k-\mathbf{f}_{k-1}$

Input: \mathbf{x}_0 , \mathbf{B}_0

Do loop for k from $k=1$

$$\mathbf{s}_k = -\mathbf{B}_{k-1}^{-1} \mathbf{f}_{k-1}$$

$$\mathbf{B}_k^{-1} = \mathbf{B}_{k-1}^{-1} - \frac{\mathbf{B}_{k-1}^{-1} \mathbf{f}_k \tilde{\mathbf{s}}_k \mathbf{B}_{k-1}^{-1}}{\tilde{\mathbf{s}}_k \mathbf{B}_{k-1}^{-1} \mathbf{y}_k} \quad \text{or} \quad \mathbf{B}_k = \mathbf{B}_{k-1} + \frac{\mathbf{f}_k \tilde{\mathbf{s}}_k}{|\mathbf{s}_k|^2} \quad (3.11)$$

Continue until $|\mathbf{s}_k| < \text{given tolerance}$

Here the vector \mathbf{s}_k is the correction to the approximate root \mathbf{x}_k and the matrix \mathbf{B}_k is an approximation of the exact Jacobian at \mathbf{x}_k .

It has been shown that this algorithm converges superlinearly to the solution \mathbf{x}^* , if a converging sequence $\{\mathbf{x}_k\}$ exists [Broyden 1973]. That is,

$$\lim_{k \rightarrow \infty} \frac{|\mathbf{x}_{k+1} - \mathbf{x}^*|}{|\mathbf{x}_k - \mathbf{x}^*|} = 0$$

The \mathbf{B}_k 's are uniquely determined by the following two conditions:

$$\mathbf{B}_k \cdot \mathbf{s}_k = \mathbf{y}_k \quad (3.12)$$

$$\mathbf{B}_k \cdot \mathbf{z} = \mathbf{B}_{k-1} \cdot \mathbf{z} \quad \text{for any } \mathbf{z} \text{ such that } \mathbf{s}_k \cdot \mathbf{z} = 0 \quad (3.13)$$

Equation (3.12) implies that \mathbf{B}_k is an approximate Jacobian at \mathbf{x}_k . For the one dimensional case, the approximate Jacobian corresponds to using the secant line to determine s_k at each stage. For the multi-dimensional case the first condition does not uniquely determine \mathbf{B}_k . There are various ways to fix \mathbf{B}_k ; for Broyden's method the approximate Jacobian is determined using Eq. (3.13). This condition together with Eq. (3.12) uniquely fixes \mathbf{B}_k . The resulting \mathbf{B}_k 's satisfy the recursion formula (3.11). This recursive relation offers significant reduction in the work involved in the calculation of the Jacobian, which usually requires $O(n^2)$ function evaluations at each stage. There are also several alternatives to (3.13) which give a symmetric \mathbf{B}_k [Dennis 1977, Polak 1971]. In our case, we observe that the latter methods involve more computations and do not give any improvement in convergence.

We employ Broyden's method to find periodic orbits by solving $\mathbf{f}(\mathbf{z}_0) \equiv \mathbf{R}_m \mathbf{T}^n \mathbf{z}_0 - \mathbf{z}_0 = \mathbf{0}$. As mentioned above, the required four dimensional search for \mathbf{z}_0 can be reduced to a two dimensional one for symmetric periodic orbits. This is done by utilizing the properties of the symmetric orbits discussed in section 2.4 as follows.

For a given frequency, the symmetry plane of \mathbf{z}_0 determines the symmetry plane of \mathbf{z}_k by Eq. (2.28). Choose a symmetry plane of \mathbf{z}_0 and let \mathbf{p}_0 be, say, the momentum of a point \mathbf{z}_0 on the plane. The function \mathbf{f} depends only on \mathbf{p}_0 i.e., $\mathbf{f} = \mathbf{f}(\mathbf{p}_0)$. Its value is determined by the equations for the symmetry plane of \mathbf{z}_k ; for example, $\mathbf{f} = \mathbf{q}_k - 1/2\mathbf{m}$ for \mathbf{z}_k on $\text{Fix}(\text{SR}_m)$ or $\mathbf{f} = \mathbf{p}_k - 2\mathbf{q}_k + \mathbf{m}$ for \mathbf{z}_k on $\text{Fix}(\text{TSR}_m)$ (See Eq.(2.26)). Therefore the function \mathbf{f} is a two dimensional vector. The explicit form of \mathbf{f} varies depending on the frequencies and the symmetry plane of \mathbf{z}_0 as shown in Table 2.1 of chapter 2. All cases can be neatly combined into one formula as follows.

$$\mathbf{f}(\mathbf{p}_0) = \mathbf{p}_k - \mathbf{C} \cdot (\mathbf{m} + \mathbf{R} - 2\mathbf{q}_{k-1}) \quad (3.14)$$

where $\mathbf{C} = 1/\text{OE} - \text{TS} \cdot \text{OE}/2$. Here $\text{OE}=1$ for odd n and $\text{OE}=2$ for even n , and $\text{TS}=1$ when \mathbf{z}_0 is on the symmetry plane of TS symmetry family (TS, $\text{TSR}_{(0,1)}$, etc.) and $\text{TS}=0$ otherwise. \mathbf{R} represents the translation of the symmetry plane of \mathbf{z}_0 , i.e., $\mathbf{R}=(1,0)$ for $\text{SR}_{(1,0)}$ or $\text{TSR}_{(1,0)}$, $\mathbf{R}=(0,1)$ for $\text{SR}_{(0,1)}$ or $\text{TSR}_{(0,1)}$ etc. Finally, the other half of the orbit is obtained by using the symmetry relation (27). Numerically this method can converge to the half orbit for cases when iteration of \mathbf{z}_k fails to converge to the second half of the orbit.

Choosing a good initial trial, \mathbf{p}_0 , is critical for the performance of any iterative solution technique. In the small parameter range, we choose $\mathbf{p}_0 = \mathbf{m}/n$, i.e., the rotation number of the orbit. This works well as long as the orbit is not far from the uniform rotation configuration. For an orbit with larger parameter, \mathbf{p}_0 can be chosen by extrapolating the orbits obtained with smaller parameters. We use the identity matrix for \mathbf{B}_0 . Using a Jacobian for \mathbf{B}_0 usually does not give any improvement in convergence. The recursion relation for \mathbf{B}_k rather than for its inverse (see Eq. (3.13)) is used because the former involves less computations and the step (3.12) can be easily done for a 2×2 matrix \mathbf{B}_k .

The mapping iteration scheme becomes impractical for a strongly unstable orbit because upon iterations of the map an initial numerical error grows exponentially. However, as long as the magnitude of the smallest multiplier is sufficiently large compared to the computer precision, this scheme works and is the simplest to implement.

3.4. Linear Newton's Algorithm for Variational Method

The algorithm for Newton's method can be written briefly as follows.

problem : Solve $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ where $\mathbf{g}: \mathcal{R}^n \rightarrow \mathcal{R}^n$

Definition : $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k)$, $\mathbf{s}_k = \mathbf{x}_k - \mathbf{x}_{k-1}$ and \mathbf{J}_k is Jacobian of \mathbf{g}_k

Input : \mathbf{x}_0

Do while $|\mathbf{s}_k| >$ given tolerance

 evaluate \mathbf{J}_k (step 1)

 solve $\mathbf{J}_k \mathbf{s}_{k+1} = -\mathbf{g}_k$ (step 2)

The vector \mathbf{s}_k is the correction to the approximate solution \mathbf{x}_k .

Newton's method is well known for its rapid quadratic convergence [Ortega 1970]. That is,

$$\lim_{k \rightarrow \infty} \frac{|\mathbf{x}_{k+1} - \mathbf{x}^*|}{|\mathbf{x}_k - \mathbf{x}^*|^2} = C$$

where C is a bounded constant and $\{\mathbf{x}_k\}$ is a converging sequence to a solution \mathbf{x}^* . One of its disadvantages is that it usually requires $O(n^2)$ function evaluations to calculate the Jacobian matrix \mathbf{J} in step 1 above. Furthermore, in step 2, one has to solve a linear system requiring $O(n^3)$ arithmetic

operations at each stage of the iteration; this can be reduced to $O(n^2)$ for the case that \mathbf{J} is a sparse matrix with band width r [Atkinson 1978].

We want to employ Newton's method to solve $\mathbf{g}(\mathbf{q}) \equiv \nabla W_n(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}) = 0$. The $N \times N$ Jacobian matrix, $\mathbf{J} \equiv \nabla \nabla W_n$, is given by Eq. (3.8) with the periodicity condition Eq. (3.6):

$$\mathbf{J} \equiv \nabla \nabla W_n = \begin{pmatrix} a_0 & -b_1 & 0 & \dots & 0 & \tilde{b}_0 \\ \tilde{b}_1 & a_1 & -b_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & \tilde{b}_{n-3} & a_{n-3} & -b_{n-2} & 0 \\ 0 & \dots & 0 & \tilde{b}_{n-2} & a_{n-2} & -b_{n-1} \\ -b_0 & 0 & \dots & 0 & \tilde{b}_{n-1} & a_{n-1} \end{pmatrix} \quad (3.17)$$

In the following we present an analytic inversion of \mathbf{J} . The explicit form of \mathbf{J}^{-1} is used to calculate \mathbf{s} . Namely, instead of solving the linear system $\mathbf{J} \mathbf{s} = -\mathbf{g}$, one obtains \mathbf{s} directly by calculating $\mathbf{s} = -\mathbf{J}^{-1} \mathbf{g}$. The resulting algorithm requires $O(n)$ computation time at each iterative stage while retaining the quadratic convergence of Newton's method.

The presentation of this method is divided into two parts; (A) inversion of \mathbf{J} via block-diagonalization and (B) calculation of \mathbf{s} . Below, we will denote the components of $N \times n$ dimensional vectors by $\tilde{\mathbf{x}} \equiv (\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{n-1})$ such that each \mathbf{x}_j is N dimensional. Similarly, each entry of an $N \times N$ matrix is denoted by an $N \times N$ submatrix, as in (3.17). The index k denoting k -th stage

of iterative Newton algorithm is dropped to avoid confusion with subscripts for vectorial components. Uppercase characters are used for matrices with the exception of the given matrices \mathbf{a}_i and \mathbf{b}_i .

(A) Inversion of \mathbf{J}

The primary goal of this part is to obtain a decomposition of \mathbf{J} into a product of easily invertible matrices (see Eq.(3.22)). We first show that \mathbf{J} can be block-diagonalized using a non-orthogonal transformation \mathbf{U} , which is upper-triangular with unit diagonal entries. Inversion of a triangular matrix is straightforward. Finally \mathbf{J}^{-1} is obtained as a product of the inverses of the decomposing matrices (Eq.(3.25)).

First, consider the quadratic form $\tilde{\mathbf{x}}\mathbf{J}\mathbf{x}$ where \mathbf{x} is an arbitrary vector. Successively completing the squares yields

$$\begin{aligned} \tilde{\mathbf{x}}\mathbf{J}\mathbf{x} = & \overbrace{(\mathbf{x}_0 - \mathbf{A}_0^{-1}\mathbf{b}_1\mathbf{x}_1 - \mathbf{B}_0^{-1}\mathbf{x}_{n-1})} \mathbf{A}_0 (\mathbf{x}_0 - \mathbf{A}_0^{-1}\mathbf{b}_1\mathbf{x}_1 - \mathbf{B}_0^{-1}\mathbf{x}_{n-1}) \\ & + \overbrace{(\mathbf{x}_1 - \mathbf{A}_1^{-1}\mathbf{b}_2\mathbf{x}_2 - \mathbf{B}_1^{-1}\mathbf{x}_{n-1})} \mathbf{A}_1 (\mathbf{x}_1 - \mathbf{A}_1^{-1}\mathbf{b}_2\mathbf{x}_2 - \mathbf{B}_1^{-1}\mathbf{x}_{n-1}) \\ & + \dots \\ & + \overbrace{(\mathbf{x}_{n-2} - \mathbf{A}_{n-2}^{-1}\mathbf{b}_{n-1}\mathbf{x}_{n-1} - \mathbf{B}_{n-2}^{-1}\mathbf{x}_{n-1})} \mathbf{A}_{n-2} (\mathbf{x}_{n-2} - \mathbf{A}_{n-2}^{-1}\mathbf{b}_{n-1}\mathbf{x}_{n-1} - \mathbf{B}_{n-2}^{-1}\mathbf{x}_{n-1}) \\ & + \tilde{\mathbf{x}}_{n-1} \mathbf{A}_{n-1} \mathbf{x}_{n-1} \end{aligned} \quad (3.18)$$

where the \mathbf{A}_i and \mathbf{B}_i are defined recursively by

$$\begin{aligned}
\mathbf{A}_0 &= \mathbf{a}_0 \\
\mathbf{A}_i &= \mathbf{a}_i - \tilde{\mathbf{b}}_i \mathbf{A}_{i-1}^{-1} \mathbf{b}_i \quad 1 \leq i < n-1 \\
\mathbf{A}_{n-1} &= \mathbf{a}_{n-1} - \tilde{\mathbf{b}}_{n-1} \mathbf{A}_{n-2}^{-1} \mathbf{b}_{n-1} - \mathbf{D}_0 - \tilde{\mathbf{D}}_{n-1}
\end{aligned} \tag{3.19a}$$

$$\begin{aligned}
\mathbf{B}_{-1} &= \mathbf{I} \\
\mathbf{B}_i &= \mathbf{B}_{i-1} \tilde{\mathbf{b}}_i^{-1} \mathbf{A}_i \quad 0 \leq i < n-1 \\
\mathbf{B}_{n-1} &= \mathbf{I}
\end{aligned} \tag{3.19b}$$

and we have introduced matrices \mathbf{D}_i defined by

$$\begin{aligned}
\mathbf{D}_{n-1} &= \tilde{\mathbf{b}}_{n-1} \mathbf{B}_{n-2}^{-1} \\
\mathbf{D}_i &= \mathbf{D}_{i+1} + \tilde{\mathbf{B}}_i^{-1} \mathbf{A}_i \mathbf{B}_i^{-1} \quad n-1 > i \geq 0
\end{aligned} \tag{3.19c}$$

Note that \mathbf{A}_i and $\mathbf{D}_i - \mathbf{D}_{n-1}$ are symmetric.

Thus Eq. (3.18) can be re-expressed as

$$\tilde{\mathbf{x}} \mathbf{J} \mathbf{x} = \tilde{\mathbf{y}} \mathcal{D} \mathbf{y} \tag{3.20}$$

where \mathcal{D} is a block-diagonal matrix with the submatrices \mathbf{A}_i along the diagonal. The vector \mathbf{y} is obtained from \mathbf{x} by a non-orthogonal transformation, $\mathbf{y} = \mathbf{U}\mathbf{x}$, where \mathbf{U} is the upper triangular matrix

$$\mathbf{U} = \begin{pmatrix}
\mathbf{I} & -\mathbf{A}_0^{-1} \mathbf{b}_1 & 0 & \dots & \dots & -\mathbf{B}_0^{-1} \\
0 & \mathbf{I} & -\mathbf{A}_1^{-1} \mathbf{b}_2 & 0 & \dots & -\mathbf{B}_1^{-1} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
\dots & \dots & 0 & \mathbf{I} & -\mathbf{A}_{n-3}^{-1} \mathbf{b}_{n-2} & -\mathbf{B}_{n-3}^{-1} \\
\dots & \dots & \dots & 0 & \mathbf{I} & -\mathbf{A}_{n-2}^{-1} \mathbf{b}_{n-1} - \mathbf{B}_{n-2}^{-1} \\
0 & \dots & \dots & \dots & 0 & \mathbf{I}
\end{pmatrix} \tag{3.21}$$

Since \mathbf{x} is arbitrary in Eq. (3.20), we obtain the desired decomposition of \mathbf{J} ;

$$\mathbf{J} = \tilde{\mathbf{U}} \mathbf{D} \mathbf{U} \quad (3.22)$$

The inverse of \mathbf{U} is obtained by solving for \mathbf{x}_i in terms of \mathbf{y}_i ;

$$\mathbf{x}_{n-1} = \mathbf{y}_{n-1} \quad (3.23a)$$

$$\mathbf{x}_i = \mathbf{y}_i + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \mathbf{x}_{i+1} + \mathbf{B}_i^{-1} \mathbf{y}_{n-1} \quad n-2 \geq i \geq 0 \quad (3.23b)$$

$$= \mathbf{y}_i + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} (\mathbf{y}_{i+1} + \mathbf{A}_{i+1}^{-1} \mathbf{b}_{i+2} \mathbf{x}_{i+2} + \mathbf{B}_{i+1}^{-1} \mathbf{y}_{n-1}) + \mathbf{B}_i^{-1} \mathbf{y}_{n-1}$$

= ...

$$= \mathbf{y}_i + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \mathbf{y}_{i+1} + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \mathbf{A}_{i+1}^{-1} \mathbf{b}_{i+2} \mathbf{y}_{i+2} + \dots + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \dots \mathbf{A}_{n-3}^{-1} \mathbf{b}_{n-2} \mathbf{y}_{n-2}$$

$$+ (\mathbf{B}_i^{-1} + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \mathbf{B}_{i+1}^{-1} + \dots + \mathbf{A}_i^{-1} \mathbf{b}_{i+1} \mathbf{A}_{i+1}^{-1} \mathbf{b}_{i+2} \dots \mathbf{A}_{n-2}^{-1} \mathbf{b}_{n-1} \mathbf{B}_{n-1}^{-1}) \mathbf{y}_{n-1}$$

Each coefficient of \mathbf{y}_j corresponds to $(\mathbf{U}^{-1})_{ij}$. That is, after straightforward matrix manipulations by using (3.19),

$$\begin{aligned} (\mathbf{U}^{-1})_{ij} &= \mathbf{b}_i^{-1} \tilde{\mathbf{B}}_{i-1} \tilde{\mathbf{B}}_{j-1}^{-1} \mathbf{b}_j & 0 \leq i \leq j \leq n-2 \\ (\mathbf{U}^{-1})_{in-1} &= \mathbf{b}_i^{-1} \tilde{\mathbf{B}}_{i-1} \mathbf{B}_i & 0 \leq i \leq n-1 \\ (\mathbf{U}^{-1})_{ij} &= 0 & i > j \end{aligned} \quad (3.24)$$

Now \mathbf{J}^{-1} is obtained from the inverses of \mathbf{D} and \mathbf{U} :

$$\mathbf{J}^{-1} = \mathbf{U}^{-1} \mathbf{D}^{-1} \tilde{\mathbf{U}}^{-1}, \text{ i.e.,}$$

$$(\mathbf{J}^{-1})_{ij} = \sum_{l=\max(i,j)}^{n-1} (\mathbf{U}^{-1})_{il} \mathbf{A}_l^{-1} (\tilde{\mathbf{U}}^{-1})_{lj} \quad (3.25)$$

Inserting Eq.(3.24) into Eq.(3.25) and using Eq.(3.19), we obtain the final expression for \mathbf{J}^{-1} :

$$(\mathbf{J}^{-1})_{ij} = \mathbf{b}_i^{-1} \tilde{\mathbf{B}}_{i-1} (\mathbf{D}_{\max(i,j)} - \mathbf{D}_{n-1} + \tilde{\mathbf{D}}_i \mathbf{A}_{n-1}^{-1} \mathbf{D}_j) \mathbf{B}_{j-1} \tilde{\mathbf{b}}_j^{-1} \quad (3.26)$$

Thus \mathbf{J}^{-1} has been obtained in analytic form. It is explicitly symmetric. The matrices \mathbf{A}_i , \mathbf{B}_i and \mathbf{D}_i are obtained recursively as shown in Eq.(3.19). However, we do not evaluate \mathbf{J}^{-1} because this would require $\mathcal{O}(n^2)$ arithmetic operations and $\mathcal{O}(n^2)$ storage space for entries of the \mathbf{J}^{-1} .

(B) Calculation of $\mathbf{s} = -\mathbf{J}^{-1} \mathbf{g}$

By virtue of Eq.(3.26), the analytic form of \mathbf{s} can be directly obtained.

$$\begin{aligned} -\mathbf{s}_i &= (\mathbf{J}^{-1})_{ij} \mathbf{g}_j \\ &= \tilde{\mathbf{Q}}_i \sum_{j<i} \mathbf{B}_{j-1} \tilde{\mathbf{b}}_j^{-1} \mathbf{g}_j + \mathbf{b}_i^{-1} \tilde{\mathbf{B}}_{i-1} \sum_{j \geq i} \mathbf{Q}_j \mathbf{g}_j + \tilde{\mathbf{g}}_i \mathbf{A}_{n-1}^{-1} \sum_{j=0}^{n-1} \mathbf{P}_j \mathbf{g}_j \end{aligned} \quad (3.27)$$

where the matrices \mathbf{P}_i , \mathbf{Q}_i are defined as follows:

$$\begin{aligned} \mathbf{P}_{n-1} &= \mathbf{I} \\ \mathbf{P}_i &= \mathbf{P}_{i+1} \tilde{\mathbf{b}}_{i+1}^{-1} \mathbf{A}_i^{-1} + \tilde{\mathbf{B}}_i^{-1} & n-2 \geq i \geq 0 \\ \mathbf{Q}_i &= \mathbf{P}_i - \mathbf{D}_{n-1} \mathbf{B}_{i-1} \tilde{\mathbf{b}}_i^{-1} & n-1 \geq i \geq 0 \end{aligned} \quad (3.28)$$

Equation (3.27) can be further simplified by introducing two vectors \mathbf{u}_i and \mathbf{d}_i ;

$$\begin{aligned}
\mathbf{u}_0 &= \mathbf{d}_{n-1} = 0 \\
\mathbf{u}_i &= \mathbf{u}_{i-1} + \mathbf{B}_{i-2} \tilde{\mathbf{b}}_{i-1}^{-1} \mathbf{g}_{i-1} & 1 \leq i \leq n-1 \\
\mathbf{d}_i &= \mathbf{d}_{i+1} + \mathbf{Q}_i \mathbf{g}_i & n-2 \geq i \geq 0
\end{aligned} \tag{3.29}$$

so that the recursive relations determine \mathbf{u}_i upwardly from \mathbf{u}_0 and \mathbf{d}_i downwardly from \mathbf{d}_{n-1} . The final expression for \mathbf{s} is

$$-\mathbf{s}_i = \tilde{\mathbf{Q}}_i \mathbf{u}_i + \mathbf{b}_i^{-1} \tilde{\mathbf{B}}_{i-1} \mathbf{d}_i + \tilde{\mathbf{P}}_i \left[\mathbf{A}_{n-1}^{-1} \sum_{j=0}^{n-1} \mathbf{P}_j \mathbf{g}_j \right] \tag{3.30}$$

Once the i -independent term inside the bracket of Eq.(3.30) is obtained, the evaluation of each component \mathbf{s}_i involves only two additions and four multiplications of N dimensional vectors and/or $N \times N$ matrices. As shown in Eq.(3.19), (3.28) and (3.29), all matrices and vectors are obtained recursively. To obtain \mathbf{D}_i from \mathbf{D}_{i+1} for example, one addition and two multiplications of matrices are required. Thus $3n-2$ matrix operations are required to obtain \mathbf{D}_i for all i . Therefore, the total number of arithmetic operations involved in calculating \mathbf{s} at each stage of Newton method is proportional to n , which is the period of the orbit, neglecting the factor involving N . The function evaluations are involved only in calculating \mathbf{a}_i and \mathbf{b}_i as well as \mathbf{g}_i . Furthermore, the number of $N \times N$ matrices and N -dimensional vectors which need to be stored is proportional to n . Consequently, the size of storage space for this algorithm, which could have been $O(n^2)$, is reduced to $O(n)$.

3.5. Discussion

In this section we mainly discuss the variational method which we presented in section 3.4 since the direct map iteration method seems rather straightforward. We remark, however, that there are cases in which iterative techniques are superior to variational methods; for example, when the desired orbits are stable or only slightly unstable. In fact, most of orbits studied in chapter 2 has been obtained by the method presented in section 3.3. Since the applicability of any numerical schemes finding orbits is not precisely defined one should always try several methods to find the one most suitable to any particular problem.

The method presented in section 3.4 is suitable for finding strongly unstable periodic orbits for a wide range of Lagrangian systems. Its main advantages are that the storage space and computation time are proportional to the period of the orbit. When $N=1$ (i.e., for area preserving maps), our method is equivalent in its efficiency to the Green function method proposed by B. Mestel and I.C. Percival [Mestel 1987, Chen 1987].

The main disadvantage of any method for the iterative solution of a set of nonlinear simultaneous equations is the critical choice of a good initial trial solution [Press 1986]. Convergence of an algorithm to a solution is guaranteed only when the solution exists, and the initial trial is within its basin of attraction. This basin typically has a very complicated, fractal structure. In general, for systems of nonlinear equations there is no sufficient test for the existence of solutions; however, for the case of twist mappings,

periodic solutions can be shown to exist for any period (see Appendix B of chapter 2).

Though existence of a solution is guaranteed, our method is by no means immune to the problem of finding a good initial trial. This choice is further complicated because it is necessary to guess values for all $N \times N$ variables of an orbit, unlike methods based on iteration for which only $2N$ initial conditions are required. Even though there have been various techniques devised to supply a good initial trial [Mestel 1987, Chen 1987], this is still a difficult problem. In our applications, as discussed below, we use extrapolation from solvable cases to obtain a trial.

Beside the linear dependence on the period of the orbit, another benefit of the present method is that the \mathbf{A}_t 's give information on the index of the action function of the orbit : the index of a function is defined as the number of downward directions at a stationary point [Milnor 1963]. The index has been shown to be closely related to the orbital stability even though it is not a sufficient test for stability for general N (see section 2.5). That is,

$$\prod_{i=1}^N R_i = \left(-\frac{1}{4}\right)^N \frac{\prod_{t=0}^{n-1} \det \mathbf{A}_t}{\prod_{t=0}^{n-1} \det \mathbf{b}_t} \quad (3.31)$$

where R_i is the residue associated with the corresponding reciprocal pair of multipliers, i.e., $R_i = 1/4 (2 - \lambda_i - 1/\lambda_i)$. Since the \mathbf{b}_t 's have been assumed to be positive definite (Eq. (3.5)), Eq. (3.31) implies that an action minimizing orbit is hyperbolic for $N=1$ [MackKay 1983].

Implementation of the method can be confronted with an overflow problem when \mathbf{J} is nearly singular. That is, $\det(\mathbf{J})$ can become extremely small, which could result in overflow when computing \mathbf{J}^{-1} , even though $\mathbf{s} = -\mathbf{J}^{-1}\mathbf{g}$ is finite. Indeed this is the situation when the action is near an inflection point: as can be seen from Eq. (3.31), one of residues is very close to zero. A typical example is when an orbit is about to undergo tangent bifurcation where it has a pair of multipliers at $+1$. Another case is when the perturbation is so small that orbits are very close to the configurations of uniform rotation.

This problem cannot be resolved without loss of efficiency. However, its effect seems to be greatly lessened in our method by using $\det(\mathbf{A}_i)$ instead of $\det(\mathbf{J})$. Namely, each $\det(\mathbf{A}_i)$ can be finite even when $\det(\mathbf{J})$ would underflow. In examples we observe that $\det(\mathbf{A}_i)$ indeed tends to be finite with the exception of $\det(\mathbf{A}_{n-1})$ which becomes very small for the nearly integrable case. However, \mathbf{A}_{n-1} is not involved in determining the other matrices so that the algorithm less often encounters such a problem. Thus our method works for the orbits whose residues are fairly close to zero as well as for strongly unstable orbits.

We illustrate the present method with an $N=2$ example for which the generating function is

$$F(\mathbf{q}, \mathbf{q}') = \frac{1}{2} (\mathbf{q}' - \mathbf{q}) \cdot \mathbf{b} \cdot (\mathbf{q}' - \mathbf{q}) - V(\mathbf{q}) \quad (3.32)$$

where $\mathbf{b} = \mathbf{I}$ and

$$V(q_1, q_2) = \frac{-1}{(2\pi)^2} \{k_1 \cos(2\pi q_1) + k_2 \cos(2\pi q_2) + h \cos[2\pi(q_1, q_2)]\}$$

The resulting map for the system (3.32) is equivalent to two coupled standard maps. Since the system is periodic in \mathbf{q} , the orbits can be depicted on a torus. Figure 3.1 shows the symmetric orbit of frequency $\omega = (12,16)/21$ with $\mathbf{q}_0=(1/2,0)$ along the path in parameter space: $k_1=\varepsilon$, $k_2=.6\varepsilon$, $h=.4\varepsilon$ and ε varies from 0 to 8.0 with a step size $\Delta\varepsilon=.16$ (for definition of symmetric orbits, see section 2.2). For $\varepsilon=0$ the configuration corresponds to a uniform rotation, $\mathbf{q}_j = j\omega + \mathbf{q}_0$. The multipliers of the orbit at $\varepsilon=8.0$ are $\lambda_1 \approx 10^{18}$ and $\lambda_2 \approx -10^{12}$. The orbit at ε is used as the initial trial to find the orbit at $\varepsilon+\Delta\varepsilon$. The required number of iterations of the Newton routine for each ε ranges from 3 to 5. As can be noticed from the figure, the points on the orbit tend to cluster together as ε becomes larger. Analogous clustering occurs for the area preserving case and leads to the formation of cantori.

The construction of our method relies only on the band structure of \mathbf{J} . The particular structure of \mathbf{J} with band width $r=3N$ in Eq.(3.17) originates from the nature of the second difference equation (3.8) or the equation of motion (3.4), for general Lagrangian systems introduced in section 3.2. Therefore the method presented in section 3.4 is expected to be applicable to a wide range of dynamical systems.

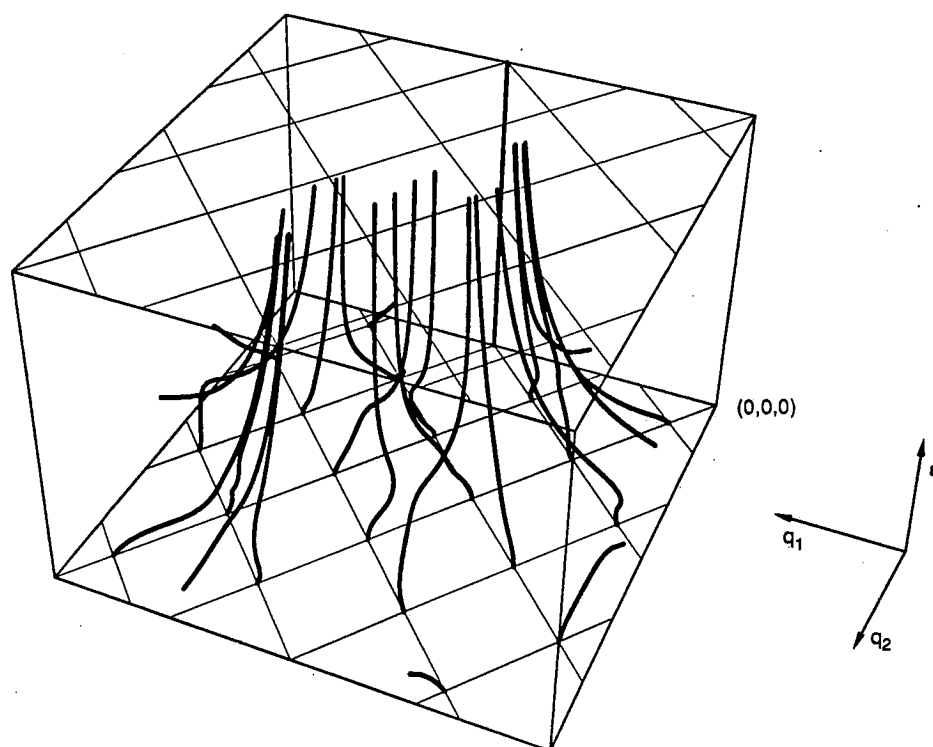


Fig. 3.1. Evolution of the points on a periodic orbit for the mapping Eq. (3.32). The box represents (q_1, q_2, ϵ) space which is $\mathcal{T}^2 \times \mathcal{R}$. The parameter ϵ varies from 0 to 8.0 in the vertical direction. The unperturbed orbit at the bottom of the box has a point at each intersection of the tilted grid lines whose slopes are determined from the frequency $\omega=(12,16)/21$.

CHAPTER 4

GEOMETRICAL STRUCTURE AND TRANSPORT IN PHASE SPACE

4.1 Introduction

The structure of phase space of Hamiltonian systems is generically a mixture of regular and chaotic (stochastic) motions. Chaotic regions are defined as the collection of orbits with a positive Lyapunov exponent : Birkhoff called such regions 'irregular components'. As the nonlinearity of the system is increased the stochasticity of phase space is generally increased. In the stochastic region of phase space, the motion is in practice random on a long time scale due to exponentially growing initial numerical errors. Thus, to an extent, transport in phase space can be modelled as a random process. However, the extent of randomness should depend on the details of the structure of phase space which is determined by the dynamics of a given map. This chapter aims at a study of geometrical structure of phase space and its application to possible descriptions of transport in phase space. Naturally, such a study is heavily based on an understanding of the dynamics of a given map.

Many of the properties of area preserving maps are well known. Some of these are reviewed in section 4.2, emphasizing the existence of partial barriers. Partial barriers are the remnants of invariant circles or separatrices and they can be constructed from segments of stable and unstable manifolds of hyperbolic invariant orbits. The action formulation for orbits, as introduced in section 2.2, turns out to play a central role in the analysis: existence of minimizing orbits and their ordering property, areas under partial barriers, flux through partial barriers, etc.

The results from section 4.2 are directly applicable to construction of Markov models for transport in phase space, as shown in section 4.3. A Markov process is defined by a set of states and transition probabilities between states. A region in phase space bounded by partial barriers is associated with a state of a Markov process. The transition probability is determined by the flux across the partial barrier between adjacent regions. The validation of the models and their results are discussed without showing the details on how to solve the corresponding master equations.

An extension of ideas employed in area preserving maps to the case of $2N$ dimensional symplectic maps is not trivial when $N > 1$. The main reason is that stable and unstable manifolds of minimizing hyperbolic orbits are N dimensional and, hence, do not divide phase space when $N > 1$. In fact, even an invariant torus does not divide phase space, resulting in the Arnol'd diffusion. Recently, however, it has been shown that in such systems under certain special conditions there exists a normally hyperbolic invariant manifold which has stable and unstable manifolds of co-dimension one. Such a normally hyperbolic invariant manifold is a set analogous to the hyperbolic periodic orbit in area preserving maps, and segments of its stable and unstable manifolds can be used to bound phase space since they are co-dimension one in the ambient phase space. In section 4.4 we show, without proofs, two propositions which state the existence and the persistence of such manifolds. An application to transport and the practicality of the theory are also discussed. In section 4.5 we close this chapter with some discussions.

4.2 Dynamics in Area Preserving maps

Many properties of area preserving maps have been determined by extensive research over the past decade [MacKay 1987 and references therein]. This section is devoted to a brief review of some of these results, emphasizing the structure of phase space in terms of periodic and quasiperiodic orbits together with their stable and unstable manifolds. It will be shown that the segments of stable and unstable manifolds can be used to construct partial barriers in phase space. There exist homoclinic orbits at intersections of stable and unstable manifolds. It will also be shown that action functions of homoclinic orbits can be used to define flux through the partial barriers. These ideas will be directly used to study transport in area preserving maps in section 4.3.

We will consider area preserving maps on the cylinder of the form

$$p_{t+1} = p_t - k f(x_t), \quad x_{t+1} = x_t + p_{t+1} \quad (4.1)$$

where x is the configuration variable of period 1, p is the canonical momentum, k is the nonlinearity parameter, and $f(x+1) = f(x)$. Furthermore, we assume that the map is reversible: $f(-x) = -f(x)$ (See the section 2.4). In the following sections, the theories are illustrated with the standard map and the sawtooth map; corresponding to $f(x) = 1/2\pi \sin(2\pi x)$ and $f(x) = x$ with discontinuities at $x = \pm 1/2$, respectively.

4.2.1 Area, flux and action difference of homoclinic orbits

Let us suppose there is a curve $c(\lambda)$ in phase space which is parametrized by $\lambda \in [0,1]$. The area A under c is the integrated area under the curve; $A = \int_C p \, dx$. The integration can give a negative area if the curve is below $p = 0$; the area is algebraic. Let c' be the image of c under the map and A' be the area under c' . Then using the generating function $F(x,x')$ of the map as given by Eq. (2.7),

$$\begin{aligned} \frac{dF}{d\lambda} &= \frac{\partial F}{\partial x'} \frac{dx'}{d\lambda} + \frac{\partial F}{\partial x} \frac{dx}{d\lambda} \\ &= p' \frac{dx'}{d\lambda} - p \frac{dx}{d\lambda} \end{aligned}$$

By integrating both sides with respect to λ ,

$$\begin{aligned} \Delta F &\equiv F[x(1),x'(1)] - F[x(0),x'(0)] \\ &= \int_0^1 p' \frac{dx'}{d\lambda} d\lambda - \int_0^1 p \frac{dx}{d\lambda} d\lambda \\ &= A' - A \end{aligned}$$

(4.2)

Now suppose that two orbits $\{ Y_t \}$ and $\{ Z_t \}$ are homoclinic to each other. Namely,

$$\lim_{t \rightarrow \pm\infty} |Y_t - Z_t| = 0$$

Then, there exists a stable (unstable) segment between Y_t and Z_t which contracts (stretches) in forward (backward) time. Let c_t^s (c_t^u) be the stable (unstable) segment between Y_t and Z_t . The area A_t^s (A_t^u) under c_t^s (c_t^u) is obtained using Eq. (4.2) recursively:

$$\begin{aligned} A_t^s &= A_{t+1}^s - \Delta F_t = A_{t+2}^s - \Delta F_{t+1} - \Delta F_t = \dots \\ &= - \sum_{j=t}^{\infty} \Delta F_j \quad (\text{since } \lim_{k \rightarrow \infty} A_k^s = 0) \end{aligned} \quad (4.3)$$

Similarly,

$$A_t^u = \sum_{j=-\infty}^{t-1} \Delta F_j \quad (4.4)$$

By combining Eq. (4.3) and (4.4), we obtain the algebraic area below the unstable segment c_t^u and above the stable segment c_t^s .

$$A_t^u - A_t^s = \sum_{j=-\infty}^{\infty} \Delta F_j \equiv \Delta W \quad (4.5)$$

That is, the area is given as the difference in actions of two orbits, $\{Y_t\}$ and $\{Z_t\}$. Note that it is independent of the reference time t . This is because the map is area preserving. Equation (4.5) is used to calculate the turnstile area of partial barriers as will be shown in the next sections.

Another quantity of interest (see below) is the sum of areas, denoted by A_t , under the all forward iterates of the stable segment c_t^s and all

backward iterates of the unstable segment C_t^u . Using Eq. (4.3) and (4.4) and reordering the summations, we obtain

$$\begin{aligned}
 A_t &= \sum_{k=1}^{\infty} A_{t+k}^s + \sum_{k=-\infty}^0 A_{t+k}^u \\
 &= - \sum_{k=1}^{\infty} k \Delta F_{t+k} - \sum_{k=-\infty}^0 k \Delta F_{t+k} \\
 &= - \sum_{k=-\infty}^{\infty} k \Delta F_{t+k}
 \end{aligned} \tag{4.6}$$

Notice that A_t^s was not included in the summation. In the case that ΔW vanishes one can use A_t^s instead of A_t^u since they have the same area by Eq. (4.5). On the other hand when $\Delta W \neq 0$, the right hand side of Eq. (4.6) is not independent of the time t ; in fact, $A_{t+1} = A_t + \Delta W$. When the action difference between the two relevant orbits vanishes, A_t does not depend on the choice of origin of time. Assuming this, we will drop the subscript 't' below. As we will see below, these two equations, (4.5) and (4.6), play a central role in the transport theory.

4.2.2 Cantori and Partial Barriers

The existence of a quasiperiodic minimizing orbit for any irrational frequency ω in any twist map given by Eq. (4.1) is guaranteed by the Aubry-Mather theorem (see section 2.5). When the map is nearly integrable, these

orbits lie on invariant tori, KAM tori (the KAM theorem). In area preserving maps, an invariant torus is a perfect barrier to transport since its co-dimension is one in the ambient phase space: it has topology of a circle in the plane. An invariant torus may cease to exist above some critical value of the nonlinearity, however, there still exists a quasiperiodic orbit which now densely covers an invariant Cantor set. A Cantor set is a perfect, totally disconnected and compact set. An orbit whose projection to the configuration coordinates forms a Cantor set is called a 'cantorus' [Percival 1979]. The complement of a cantorus is a countably infinite set of gaps. A phase area element can cross the gaps by the mapping dynamics; thus cantori are not perfect barriers. But they still play a role as partial barriers because the leakage rate per iteration of the map can be quite small.

We can visualize this leakage as occurring through a special region, the so-called 'turnstile'. To construct a turnstile consider a gap on a cantorus. Denote the left endpoint by X_0 and the right by Z_0 . The action difference between the orbits $\{X_t\}$ and $\{Z_t\}$ vanishes as proved by Mather [Mather 1986]. The iterates of the gap never overlap since $\{X_t\}$ and $\{Z_t\}$ are ordered: every minimizing orbit is ordered [Aubry 1983]. Since the length of the union of all iterates of a gap should be finite, the gap length goes to zero as $t \rightarrow \pm \infty$. Thus the orbits $\{X_t\}$ and $\{Z_t\}$ are homoclinic to each other (actually to the cantorus) and, hence, there exist a stable and an unstable segment between the endpoints of the gap as discussed in the previous section. Furthermore, a cantorus typically has zero length: it can be proven that hyperbolic cantori have zero length [MacKay 1987]. Thus the

union of gaps has length one. The cantorus partial barrier is then formed by the forward iterates of a stable segment together with the backward iterates of an unstable segment (see Fig. 4.1).

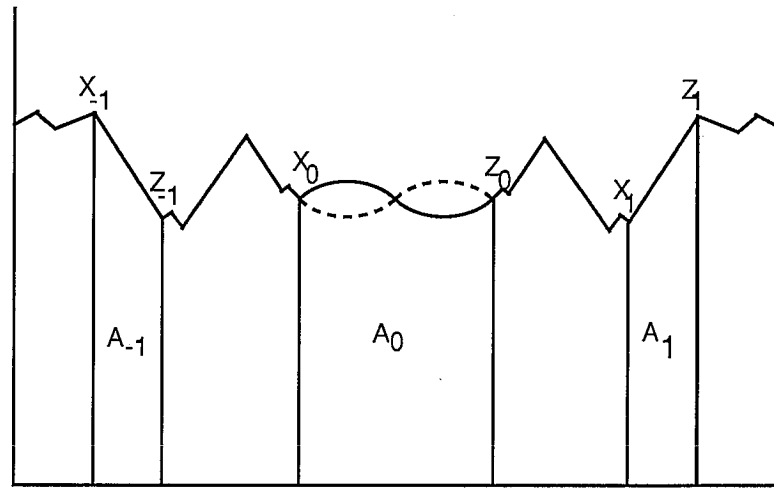


Fig. 4.1 Cantorus partial barrier. $\{X_t\}$ is the orbit of the left gap endpoint and $\{Z_t\}$ of the right gap endpoint.

The existence of a minimax orbit, denoted by $\{Y_t\}$, inside the gap is guaranteed by the fact that $\{X_t\}$ and $\{Z_t\}$ are minimizing and ordered. Since the orbit $\{Y_t\}$ is also homoclinic to the cantorus, it should lie on the intersection point of the stable and the unstable segments of each gap, which forms the center of turnstile. It is also possible that there are several minimax orbits, but the resulting structure can still be treated as one turnstile with an effective area. The area of the left (right) lobe of turnstile gives the upward (downward) flux per iteration through the cantorus partial barrier.

Using Eq. (4.5), the flux is directly obtained from the action difference between the orbits $\{ Y_t \}$ and $\{ Z_t \}$. The area under a cantorus partial barrier is the sum of areas under the gaps since the area under the cantorus itself is zero. It is obtained from the orbits $\{ X_t \}$ and $\{ Z_t \}$ by using Eq. (4.6). The lobes have the same areas since ΔW between the orbits $\{ X_t \}$ and $\{ Z_t \}$ vanishes. This also follows from zero net flux (see section 2.2).

A set of cantorus partial barriers in phase space partitions phase space into separated regions; transitions between these regions occur through the turnstiles of the partial barriers. This is discussed in section 4.3.2 and 4.3.3.

4.2.3 Partial Separatrices and Resonances

For integrable Hamiltonian systems, the stable manifold of a hyperbolic periodic point joins smoothly with the unstable manifold of a nearby hyperbolic periodic point, forming a separatrix. The motion on the separatrix, therefore, is heteroclinic between the two fixed points. Thus a separatrix is an invariant set. The simplest example is the separatrix which divides the phase space of a one-dimensional pendulum into vibrational and rotational motions.

Upon a small non-integrable perturbation, the stable and unstable manifolds typically intersect transversely and the smooth separatrix is lost. The transversality dramatically changes the structure of phase space near the broken separatrix, forming a heteroclinic tangle. In this section we

describe a way to construct a partial separatrix, which is almost invariant in the sense of transport, from pieces of stable and unstable manifolds of minimizing hyperbolic periodic points.

For the maps given by Eq. (4.1) with $k \neq 0$, there exists at least a pair of minimizing and minimax periodic orbits of a given frequency m/n . The minimizing orbit is always hyperbolic (see section 2.5) and, hence, it has stable and unstable manifolds. The minimax orbit is typically elliptic but it can be hyperbolic with reflection: the following construction holds for either case. A partial separatrix for an m/n resonance is constructed in the following way:

(1) Between two neighboring points of a given m/n minimizing orbit, M^l and M^r , there exists a m/n_+ minimizing heteroclinic orbit whose frequency limits to m/n from above. Such an orbit lies at the intersections of the right-going branch of the unstable manifold of M^l and right-going branch of the stable manifold of M^r . Choose any such intersection point, M_t^+ .

(2) Connect M_t^+ to M^l along the unstable manifold of M^l and to M^r along the stable manifold of M^r .

(3) The upper partial separatrix is obtained by connecting the $(n-1)$ successive images of the curve of step (2) together with itself.

(4) The n -th image of the curve of step (2) must intersect with the original curve, forming a turnstile. The right endpoint of the turnstile is M_{t+n}^+ . The turnstile center, S_t^+ , is a point of the minimax heteroclinic orbit.

The lower partial separatrix is obtained similarly with the heteroclinic orbit, $m/n-$, which slides to the left. For the existence of such heteroclinic orbits, see the original paper by MacKay et al [MacKay 1987]. A resonance is defined as the region bounded by the upper and lower separatrices. As an example, Fig. 4.2 shows the resonance of frequency $1/3$ for the standard map in symmetry coordinates.

To find the area under the upper partial separatrix of a resonance, choose the orbits $\{ M_t^+ \}$ and $\{ M_{t+n}^+ \}$ and apply the Eq. (4.6). The action difference between these orbits is zero since two orbits are essentially the same.

$$A^+ = - \sum_{t=-\infty}^{\infty} t [F(x_{t+n}^+, x_{t+n+1}^+) - F(x_t^+, x_{t+1}^+)] \quad (4.7)$$

The convergence of the series is guaranteed since the orbit $\{ M_t^+ \}$ is ordered and, hence, the iterates of a gap between M_t^+ and M_{t+n}^+ never overlap. To get a more practical form of Eq. (4.7), one can rearrange the summation. But it should be done carefully to maintain the convergence. By adding and subtracting the action of the periodic orbit of period n , and shifting t by n in the last term, one can obtain

$$A^+ = n \sum_{i=-\infty}^{\infty} \sum_{t=1}^n [F(x_{t+in}^+, x_{t+in+1}^+) - F(x_t^+, x_{t+1}^+)] \quad (4.8)$$

The expression for A^- is identical except for signs. The area of resonance is, therefore, given by the difference, $A^+ - A^-$.

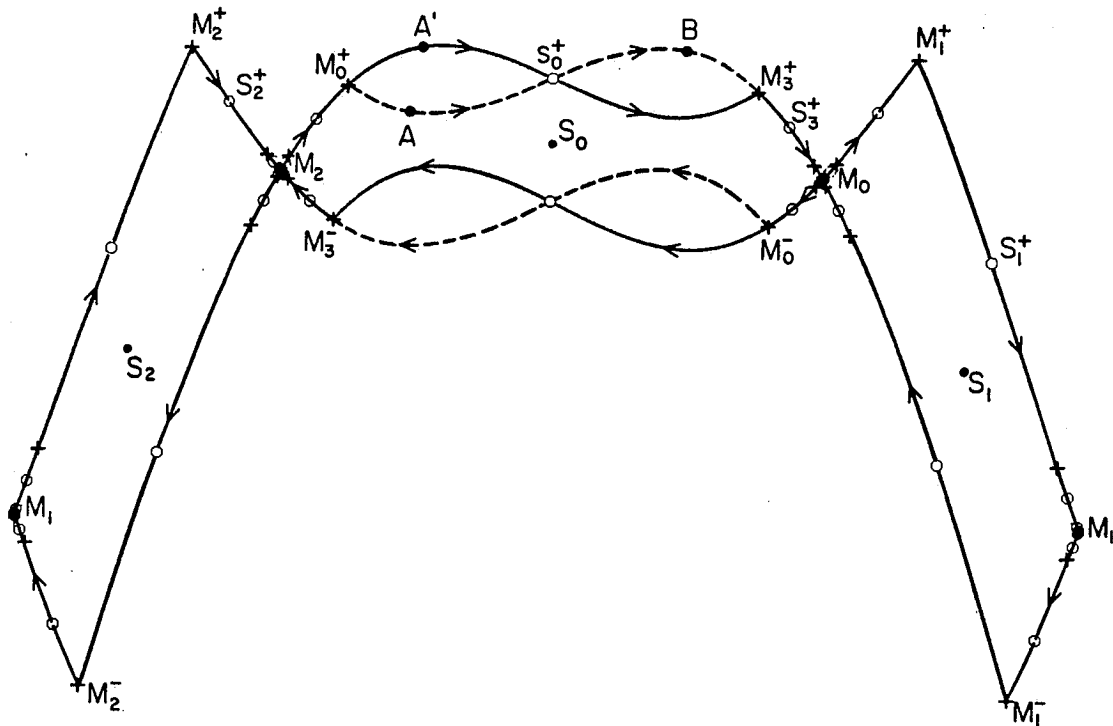


Fig. 4.2 Resonance of frequency $1/3$ for the standard map in symmetry coordinates. The resonance boundary is indicated by solid lines. Dashed lines represent the other boundaries forming the turnstiles in the partial separatrices. [Mackay 1987]

The turnstiles in the upper and lower separatrices play the same role as the one in a cantorus partial barrier. The flux exchanged through the upper separatrix is, using Eq. (4.5), the action difference between the orbits $\{ M_i^+ \}$ and $\{ S_i^+ \}$. Similarly, the flux through the lower separatrix is obtained from the orbits $\{ M_i^- \}$ and $\{ S_i^- \}$. Sum of these two fluxes gives the total flux which a single resonance exchanges with the rest of phase space per iteration of the map.

The resonances defined as above unlikely overlap one another since unstable (stable) manifolds never intersect with unstable (stable) manifolds. Importantly, it has been observed that the total area of resonances fills phase space when there exists no invariant circle. This enables us to partition phase space completely into resonances for supercritical map cases where no (rotational) invariant circle exists. This idea is discussed in section 4.3.4.

4.3 Markov Models for Transport in Area Preserving Maps

Chaotic motion is characterized by positive Lyapunov exponents. This implies that the history of such motion is rapidly lost and, hence, that the motion could be modelled to some extent as a random process. To accomplish this we first consider only the chaos of orbits in a connected chaotic region. For area preserving maps, an orbit cannot pass through invariant circles. Thus, such a connected region is bounded by invariant circles if they exist (Birkhoff's theorem). Inside such a region, however, the dynamics is more complicated than just purely random. In fact, we have seen in the previous section that there still exist partial barriers formed by cantori and broken separatrices. Furthermore, these exist even after the last invariant circle of the map is broken up, which makes the role of partial barriers even more important.

In this section, we describe several methods which attempt to partition such chaotic regions using the partial barriers developed in the previous section. Transitions between these regions will be modelled as a Markov processes. Quantities relevant to transport are obtained by solving the master equation for the probability distribution, which we do not attempt here.

4.3.1 Markov Approximation for Stochastic Motions

Consider a random motion of a particle or a phase space point, through a sequence of states labeled by an index i : for our present purpose, the index i ranges over a countable set; however, more generally one could treat i as a continuous random variable [Balescu 1975]. A random process is completely characterized by the knowledge of an infinite sequence of joint probability distribution functions, $P_n(i_n, i_{n-1}, \dots, i_1; t_n, t_{n-1}, \dots, t_1)$: P_n is the probability that a particle is in state i_k at time t_k , $1 \leq k \leq n$. In spite of this intractable infinity, in many practical cases high order joint probabilities can be fortunately expressed as a combination of lower order ones. For instance, if the process is completely uncorrelated in time (purely random), all order P_n with $n \geq 2$ are just the n -fold product of P_1 's each time. A Markov process is the next simplest process; here the second order joint probability generates all higher ones.

To proceed further it is convenient to introduce the transition probability, $p_2(i_1; t_1 \rightarrow i_2; t_2)$, which is the conditional probability that a particle is in state i_2 at time t_2 given that it was in state i_1 at time t_1 . Clearly, this satisfies the relation

$$P_2(i_2, i_1; t_2, t_1) = P_1(i_1; t_1) p_2(i_1; t_1 \rightarrow i_2; t_2) \quad (4.9)$$

Higher order transition probabilities are denoted as $p_n(i_1; t_1 \rightarrow \dots \rightarrow i_{n-1}; t_{n-1} \rightarrow i_n; t_n)$. A Markov process is defined by

$$p_n(i_1; t_1 \rightarrow \dots \rightarrow i_{n-1}; t_{n-1} \rightarrow i_n; t_n) = p_2(i_{n-1}; t_{n-1} \rightarrow i_n; t_n) \quad (4.10)$$

The essence of the Markov approximation is that upon transition from t_{n-1} to t_n , all influence of the history of the state for $t < t_{n-1}$ is lost. It should be then clear that all higher order joint probability functions are generated from the second order ones using Eq. (4.10); e.g., $P_3(i_3, i_2, i_1; t_3, t_2, t_1) = P_2(i_3, i_2; t_3, t_2) P_2(i_2, i_1; t_2, t_1) / P_1(i_2; t_2)$.

Thus, a Markov process is defined by transition probabilities together with a countable set of accessible states. The master equation governing the evolution of the probability density, $P(i; t)$, is

$$P(i; t) = \sum_j P(j; t') p(j; t' \rightarrow i; t) \quad (4.11)$$

which is obtained by summation over i_1 on both sides of Eq. (4.9). Here we drop the subscripts without confusion.

In the following sections we show three approaches to construct Markov models for transport in area preserving maps. We have already seen that the segments of stable and unstable manifolds of cantori and minimizing periodic orbits form partial barriers to transport. Basically these barriers are used to partition phase space. Thus the first step of each model construction will be to find an appropriate partition of phase space, each region of which is associated with a state. The next step is to determine the transition probability between states. These two steps define a Markov model.

4.3.2 Markov Chain via Cantori

The set of cantori is an uncountable infinite set since the irrationals are dense in the reals. Suppose there exists a countable set of cantori for which ΔW as a function of frequency has sharp local minima. Then, since ΔW gives the flux through a cantorus partial barrier, transport through such a minimizing cantorus would be much slower than through cantori with larger ΔW . Thus the minimizing cantori tend to be limiting barriers and the cantori with large ΔW between minimizing cantori can be ignored. If, furthermore, one assumes that the motion in the region bounded by two adjacent minimizing cantori is purely random, then successive transitions through minimizing cantori are uncorrelated.

This model gives a Markov partition of phase space. The regions bounded by adjacent minimizing cantori are labeled by an index i . Using the result of section 4.2.2, the transition probability from state i to state j per iteration is

$$p_{ij} = \Delta W_{ij} / A_i$$

where A_i is the area accessible to state i ; that is, the area of the connected chaotic region between adjacent minimizing cantori (Eq. (4.6)). Therefore the dynamics is described by p_{ij} over a chain consisting of all countable states i . This defines the Markov chain.

When p_{ij} is small for $i \neq j$, i.e., $p_{ii} \approx 1$, the Markov chain can be taken to have continuous time:

$$\frac{dP_j}{dt} = \sum_i P_i p_{ij} \quad (4.12)$$

where $P_i(t)$ is the particle density in state i . When we set $P_j(0) = 0$ for $j \neq i$, $P_i(t)/P_i(0)$ gives the fraction of particles remaining in state i at time t , which is called the survival probability.

Now let's consider a region near an invariant circle which bounds a chaotic region; suppose that this is the only invariant circle in some region of interest. Since it bounds the stochastic motion, we are especially interested in the dynamics in the one side of the boundary circle, let say, below the circle. Assume that this boundary circle is noble. Then the scaling results from the renormalization transformation analysis are applicable in its neighborhood. Numerical results done with standard map shows that ΔW has sharp local minimum at the frequencies of noble cantori [MacKay 1984]. Therefore, to an extent, it is a reasonable approximation to neglect all cantori but the locally minimizing ones.

Below the boundary circle there is an infinity of elliptic periodic orbits which limit to the boundary circle. Let the frequency of the boundary circle be ν . The frequencies of the periodic orbits, ν_i , are the even convergents of ν in its continued fraction expansion. The level i periodic orbit defines the state i in the following way. Between successive levels there are an infinity of cantori corresponding to irrationals in the range $\nu_i < \omega < \nu_{i+1}$. We choose the cantorus, ω_i , which has minimum ΔW in the range. Then the area bounded by the adjacent cantori ω_i and ω_{i-1} corresponds to state i . This situation is depicted in Fig. 4.3.

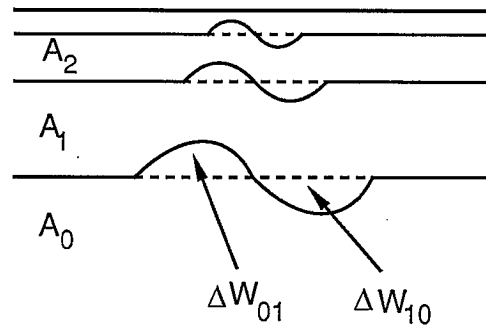


Fig. 4.3 Markov chain near a boundary circle

Using the results from the renormalization analysis, Hanson et al. show that the survival probability $F(t) (\equiv P(t) / P(0))$ of Eq. (4.12) decays algebraically [Hanson 1985]: $F(t) \sim t^{-z}$ with $z \sim 3.05$. The decay of this result is rather faster compared to the numerical result, $z \sim 1.5$, obtained by Karney and Chirikov et al. [Karney 1983, Chirikov 1984]. The most important reason for this discrepancy might be due to the approximation that the motion inside each state is random. In fact, the dynamics between two minimizing cantori is more complicated than we have assumed.

4.3.3 Markov Tree Model

The most important dynamics heretofore neglected inside a state might be the dynamics near stable islands. In fact, the structure of phase space is an infinity of self-similar structures on every scale; islands around islands ad infinitum. To include such complications, it is essential to employ

the renormalization transformation analysis, in this case called the class renormalization [Meiss 1986].

The hierarchy of orbits is associated with the 'classes' of orbits; we define a class C orbit as an ordered orbit which encircles a class $C-1$ elliptic periodic orbit. The class zero orbits are the rotational ordered orbits, for example the rotational cantori of Fig. 4.3. Between any two adjacent class zero cantori, there is an infinity of class zero elliptic periodic orbits. Every elliptic orbit is enclosed by class one boundary circle and the inside is inaccessible from outside just as the region above the class zero boundary circle is inaccessible from below. For simplicity, only one significant island chain between two adjacent minimizing cantori is considered. Outside the class one boundary circle v_1 , there is also an infinity of cantori in the range $0 < \omega < v_1$. In the same way we constructed class zero cantori, we retain only a countable set of minimizing cantori; these tori are of class one. Between any two adjacent class one cantori, there exists a most significant class one elliptic periodic orbit. The construction of minimizing cantori and boundary circles encircling significant resonances is continued ad infinitum (see Fig. 4.4).

A connected region between two adjacent class j cantori and outside the outermost class $j+1$ cantorus is defined to be a state. The dynamics is again assumed to be purely random inside such regions. Transition from one region then occurs either by a change in level through the turnstiles of adjacent class j cantorus or to higher class through the turnstile of the adjacent class $j+1$ cantorus. The transition rate is again obtained from the

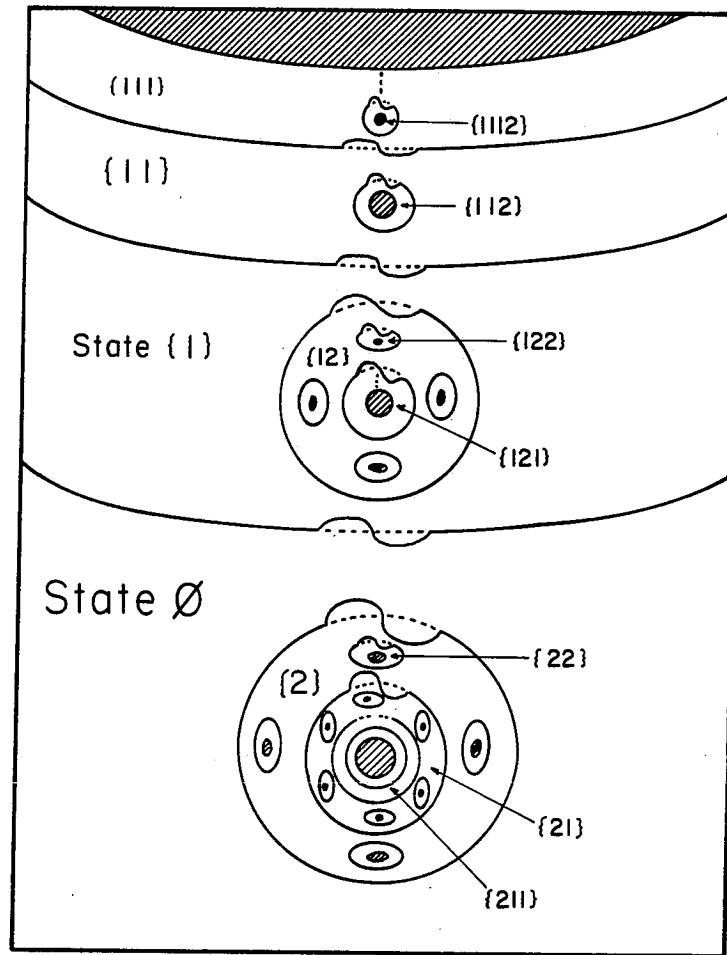


Fig. 4.4 Schematic of the connected chaotic region of an area preserving map near a boundary circle. Inaccessible areas are shaded. Boundaries of the labeled states are minimizing cantori, with their turnstiles shown. Each class one island shown represents one member of an island chain [Meiss 1986].

turnstile area through which transition occurs divided by the area of the state.

In particular, a state, S , is specified by a sequence of two symbols, say, $\{L, C\}$:

$$S = \{ \sigma_1, \sigma_2, \dots, \sigma_N \}$$

Starting from the region between level zero and level one cantori of class zero and outside the level zero cantorus of class one, denoted by \emptyset (null), σ_1 is L (C) if transition occurs in the direction increasing level (class) by one. Coding proceeds in the same way until we stop transition when we reach the state to be labeled. Thus the dynamics is depicted as random transitions on a binary tree. The transition probabilities are determined from ΔW of the cantorus between two regions divided by the area of the base state. This defines the Markov tree.

The result of Meiss and Ott gives an evident improvement compared to the Markov chain model [Meiss 1985, 1986]: $z \sim 1.96$. This model clearly takes account of the trapping behavior of an orbit near boundary circles. The model is quite valuable in that the trapping behavior seems to dominate the long time behavior of Hamiltonian systems.

The models described in section 4.3.2, 4.3.3 rely on the assumption that there exists such a countable set of flux minimizing cantori. Numerical evidence from the critical standard map implies that this is a reasonable approximation. However, these models are not generalized to the cases all fluxes are of the same order of magnitude. In particular, Mather proved that

ΔW is a continuous function of irrationals so that nearby cantori have nearly the same value of ΔW [Mather 1986].

4.3.4 Markov Chain of Resonances

A resonance is the region bounded by the upper and lower separatrices of an ordered hyperbolic periodic orbit. It has been shown that for supercritical cases the collection of resonances for all frequencies fill phase space [MacKay 1987, Chen 1987]. This fact allows a partition of phase space into a set of resonances. This partition should be more practical compared to one using cantori since a set of resonances is always countable.

The total flux exchanged per iteration by a resonance with the rest of phase space is determined by the turnstiles in the upper and lower separatrices. Transition between two resonances occurs through the overlap of the turnstiles of each resonance (see Fig. 4.5). When turnstiles overlap, the unstable manifold of lower turnstile of upper resonance m/n intersects with the stable manifold of upper turnstile of lower resonance m'/n' . Typically there are two intersection points, H_0^l and H_0^r , which are homoclinic to each other. They are heteroclinic orbits which are asymptotic to the m/n minimizing orbit in the backward time and to the m'/n' orbit in the forward time: existence of such heteroclinic connection gives a precise criterion for resonance overlap. Using Eq. (4.5), the action difference between these two orbits gives the flux between the two resonances.

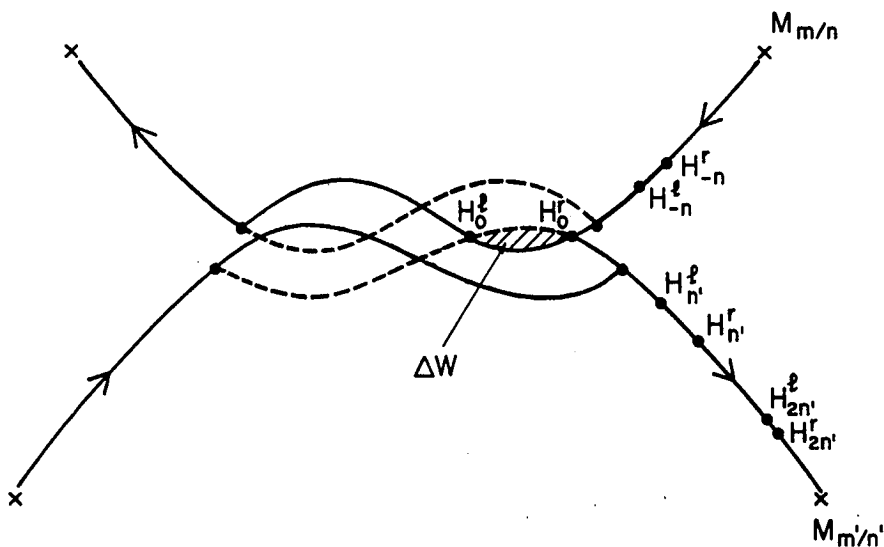


Fig. 4.5 Overlap of turnstiles [Mackay 1987]

Each resonance is a chain of islands (see Fig. 4.2). The picture described in section 4.2.3 shows that flux from a resonance with the rest of phase space occurs only through its upper and lower turnstiles of the main island: we will call the island which has turnstiles the main island. That is, no flux exchange is allowed in the other islands of the chain until its iterate is mapped to the main island. This gives a picture of flux through a resonance as far as a single isolated resonance is concerned. One should, however, be reminded that this is not the only one choice for locating a resonance. In fact, because of the area preservation, any iterate of such a resonance is a resonance with the same turnstile area as well. Thus one should be careful

in the construction of a partition of phase space with resonances to ensure that they never overlap. This kind of construction is rather simple for the reversible maps given by Eq. (4.1) because all minimax orbits appear to lie on the dominant symmetry line [MacKay 1982]. Namely, in this case the main island with turnstile is placed on the dominant line for all resonances.

A map of the form given by Eq. (4.1) is periodic in the momentum. We denote by Ω_l a unit cell in the range $l \leq p \leq l+1$ for integer l and choose in Ω_0 a set of R resonances with frequencies m_r/n_r , $r = 1, \dots, R$. The corresponding resonances in Ω_l are the translates of resonances in Ω_0 , the frequencies of which are given by $m_r/n_r + l$. For a resonance m_r/n_r , n_r islands are labeled by s , $s = 0, 1, \dots, n_r - 1$, where island s is the s -th iterate of the main island $s = 0$. Each island is called a state, labeled by $(r, s; l)$. The transition probabilities are obtained as follows.

Let A_r be the area of the one of the island for resonance r . Denote by $O(r, r'; l - l')$ the area of overlap of the turnstiles between islands $(r, 0; l)$ and $(r', 0; l')$, where $r \neq r'$. Its translational invariance is due to periodicity of the map. Define $O(r, r; 0)$ as the area of island $(r, 0; l)$ without its upper and lower turnstiles, and $\bar{A}_r \equiv \sum_{r', r} O(r, r'; l')$. For supercritical cases the partition by resonances becomes complete and, hence, $\bar{A}_r \rightarrow A_r$ as $R \rightarrow \infty$.

Then the following statements should be clear.

- (1) A transition from $(r, s; l)$ to $(r' \neq r, s'; l')$ is possible only when $s=0$ and $s'=1$.
- (2) The remaining area of $(r, 0; l)$ in (1), $O(r, r; 0)$, is always mapped to $(r, 1; l)$.
- (3) A transition from $(r, s \neq 0; l)$ occurs only to $(r, s + 1; l)$.

Thus the transition probability from a state $(r, s; l)$ to a state $(r', s'; l')$ is determined as

$$p(r, s \rightarrow r', s'; l - l') = \frac{O(r; r'; l - l')}{\bar{A}_r} \delta_{s,0} \delta_{s',1} (1 - \delta_{r,t}) \\ + \frac{O(r; r; 0)}{\bar{A}_r} \delta_{s,0} \delta_{s',1} \delta_{r,t} + (1 - \delta_{s,0}) \delta_{s',s+1} \delta_{r,t}$$

This defines the Markov model for dynamics as a random process on a chain of resonances. Note that \bar{A}_r was used instead of A_r even though the former is a disconnected area contrary to the Markov approximation. This is to obtain more consistent results for practical cases where only a finite set of resonances (finite R) is considered.

Refinements of the above partition can be defined as follows: Under n -th iteration of the map the turnstile of a period n resonance is evolved along the direction of stable manifold, forming another turnstile. It is noticeable that the incoming lobe of this n -th turnstile decreases the effective area of the predefined resonance. Thus we see that there is a mechanism that the transition rate is modified every n -th iteration. This mechanism seems to be very complicated since it acts on every resonance of different period and the iterates of incoming lobe can even be outside the predefined resonance. This mechanism, however, would decrease the survival probability when we retain the Markov approximation in each refined resonance.

This model was first solved by Dana et al [Dana 1989]. They applied the model to the sawtooth map example for which all heteroclinic orbits can

be obtained analytically. The scaling of diffusion coefficient in k around criticality ($k_c = 0$) was shown to be in good agreement with numerical experiment. For large k , the prediction of the theory is visibly below the result of Cary and Meiss (see chapter 5) as well as numerical results. However, they found a significant improvement when they considered the refined partition of resonance after n iterations.

Practical implementation of this method to the general maps might have difficulties because one has to provide the areas of overlapping turnstiles. To determine them analytically, one has to find all heteroclinic orbits which lie at the intersections of stable and unstable manifolds of minimizing periodic orbits in consideration, for which there is no generally reliable method as yet [Chen 1988].

4.4 Invariant Manifolds of $N+1$ d.o.f Hamiltonian Systems

The models developed in section 4.3 rely much on the underlying orbit dynamics; existence of minimizing orbits and their ordering properties, and existence of cantori, etc. The most important distinctive feature of area preserving maps might be that minimizing hyperbolic orbits have stable and unstable manifolds of co-dimension one and therefore it is possible to partition phase space with their segments. In general, however, for $2N$ -dimensional symplectic maps stable and unstable manifolds of hyperbolic fixed points are N -dimensional and, hence, are not of co-dimension one. Consequently, their segments don't divide phase space. In fact, even KAM tori are not of co-dimension one; an $(N+1)$ -dimensional torus cannot divide a $(2N+1)$ -dimensional energy surface. Hence they are not perfect barriers to transport any more. Furthermore, as pointed out in section 2.5, the minimizing condition should be weakened and generally the ordering property is not guaranteed. Therefore, it seems unlikely that the ideas established in area preserving maps would extend to higher dimensional cases.

Nevertheless, it is possible that there might exist some invariant manifolds of co-dimension one in higher dimensional phase space so that the phase space could be partitioned with their segments. Recently it has been shown that a formulation for transport similar to the area preserving case is possible for higher dimensional systems under certain special conditions [Rom-Kedar 1988, Wiggins 1989]. Namely, they showed that in

systems satisfying some conditions, there exists a set analogous to hyperbolic fixed points in area preserving maps with stable and unstable manifolds of co-dimension one. Proofs of the existence and persistence of such manifolds rely on preceding sophisticated theorems. However, once such manifolds are shown to exist, the transport theory will be constructed in the very similar way as in area preserving maps. We remark that the practicality of this theory is still not determined.

4.4.1 Nearly Integrable Hamiltonian Systems of $N+1$ degrees of freedom

Let us consider $N+1$ d.o.f Hamiltonian systems of the form

$$H(x, I, \theta; \mu, \varepsilon) = H_0(x, I) + \varepsilon \tilde{H}(x, I, \theta; \mu, \varepsilon) \quad (4.13)$$

where $(x, I, \theta) \in \mathcal{R}^2 \times \mathcal{R}^N \times \mathcal{T}^N$, and $\mu \in \mathcal{R}^p$ is a vector parameter. J is the antisymplectic matrix as given in Eq. (2.9). \tilde{H} represents a small nonintegrable perturbations to the integrable Hamiltonian H_0 ; $0 < \varepsilon \ll 1$. Further, we assume that the Hamiltonian functions, H_0 and \tilde{H} , are C^{r+1} , $r \geq 1$.

First, let us consider the motions in the unperturbed system of Eq.(4.13)

$$\begin{aligned} \dot{x} &= J D_x H_0(x, I) \\ \dot{I} &= 0 \\ \dot{\theta} &= D_\theta H_0(x, I) \end{aligned} \quad (4.14)$$

The motion has one degree-of-freedom in x with trivial rotation in (I, θ) and, hence, is totally integrable. We will have two more assumptions as follows.

(1) There exists an open set $U \subset \mathcal{R}^N$ such that for each $I \in U$ the x component of the motion has a hyperbolic fixed point which varies smoothly with I , denoted by $\gamma(I)$, which has an orbit $x^I(t)$ homoclinic to $\gamma(I)$: that is,

$$D_x H_0(\gamma(I), I) = 0 \quad \text{and} \quad \lim_{t \rightarrow \pm\infty} x^I(t) = \gamma(I) \quad \text{for } I \in U$$

(2) $D_I H_0(x, I) \neq 0$

Now consider the set of points \mathcal{M} in $\mathcal{R}^2 \times \mathcal{R}^N \times \mathcal{T}^N$ defined by

$$\mathcal{M} = \{ (x, I, \theta) \in \mathcal{R}^2 \times \mathcal{R}^N \times \mathcal{T}^N \mid x = \gamma(I), I \in U, \theta \in \mathcal{T}^N \}$$

The unperturbed vector field restricted to \mathcal{M} is given by

$$\begin{aligned} \dot{x} &= 0 \\ \dot{\theta} &= D_I H_0(\gamma(I), I), \quad (I, \theta) \in U \times \mathcal{T}^N \end{aligned}$$

So \mathcal{M} has the structure of an N -parameter family of N -tori with the flow on the tori being either commensurate or incommensurate depending on $D_I H_0(\gamma(I), I)$; Chirikov calls these "whiskered tori" [Chirikov 1979]. The corresponding torus for a fixed $I = I'$ is denoted as

$$\tau(I') \equiv \{ (x, I, \theta) \in \mathcal{R}^2 \times U \times \mathcal{T}^N \mid x = \gamma(I'), I = I', \theta \in \mathcal{T}^N \}$$

From the above first assumption of the homoclinic motion in x , it is clear that each torus has $N+1$ dimensional stable and unstable manifolds.

4.4.2 Existence and Persistence of Invariant Manifolds

The proposition given below accounts for the existence of relevant invariant manifolds for the unperturbed system given in Eq. (4.14). It is followed by several remarks on its implications.

Proposition A. [Wiggins 1988] \mathcal{M} is a C^r $2N$ dimensional normally hyperbolic invariant manifold of (4.14). Moreover, \mathcal{M} has C^r $2N+1$ dimensional stable and unstable manifolds denoted by $W^s(\mathcal{M})$ and $W^u(\mathcal{M})$, respectively, which intersect in the $2N+1$ dimensional manifold of homoclinic orbits

$$\Gamma = \{ (x^l(-t_0), l, \theta_0) \in \mathcal{R}^2 \times \mathcal{R}^N \times \mathcal{T}^N \mid (t_0, l, \theta_0) \in \mathcal{R}^1 \times U \times \mathcal{T}^N \}$$

Proof: The proof is trivial from the previous section. About the normal hyperbolicity of \mathcal{M} , see Wiggins [1988].

Remarks:

(1) Roughly speaking, the normal hyperbolicity means that the directions normal to \mathcal{M} are expanded or contracted more sharply than the directions tangent to \mathcal{M} . It is characterized by the generalized Lyapunov type numbers [Wiggins 1988]. It will turn out that in their role to transport the normally hyperbolic invariant manifolds are higher dimensional analog of hyperbolic fixed points in area preserving maps. Importantly, normally

hyperbolic invariant manifolds persist under perturbation as invariant manifolds.

(2) The stable and unstable manifolds have co-dimension one. So an assembly of the segments of such manifolds can separate the ambient $2N+2$ dimensional phase space.

(3) $W^s(\mathcal{M})$, $W^u(\mathcal{M})$, \mathcal{M} and, hence, $\Gamma \equiv W^s(\mathcal{M}) \cap W^u(\mathcal{M}) - \mathcal{M}$ intersect transversely the constant energy levels $H_0(x, l) \equiv h$ provided $D_l H_0(x, l) \neq 0$. Furthermore, $\mathcal{M} \cap h$ is diffeomorphic to S^{2N-1} which is compact and boundaryless; these two are the lemmas proven by Wiggins [Wiggins 1989]. This implies that $(\Gamma \cap h) \cup (\mathcal{M} \cap h)$ divides the energy level surface $H_0(x, l) \equiv h$ into an inside and an outside.

(4) In the standard way (see section 1.3), one can construct the Poincaré surface section, Σ , which is transverse to the constant energy levels. Let the intersection $\Sigma \cap h$ be Σ_h and the return map restricted to Σ_h be P_h . Σ_h is $2N$ dimensional and, thus, P_h is a $2N$ dimensional symplectic map. The invariant manifolds in (3) are also transverse to Σ_h . Then we have

$$\mathcal{M} \cap \Sigma_h \quad \text{is } 2N-2 \text{ dimensional}$$

$$W^s(\mathcal{M}) \cap \Sigma_h \quad \text{is } 2N-1 \text{ dimensional}$$

$$W^u(\mathcal{M}) \cap \Sigma_h \quad \text{is } 2N-1 \text{ dimensional.}$$

Likewise, $\mathcal{M} \cap \Sigma_h$ is compact and boundaryless and $(\Gamma \cap \Sigma_h) \cup (\mathcal{M} \cap \Sigma_h)$ divides Σ_h into inside and an outside. Thus $(\Gamma \cap \Sigma_h) \cup (\mathcal{M} \cap \Sigma_h)$ is a barrier to transport. For $N = 1$ (area preserving maps), $\mathcal{M} \cap \Sigma_h$ is a hyperbolic fixed point. For $N = 2$ (four-dimensional maps), $\mathcal{M} \cap \Sigma_h$ is a one parameter family of one tori, i.e., two-dimensional cylinder.

Now we go back to the perturbed system, Eq. (4.13). The question is whether those invariant manifolds of the integrable system, Eq. (4.14), still exist under perturbation. The answer is given by the proposition below.

Proposition B. [Wiggins 1988] There exists $\varepsilon_0 > 0$ such that for $0 < \varepsilon \leq \varepsilon_0$ the perturbed system (4.13) possesses a C^r $2N$ dimensional normally hyperbolic locally invariant manifold

$$\mathcal{M}_\varepsilon = \{ (x, I, \theta) \in \mathcal{R}^2 \times \mathcal{R}^N \times \mathcal{T}^N \mid x = \tilde{\gamma}(I, \theta; \varepsilon) = \gamma(I) + \alpha(\varepsilon),$$

$$I \in \tilde{U} \subset U \subset \mathcal{R}^N, \theta \in \mathcal{T}^N \}$$

where $\tilde{U} \subset U$ is a compact, connected N dimensional set. Moreover, \mathcal{M}_ε has local C^r stable and unstable manifolds, $W_{\text{loc}}^s(\mathcal{M}_\varepsilon)$ and $W_{\text{loc}}^u(\mathcal{M}_\varepsilon)$, which are of the same dimension and C^r close to $W_{\text{loc}}^s(\mathcal{M})$ and $W_{\text{loc}}^u(\mathcal{M})$, respectively.

Proof: The proof of the proposition uses Fenichel's theorem on the existence and persistence of the overflowing and inflowing invariant manifolds under the C^r vector field. See Wiggins [1988].

Remarks:

(1) Transversal intersections are structurally stable under perturbations. Therefore, h_ε intersects $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ transversely, Σ intersects $W^s(\mathcal{M}_\varepsilon)$ and $W^u(\mathcal{M}_\varepsilon)$ transversely, and Σ intersects h_ε transversely.

Then we have

$\mathcal{M}_\varepsilon \cap \Sigma_{h_\varepsilon} \equiv \widehat{\mathcal{M}}_\varepsilon$ is $2N-2$ dimensional

$W^s(\mathcal{M}_\varepsilon) \cap \Sigma_{h_\varepsilon} \equiv W^s(\widehat{\mathcal{M}}_\varepsilon)$ is $2N-1$ dimensional

$W^u(\mathcal{M}_\varepsilon) \cap \Sigma_{h_\varepsilon} \equiv W^u(\widehat{\mathcal{M}}_\varepsilon)$ is $2N-1$ dimensional.

(2) Dynamics on $\widehat{\mathcal{M}}_\varepsilon$ might be quite complicated [Graff 1974]. In fact the KAM theorem applies to the survival of most nonresonant tori of Eq. (4.14) on $\widehat{\mathcal{M}}_\varepsilon$.

4.4.3 Transversal Homoclinic Manifold and Homoclinic Tangle

Suppose $W^s(\widehat{\mathcal{M}}_\varepsilon)$ and $W^u(\widehat{\mathcal{M}}_\varepsilon)$ intersect transversely. Then the intersection, $W^s(\widehat{\mathcal{M}}_\varepsilon) \cap W^u(\widehat{\mathcal{M}}_\varepsilon)$, is $2N-2$ dimensional: transversality means that $\dim(T_p W^s(\widehat{\mathcal{M}}_\varepsilon) + T_p W^u(\widehat{\mathcal{M}}_\varepsilon)) = \dim(T_p W^s(\widehat{\mathcal{M}}_\varepsilon)) + \dim(T_p W^u(\widehat{\mathcal{M}}_\varepsilon)) - \dim(T_p W^s(\widehat{\mathcal{M}}_\varepsilon) \cap T_p W^u(\widehat{\mathcal{M}}_\varepsilon))$, where $T_p W^s(\widehat{\mathcal{M}}_\varepsilon)$ is the tangent space of $W^s(\widehat{\mathcal{M}}_\varepsilon)$ at $p \in W^s(\widehat{\mathcal{M}}_\varepsilon) \cap W^u(\widehat{\mathcal{M}}_\varepsilon)$ and similarly for $W^u(\widehat{\mathcal{M}}_\varepsilon)$. An additional requirement for the transversal intersection of stable and unstable manifolds in $N \geq 2$ cases is that the intersection should be homotopic to $\widehat{\mathcal{M}}_\varepsilon$. This is because the intersection for our purpose should be the one which is asymptotic to $\widehat{\mathcal{M}}_\varepsilon$, rather than to just a part of $\widehat{\mathcal{M}}_\varepsilon$. An intersection satisfying these two requirements of transversality and homotopy is referred to as a 'transversal homoclinic manifold'.

To determine the existence of a transversal intersection of $W^s(\widehat{\mathcal{M}}_\varepsilon)$ and $W^u(\widehat{\mathcal{M}}_\varepsilon)$, a Melnikov type method can be used [Wiggins 1988]; he shows that the distance function of the usual Melnikov method for planar homoclinic orbits generalizes to $N \geq 2$ cases. Furthermore, the method results in criteria

for the existence of a transversal homoclinic manifold satisfying homotopy condition.

Once such a transversal homoclinic manifold exists, the existence of homoclinic tangle similar to the one in planar maps is guaranteed by the invariance of $W^s(\widehat{\mathcal{M}}_\varepsilon)$ and $W^u(\widehat{\mathcal{M}}_\varepsilon)$; the λ -lemma is also extendable to $N \geq 2$ cases [Guckenheimer 1983] [Wiggins 1988]. Thus we obtain a picture of lobes and turnstiles constructed with segments of stable and unstable manifolds, which is very similar to the area preserving map case.

A region in phase space now can be defined as a phase volume bounded by segments of $W^s(\widehat{\mathcal{M}}_\varepsilon)$ and $W^u(\widehat{\mathcal{M}}_\varepsilon)$. In the same way as for area preserving maps, the transition rate between regions per iteration might be given as the ratio of lobe volume to the volume of the region, if they are known. Rom-Kedar and Wiggins have presented a transport theory based on the lobe dynamics, which we don't attempt show in detail [Rom-Kedar 1989, Wiggins 1989]. Their theory is slightly different from the ones shown in section 4.3 in that their theory takes account of the entire geometry of homoclinic orbits in terms of all iterates of turnstiles. However, how to calculate them along infinite time seems to be an intractable task. They show several numerical techniques for area preserving maps, to which the action function method of section 4.3 is also applicable.

4.5 Discussion

The Markov models for transport described in section 4.3 are fairly successful in describing transport for area preserving maps. Their construction of the models relies much on the details of underlying dynamics. However, they still have some limits in that the dynamics in some region of phase space, depending on the choice of partitions, is blurred out in terms of the Markov approximation. In fact, the geometry of Hamiltonian phase space is infinitely complicated at least in two respects: the self-similar structure on every scale and the endless stretching of unstable manifolds due to symplecticity. Thus the deviation of the theoretical predictions from experimental results seems due to neglect of geometric structure on finer scales. In this sense, despite its several weakness, the Markov tree model has a more transparent connection than the others to the actual dynamics since it takes in consideration an infinity of classes each of which is again an infinity of levels.

An extension of the ideas employed in area preserving maps to higher dimensional cases is not yet possible since the orbit dynamics are quite different in each case. The theory described in section 4.4 suggests a rather natural generalization of the idea of partitioning phase space with segments of co-dimension one manifolds. However, their transport theory based on the lobe dynamics has difficulties in practical implementation and it seems questionable whether the existence of such normally hyperbolic invariant manifolds is typical in more general systems; for instance, such

manifolds do not seem to exist in Freoschle's example (eq. 2.1) with any finite coupling parameter h (see figures 2.10 and 2.11). Furthermore, their theory does not have a transparent connection to the symplectic nature of dynamics which is expected to be the most important restriction to the motions in phase space; preservation of the phase volume is the consequence of preservation of the symplectic form (see section 1.2 and also 2.2).

CHAPTER 5

STATISTICAL METHOD FOR DIFFUSION IN SYMPLECTIC MAPS

5.1. Introduction

We have seen in the previous chapters that symplectic maps are generically not integrable, but also not ergodic. The rigorous treatment of the subset of apparently chaotic orbits by statistical methods is a long sought, but elusive goal. One statistical property of interest is the diffusion coefficient. This has been extensively studied for area preserving maps. Notably, for the standard map, Chirikov showed numerically that the momentum diffusion of an ensemble of particles approaches the "quasilinear" value as the nonlinearity parameter increases [Chirikov 1979]. Quasilinear diffusion results if the force is a delta correlated random process. For the standard map, the approach to quasilinearity is not monotonic: the oscillations of the diffusion coefficient were explained by Rechester et al using an expansion called the Fourier path method [Rechester 1980, 1981]. This was extended and modified to apply to many maps using the characteristic function formalism [Cary 1981, Meiss 1983]. Here it was shown that the oscillations are due to short time correlations. The characteristic function method allows the exact calculation of the diffusion coefficient for Arnol'd's "cat maps" [Cary 1981, Antonsen 1981], and can be used for maps with nonlinear twist terms [Murray 1985, Hatori 1985, Lichtenberg 1987].

In general these expressions for the diffusion coefficient agree well with numerical experiments, providing the measurement time is moderate.

As the time is increased, long time tails in the correlation functions become increasingly important [Karney 1983, Meiss 1986], and the diffusion coefficient develops sharp peaks [Karney 1982, Ichikawa 1987]. The height of the peaks grows with the computation time; thus the series expressions for D are not convergent. These anomalies are due to the presence of stable "accelerator modes" in the standard map. We will discuss these further in 5.6 for the four dimensional case.

In this paper we generalize the characteristic function formalism for mappings of the form

$$\mathbf{x}_{t+1} - 2\mathbf{x}_t + \mathbf{x}_{t-1} = -\nabla V(\mathbf{x}_t) , \quad (5.1)$$

Here \mathbf{x} is an N -dimensional configuration vector, and $V(\mathbf{x})$ is an arbitrary periodic function of \mathbf{x} : $V(\mathbf{x} + 2\pi\mathbf{m}) = V(\mathbf{x})$, for any integer vector \mathbf{m} . Defining the momentum coordinate by

$$\mathbf{y}_t = \mathbf{x}_t - \mathbf{x}_{t-1} \quad (5.2)$$

this map is easily shown to be symplectic (it preserves the area $\oint \mathbf{y} \cdot d\mathbf{x}$). Since V is periodic, the configuration space can be taken to be the N -torus. We allow arbitrary values for \mathbf{y} , thus the phase space is $\mathcal{R}^N \times \mathcal{T}^N$.

For an example we again use a four-dimensional (two degree-of-freedom) map first studied by Froeschlé [Froeschlé 1972, 1973].

$$V(\mathbf{x}) = -a \cos(x^1) - b \cos(x^2) - c \cos(x^1+x^2) \quad (5.3)$$

where 'a' and 'b' represent the strengths of standard maps for each degree of freedom, and 'c' couples these together. The map appears not to have any global invariants unless all but one of the three parameters are zero. Many of the expressions we will give below will be for the simple case $b=0$, but this case does not appear to be special in any way.

5.2. Diffusion tensor

A statistical analysis describes the behavior of an ensemble of initial conditions. We choose a uniform distribution over some invariant region in phase space, denoted as \mathfrak{R} . It is convenient to use the Lagrangian form of the map, eq. (5.1), and to average over an ensemble of initial conditions ($\mathbf{x}_0, \mathbf{x}_1$). Averages over this set of initial conditions are denoted $\langle \rangle_{\mathfrak{R}}$. Typically we will assume the configuration is arbitrary, so that \mathfrak{R} is $\mathcal{T}^N \times \mathcal{T}^N$. Ideally, one would like to use a uniform distribution on the set of chaotic initial conditions for the ensemble, since the regular initial conditions do not diffuse. However, we know of no analytical techniques to determine this ensemble (which in the two dimensional case is a fat fractal [Umberger 1985]), nor even to show that it occupies a set of positive measure for a typical map.

The momentum diffusion tensor is defined to be the asymptotic rate of spread of the second moment of the momentum distribution

$$\mathbf{D} = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle \Delta \mathbf{y}_t \Delta \mathbf{y}_t \rangle_{\mathfrak{R}} \quad (5.4)$$

where $\Delta \mathbf{y}_t(\mathbf{x}_0, \mathbf{x}_1)$ is the momentum change in t iterations for an initial condition $(\mathbf{x}_0, \mathbf{x}_1)$. The factor of two is due to the definition of diffusion in the

Fokker-Planck equation [Lichtenberg 1987]. Using (1) and (2), this can be easily rewritten as

$$\mathbf{D} = \frac{1}{2} \mathbf{C}(0) + \sum_{t=1}^{\infty} \mathbf{C}(t)$$

$$\mathbf{C}(t) \equiv \left\langle \nabla V(\mathbf{x}_0) \nabla V(\mathbf{x}_t) \right\rangle_{\mathfrak{R}} \quad (5.5)$$

where \mathbf{C} , the force correlation tensor, has been assumed to decay more rapidly than t^{-1} . Finally, the correlation function can be expressed in terms of the Fourier transform of ∇V

$$\nabla V = \sum_{\mathbf{m}=-\infty}^{\infty} \mathbf{f}_{\mathbf{m}} e^{i\mathbf{m} \cdot \mathbf{x}} \quad (5.6)$$

where $\mathbf{f}_{\mathbf{m}} = \mathbf{f}_{-\mathbf{m}}^*$, so that

$$\mathbf{C}(t) = \sum_{\mathbf{m}, \mathbf{n}=-\infty}^{\infty} \mathbf{f}_{\mathbf{m}} \mathbf{f}_{\mathbf{n}} \left\langle \exp(i\mathbf{m} \cdot \mathbf{x}_t + i\mathbf{n} \cdot \mathbf{x}_0) \right\rangle_{\mathfrak{R}} \quad (5.7)$$

Since the region \mathfrak{R} is invariant under the mapping and the mapping is volume preserving, Eq. (5.7) implies the time reversal symmetry $\mathbf{C}(t) = \tilde{\mathbf{C}}(-t)$, where “ \sim ” designates transpose.

For the example of eq. (5.3) the nonzero $\mathbf{f}_{\mathbf{m}}$ are

$$\mathbf{f}_{(1,0)} = (a/2i, 0), \quad \mathbf{f}_{(0,1)} = (0, b/2i), \quad \mathbf{f}_{(1,1)} = (c/2i, c/2i), \quad (5.8)$$

plus their complex conjugates.

The task is now to obtain expressions for the ensemble average of the exponential appearing in eq. (5.7).

5.3. Characteristic Function Method

Any statistical property of the ensemble can be obtained from the joint probability distribution $P(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t)$ which is the probability that the trajectory lands at the points \mathbf{x}_k for $k=0, 1, \dots, t$. The t^{th} -characteristic function is defined to be the Fourier transform of this distribution [Cary 1981, Meiss 1983]:

$$\chi_t(\mathbf{n}_t, \mathbf{n}_{t-1}, \dots, \mathbf{n}_1, \mathbf{n}_0) = \left\langle \exp(i \sum_{j=0}^t \mathbf{n}_j \cdot \mathbf{x}_j) \right\rangle_{\mathfrak{R}} \quad (5.9)$$

where the \mathbf{n}_t are arbitrary integer vectors. In particular, the correlation function depends only on the set of characteristic functions with the first and last arguments nonzero. Thus it is convenient to define the reduced function

$$\chi_t(\mathbf{n}_t, \mathbf{n}_0) \equiv \chi_t(\mathbf{n}_t, \mathbf{0}, \dots, \mathbf{0}, \mathbf{n}_0) \quad (5.10)$$

Using eq. (5.7) with this, the correlation function is

$$\mathbf{C}(t) = \sum_{\mathbf{m}, \mathbf{n}=-\infty}^{\infty} \mathbf{f}_m \mathbf{f}_n \chi_t(\mathbf{m}, \mathbf{n}) \quad (5.11)$$

For the potential (5.3), one can form linear combinations of the components of the correlation tensor to separate the contributions explicitly due to each term in (5.3) independently, and those due to coupling between them. For example, when $b=0$ we define

$$\begin{aligned}
C_{aa}(t) &\equiv C^{11} - 2C^{12} + C^{22} = -1/2 a^2 \{ \chi_t[(1,0), (1,0)] - \chi_t[(-1,0), (1,0)] \} \\
C_{ac}(t) &\equiv C^{12} - C^{22} = -1/2 ac \{ \chi_t[(1,0), (1,1)] - \chi_t[(-1,0), (1,1)] \} \\
C_{cc}(t) &\equiv C^{22} = -1/2 c^2 \{ \chi_t[(1,1), (1,1)] - \chi_t[(-1,-1), (1,1)] \}
\end{aligned}
\tag{5.12}$$

Note, however, that each of the χ_t in (5.12) depends implicitly on both parameters 'a' and 'c'.

Evaluation of χ_t requires knowledge of the exact orbit beginning at an arbitrary point $(\mathbf{x}_0, \mathbf{x}_1)$, and is impossible in general. However, a remarkably simple recursion relation can be obtained, using eq. (1), which relates χ_t to χ_{t-1} :

$$\chi_t(\mathbf{n}_t, \mathbf{n}_{t-1}, \dots, \mathbf{n}_1, \mathbf{n}_0) = \sum_{\mathbf{m}} g_{\mathbf{m}}(\mathbf{n}_t) \chi_{t-1}(\mathbf{n}_{t-1} + 2\mathbf{n}_t - \mathbf{m}, \mathbf{n}_{t-2} - \mathbf{n}_t, \mathbf{n}_{t-3}, \dots, \mathbf{n}_1, \mathbf{n}_0)
\tag{5.13}$$

Here we have defined the Fourier transform of the exponential of ∇V :

$$g_{\mathbf{m}}(\mathbf{n}) \equiv \frac{1}{(2\pi)^N} \int_0^{2\pi} d^N x \exp \left[i \left(\mathbf{m} \cdot \mathbf{x} - \mathbf{n} \cdot \nabla V(\mathbf{x}) \right) \right]
\tag{5.14}$$

for an integer vector \mathbf{m} . For the Froeschlé example, $g_{\mathbf{m}}$ can be evaluated in terms of Bessel functions:

$$g_{\mathbf{m}}(\mathbf{n}) = \sum_{j=-\infty}^{\infty} J_{m^1-j}(n^1 a) J_{m^2-j}(n^2 b) J_j[(n^1 + n^2)c] \quad (5.15a)$$

When $b=0$ the infinite sum in eq. (5.15a) collapses to a single term

$$g_{\mathbf{m}}(\mathbf{n}) = J_{m^1-m^2}(n^1 a) J_{m^2}[(n^1 + n^2)c] \quad (b=0) \quad (5.15b)$$

The recursion relation (5.13) is easily formally iterated to obtain χ_t in terms of χ_1 . In particular for the class of eq. (5.10) we obtain the exact, if somewhat formal, expression:

$$\chi_t(\mathbf{n}_t, \mathbf{n}_0) = \sum_{\mathbf{n}_1, \dots, \mathbf{n}_{t-1}} \prod_{k=2}^t g_{\mathbf{m}_k}(\mathbf{n}_k) \chi_1(\mathbf{n}_1, \mathbf{n}_0 - \mathbf{n}_2) \quad (5.16a)$$

where the \mathbf{m}_k for $k = 2, 3, \dots, t$ are defined by

$$\mathbf{m}_k = -\delta^2 \mathbf{n}_k \equiv -\mathbf{n}_{k+1} + 2\mathbf{n}_k - \mathbf{n}_{k-1} \quad (5.16b)$$

with $\mathbf{n}_{t+1} \equiv 0$. Equation (5.16) is exactly the same as for the area preserving case [Cary 1981] except that the integers m_t and n_t have been replaced by integer vectors, and we have reversed the sign of \mathbf{m} and the ordering of the index "k" for later convenience.

If we assume that \mathfrak{R} is $\mathcal{T}^N \times \mathcal{T}^N$, then the first characteristic function is simply a Kronecker delta

$$\chi_1(\mathbf{m}, \mathbf{n}) = \delta_{\mathbf{m},0} \delta_{\mathbf{n},0} \quad (5.17)$$

Using this in eq. (5.16) gives the symmetry property $\chi_t(\mathbf{m}, \mathbf{n}) = \chi_t(\mathbf{n}, \mathbf{m})$. The definition (5.11) then implies that the correlation matrix is symmetric, $\mathbf{C}(t) = \tilde{\mathbf{C}}(t)$, and therefore so is the diffusion matrix.

When we set $b=0$ in the Froeschlé map and use $\mathfrak{R} = \mathcal{T}^N \times \mathcal{T}^N$, the first three χ_t can be written explicitly:

$$\begin{aligned} \chi_2(\mathbf{m}, \mathbf{n}) &= \delta_{\mathbf{m},\mathbf{n}} J_{2(n^1-n^2)}(an^1) J_{2n^2}[c(n^1+n^2)] \\ \chi_3(\mathbf{m}, \mathbf{n}) &= J_{m^2-m^1+2(n^1-n^2)}(an^1) J_{n^2-n^1+2(m^1-m^2)}(am^1) \times \\ &\quad J_{2n^2-m^2}[c(n^1+n^2)] J_{2m^2-n^2}[c(m^1+m^2)] \end{aligned} \quad (5.18)$$

The next characteristic function, χ_4 , is an infinite sum over an intermediate index of a product of six Bessel functions. In general when all three parameters a, b, c are non-zero, each χ_t contains sums over products of $3t$ Bessel functions, though it is possible that some of these could turn out to be $J_0(0) = 1$. In fact, as we will see below, terms with a maximal number of such factors dominate when the parameters are large.

5.4 Principal Terms

Each term in the sum of eq (5.16) can be viewed as a particular path in the space (\mathbf{n}, k) . Once the path is chosen, then the \mathbf{m}_k are determined by the second differences through eq (5.13); thus \mathbf{m}_k is effectively the curvature of the \mathbf{n}_k - path. The sum is over every path in this space which begins as $(\mathbf{0}, 1)$, $(\mathbf{n}_0, 2)$, and ends as (\mathbf{n}_t, t) , $(\mathbf{0}, t+1)$.

In the area preserving case, the most important terms in this sum are those which have the minimum number of Bessel functions of nonzero argument and index. These are the so-called "principal terms". These appear to dominate for large values of the parameter 'a' since $J_m(na) \propto (a)^{-1/2}$ [Cary 1981].

The same argument can be given in this case, namely $g_{\mathbf{m}}(\mathbf{n}) \leq 1$ with equality only when $\mathbf{m} = \mathbf{n} = \mathbf{0}$. For the later case the path must have a straight segment which passes through the origin of the \mathbf{n} plane. The principal terms are those with the maximum number of such segments.

Consider the case of $\chi_t^P[(1,1), s(1,1)]$ where $s = \pm 1$, which is needed to compute C_{cc} . When t is even the principal path is given by the sequence, $\dots \rightarrow (1,1) \rightarrow (0,0) \rightarrow (-1,-1) \rightarrow (0,0) \rightarrow \dots$, as shown in Fig. 5.1(a) by the solid line. The resulting contribution to χ_t^P for $t \geq 2$ is

$$\chi_t^P [(1,1), s(1,1)] = [J_0(a) J_2(2c)]_2^t \delta_{s, (-1)^{\frac{t}{2}+1}} \quad t \text{ even} \quad (5.19)$$

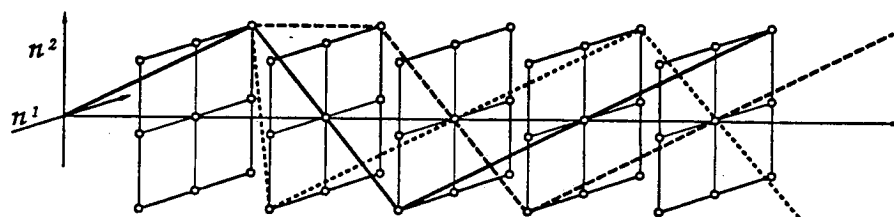
For odd t , the paths necessarily include at least one defect (i.e. an extra point with $\mathbf{m}, \mathbf{n} \neq 0$) in order to satisfy the boundary conditions at $k=0$ and t . Paths with a single defect, such as those shown by dashed lines in Fig. 5.1(a) will have the minimum number of Bessel functions. The defect can occur at any of $(t-1)/2$ positions. Each term gives the same result, so we obtain (for $t \geq 3$)

$$\chi_t^P [(1,1), s(1,1)] = \frac{t-1}{2} \left\{ [J_3(2c)]^2 \delta_{s,(-1)^{\frac{t-1}{2}}} + [J_1(2c)]^2 \delta_{s,(-1)^{\frac{t-3}{2}}} \right\} \times \\ [J_0(a)]^2 [J_0(a) J_2(2c)]^{\frac{t-3}{2}}$$

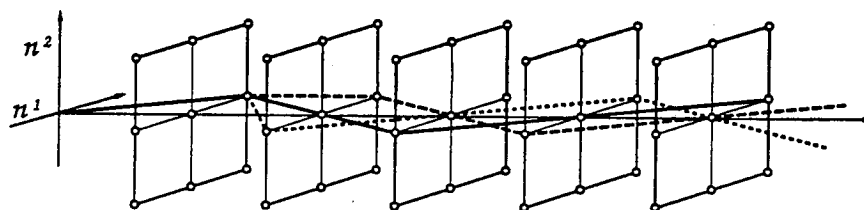
for t odd (5.20)

The paths contributing to the principal terms for $\chi_t[(1,0), s(1,0)]$ are shown in Fig. 5.1(b). When t is even the solid path is dominant, and when t is odd, both the dashed paths must be included. The results are equivalent to the above upon making the substitutions $2c \rightarrow a$ and $a \rightarrow c$.

Finally one would like to compute the principal terms for $\chi_t[(1,1), (\pm 1, 0)]$. Such paths must also have at least one defect. However, it is not difficult to see that for each path with one defect contributing to $\chi_t[(1,1), (1,0)]$, there is one which gives the identical result for $\chi_t[(1,1), (-1,0)]$. Equation (5.15) implies that these terms cancel from C_{ac} . In order to compute the effect of C_{ac} on the diffusion it is necessary to include paths with two or more defects. Thus $C_{ac} = 0$ in the principal term approximation.



(a)



(b)

Fig. 5.1 Principal terms represented by paths in (\mathbf{n}, k) space (a) for $\chi_t[(1,1), \pm(1,1)]$ and (b) $\chi_t[(1,0), \pm(1,0)]$. The solid paths are for even t ; the dashed paths for odd t cases.

5.5. Quasilinear Diffusion and Corrections

Combining eq. (5.5) and (5.11) with eq. (5.16) gives an explicit, formal expression for the diffusion tensor. As was the case for area preserving maps, the convergence of these series is doubtful due to the long-time correlations induced by regular regions in the phase space (in particular, the accelerator modes, see below). Such effects have not been studied when the number of degrees of freedom ($N+1$) is larger than two. However, moderate time and finite precision numerical experiments can be expected to give results close to those given by eq. (5.5) if a finite number of terms in t are retained.

The first term in the series for the diffusion tensor is the so-called quasilinear term:

$$\mathbf{D}_{\text{ql}} = \frac{1}{2} \mathbf{C}(0) = \frac{1}{2} \sum_{\mathbf{m}} \mathbf{f}_{\mathbf{m}} \mathbf{f}_{-\mathbf{m}} \quad (5.21)$$

where we have set \mathfrak{X} to the entire torus. For the Froeschlé example this is

$$\mathbf{D}_{\text{ql}} = \frac{1}{4} \begin{pmatrix} a^2+c^2 & c^2 \\ c^2 & b^2+c^2 \end{pmatrix} \quad (5.22)$$

The series in force correlations gives corrections to this which, because of the Bessel functions, tend to oscillate as the parameter values change. The first few terms are easy to write out (with $b=0$):

$$\begin{aligned}
C_{aa}(1) &= C_{cc}(1) = 0 \\
C_{aa}(2) &= -1/2a^2 J_2(a)J_0(c), \quad C_{cc}(2) = -1/2c^2 J_0(a)J_2(2c) \\
C_{aa}(3) &= -1/2a^2 [J_1^2(a) - J_3^2(a)] J_0^2(c), \\
C_{cc}(3) &= -1/2c^2 [J_1^2(2c) - J_3^2(2c)] J_0^2(a)
\end{aligned} \tag{5.23}$$

The coupling correlation, $C_{ac}(t)$, vanishes identically for $t < 4$.

Using the principal terms (5.19) and (5.20) for the longer time correlations we can do the sum in eq. (5.5) to obtain:

$$\begin{aligned}
D_{aa}^P &= \frac{a^2}{4} \frac{1 - [2J_1^2(a) + J_2^2(a) - 2J_3^2(a)] J_0^2(c)}{[1 + J_2(a)J_0(c)]^2} \\
D_{cc}^P &= \frac{c^2}{4} \frac{1 - [2J_1^2(2c) + J_2^2(2c) - 2J_3^2(2c)] J_0^2(a)}{[1 + J_0(a)J_2(2c)]^2}
\end{aligned} \tag{5.24}$$

Note that these are not separable into terms due to the first degree of freedom, and terms due to the second. In the principal terms approximation, $D_{ac} = 0$, or $D_{12} = D_{21} = D_{22}$. We expect these results to be a good approximation when both a and c are not too small.

An interesting limit is that of $a \gg 1$, implying that the first degree of freedom is highly stochastic and $c \ll 1$, implying that the coupling is weak. In this case the self-diffusion of the first degree-of-freedom is approximately quasilinear. This degree of freedom drives the diffusion in the second, this is

called "thick-layer diffusion" [Tennyson 1979]. Tennyson et al compute such diffusion by treating the second degree of freedom as though it was stochastically driven by the first. This technique yields only the lowest order, quasilinear term, in D_{cc} . The principal terms expression implies that $D_{cc} \rightarrow D_{ql}$ in the limit $c \rightarrow 0$; however, this is not exact because there are an infinite number of terms which contribute to D at $O(c^4)$ which have been neglected.

Another limit of interest, is $a \lesssim 1$, $c \ll 1$, where the initial conditions are chosen with the first degree of freedom in a narrow stochastic layer. This leads to "thin layer diffusion", or what we think of as Arnol'd diffusion [Lichtenberg 1983]. Our method does not give an effective technique for computing this.

5.6. Accelerator Modes

Maps of the form (5.1) are also periodic in the momentum direction: if $(\mathbf{y}_t, \mathbf{x}_t)$ is an orbit on $\mathcal{R}^N \times \mathcal{T}^N$ then $(\mathbf{y}_t + 2\pi\mathbf{j}, \mathbf{x}_t)$ is also an orbit for any integer vector \mathbf{j} . This suggests looking for generalized "periodic" orbits satisfying the conditions

$$\mathbf{y}_t = \mathbf{y}_0 + 2\pi\mathbf{j}, \quad \mathbf{x}_t = \mathbf{x}_0 \quad (5.25)$$

If $\mathbf{j} \neq \mathbf{0}$ then these orbits are called t^{th} order accelerator modes.

The condition for the first order accelerator modes, $t=1$, for eq. (1) is that $-\nabla V(\mathbf{x}) = 2\pi\mathbf{j} = 2\pi(r, s)$ where r and s are integers, and $\mathbf{y} = \mathbf{0}$. Using the potential (5.3), with $b=0$ implies that

$$\sin(x^1) = \frac{2\pi(s-r)}{a} \quad \sin(x^1 + x^2) = -\frac{2\pi s}{c} \quad (5.26)$$

Thus these can exist only if $|c| \geq 2\pi|s|$ and $|a| \geq 2\pi|r-s|$. In general there are four such fixed points for each (r, s) , of which only one is stable. The requirement for stability (all four eigenvalues of modulus unity) is determined by linearizing the map about the accelerator orbit:

$$\begin{aligned} & \alpha\gamma \geq 0, \quad (\alpha-8)(\gamma-4) \geq 16 \quad \text{and} \quad 0 \leq \alpha+2\gamma \leq 8 \\ \text{where} \quad & \alpha \equiv a \cos(x^1) = \pm \sqrt{a^2 - 4\pi^2(s-r)^2} \\ & \gamma \equiv c \cos(x^1 + x^2) = \pm \sqrt{c^2 - 4\pi^2 s^2} \end{aligned} \quad (5.27)$$

The regions of parameters where the stable first order accelerator modes exist are shown in Fig. 5.2.

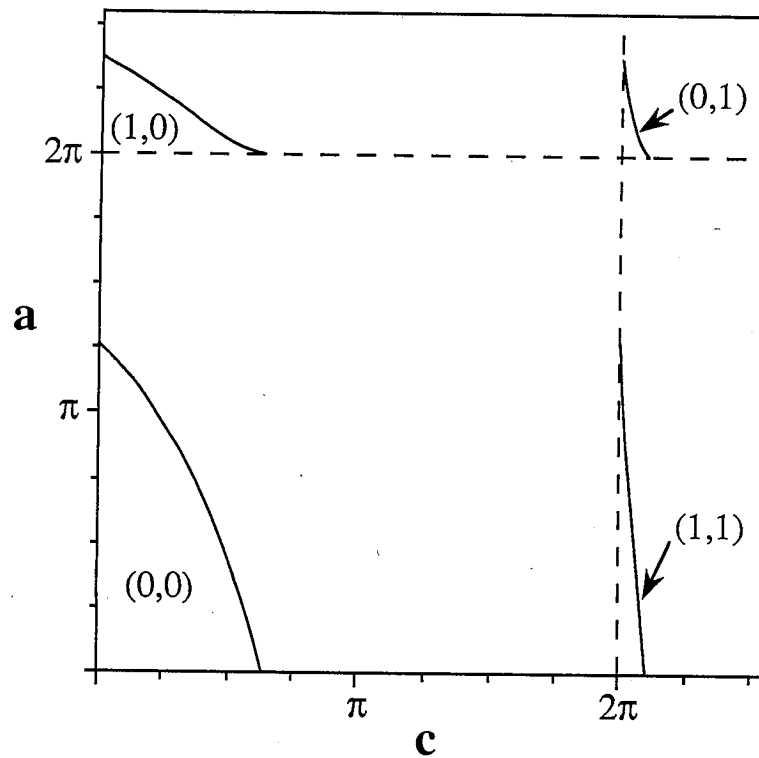


Fig. 5.2 Accelerator mode stability regions. Parameter regions for the existence of the first order accelerator modes with various \mathbf{j} are shown. An orbit trapped in the $\mathbf{j}=(r,s)$ mode propagates by $2\pi r$ in y^1 direction and by $2\pi s$ in y^2 direction upon each iteration of the map. The mode $\mathbf{j}=(0,0)$, the fixed point, does not contribute to acceleration.

In the case of area preserving maps, the existence of a stable accelerator mode causes the series for the diffusion coefficient to diverge. This is due to two effects: first, if the region \mathfrak{R} is chosen to be entire phase space, then initial conditions trapped in an accelerator mode (which by the KAM theorem have finite measure) are included in the average [Ichikawa 1983]. Since Δy increases in proportion to t for such initial conditions, D necessarily diverges. The characteristic functions must include this effect; however, extracting it from the formulae appears difficult.

Second, even if initial conditions trapped in an accelerator mode are not included in the average, numerical evidence indicates that the correlation function decays as t^{-z} for large time where z is in the range 0.3 - 0.5 [Karney 1983, Meiss 1986]. This is due to long time trapping of stochastic orbits near "sticky" islands. Consequently, the series (5.5) must be divergent even if \mathfrak{R} is chosen to be the connected stochastic region.

This divergence might be expected to be less severe in higher dimensional maps because initial conditions in the neighborhood of an elliptic orbit can still escape by the mechanism of Arnol'd diffusion. However, this escape can be slow [Nekhoroshev 1965], and may still give an algebraic decay of the correlation function.

5.7. Numerical Experiments

The diffusion tensor can be computed directly from eq. (5.4), interpreting the average as a sum over a grid of initial conditions in phase space. We iterate each initial condition for a time T , and estimate the statistical errors in D by the RMS deviation of the results for the different initial conditions. The RMS deviation is seen to decrease as inverse of the square root of the ensemble size.

Comparison of the numerical results with the principal term theory is shown in Figs. 5.3 for $a=3.0$, as c varies from 0.0 to 5.0. Here we chose a grid of 10^4 initial conditions, and iterated each $T=50$ times. The plots (a) and (b) show the values of D_{aa} and D_{cc} normalized to their quasilinear values, respectively. The contribution of D_{ac} to D_{12} is comparable to D_{cc} when c is small (Fig. 5.3c). For larger c , D_{ac} is consistent with zero (statistical errors due to the finite ensemble prevent an accurate measurement). Results for various ranges of parameters show that providing a and c are larger than 2 the principal term results agree extremely well with the numerical experiments.

In the region of parameter space where a first order accelerator mode is stable, the diffusion coefficient is enhanced when the ensemble is $\mathcal{T}^N \times \mathcal{T}^N$. As in the case of the standard map, the peak in the diffusion coefficient grows as T is increased, indicating that convergence has not been obtained. To exhibit the role of the accelerator mode, we choose initial

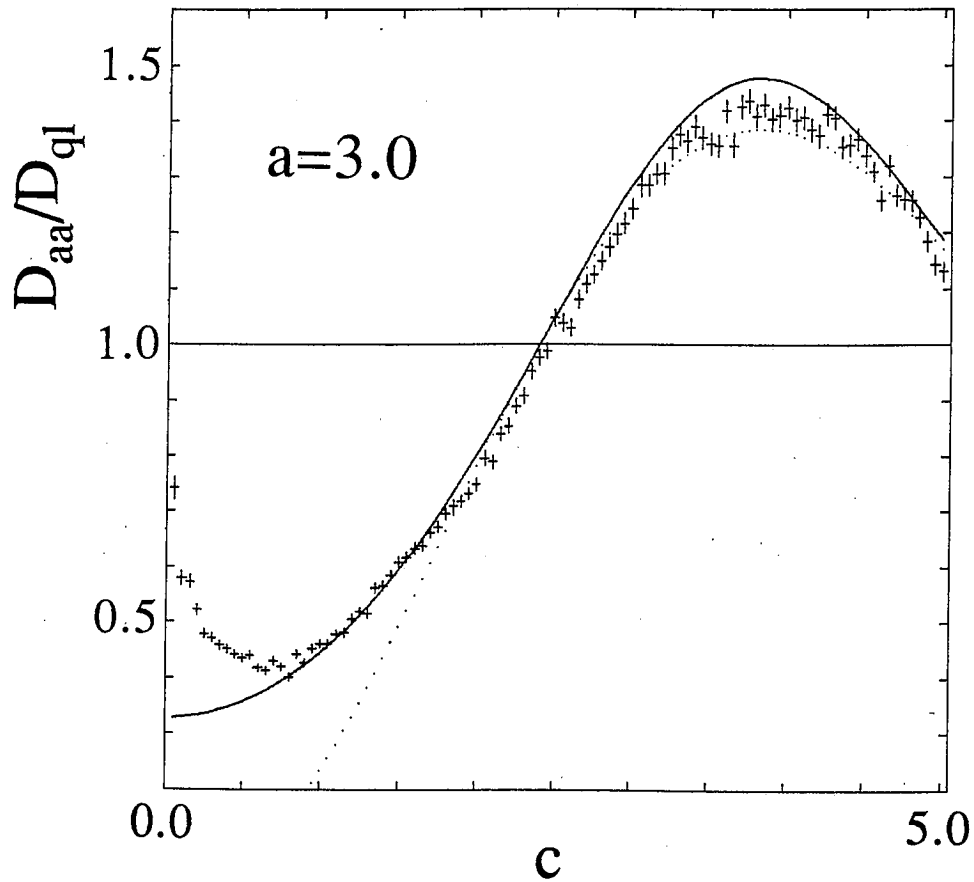


Fig. 5.3 (a) Comparison of numerical diffusion coefficients with theoretical results. The dotted curves represent the theory which includes up to the third order correction to the quasilinear result. The principal term results are drawn with the solid curves. The statistical ensemble consists of 10^4 orbits of length $T=50$ with their initial conditions chosen from an uniform grid over the whole phase space. The error bars give the RMS deviation from the average values. (a) D_{aa} is normalized to its quasilinear value.

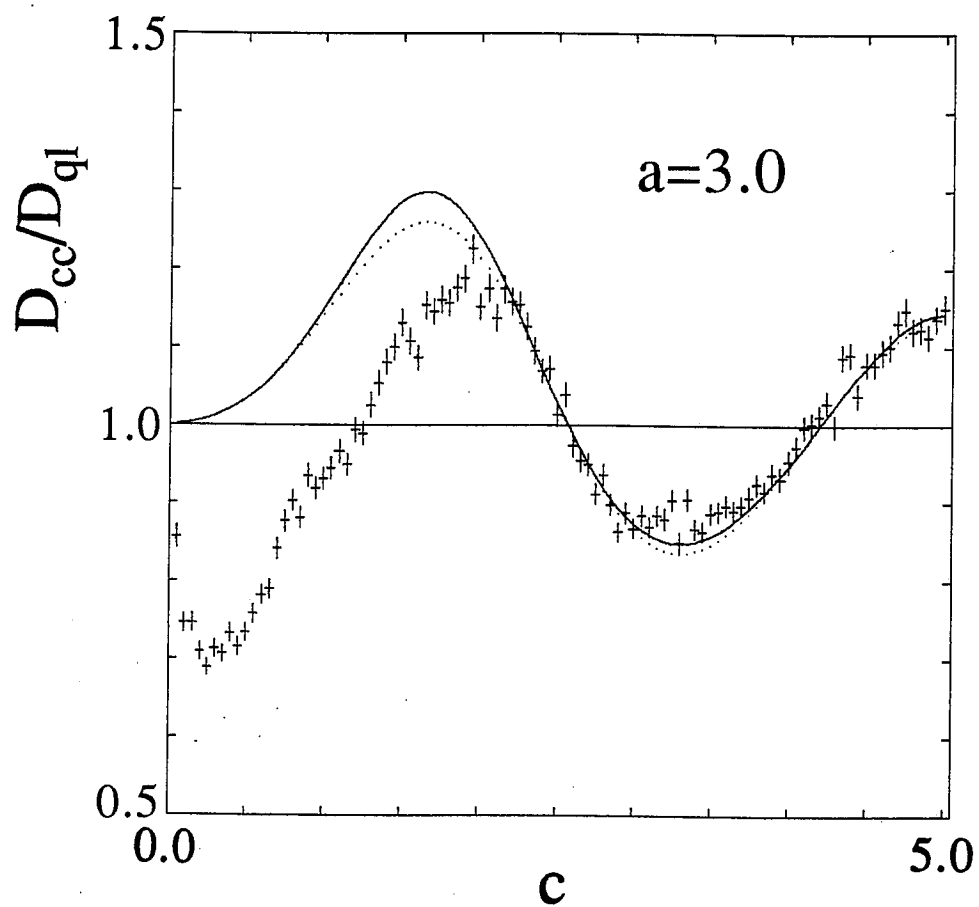


Fig. 5.3 (b) Comparison of numerical diffusion coefficients with theoretical results. D_{cc} is normalized to its quasilinear value.

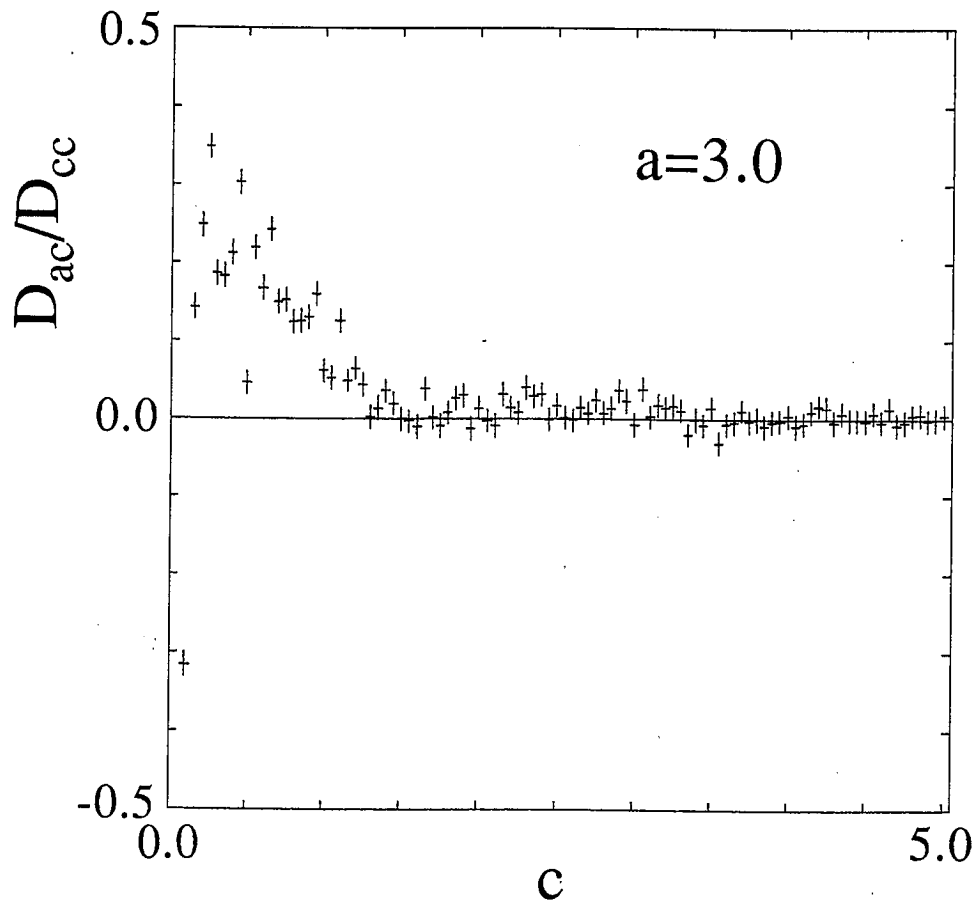


Fig. 5.3 (c) Comparison of numerical diffusion coefficients with theoretical results. D_{ac} is normalized to D_{cc} . Its fluctuation for $c \geq 2.0$ is of order of D_{ac} ; however, considering its contribution to D_{12} , it is still consistent to zero.

conditions at $\mathbf{y} = \mathbf{0}$, with \mathbf{x} chosen on a 10^2 by 10^2 grid: the location as well as the size of stable islands varies as parameters vary; however, first order accelerator modes are always at $\mathbf{y} = \mathbf{0}$. Figs. 5.4(a)-(c) demonstrate divergent behavior of D_{aa} in T due to the first order accelerator mode with $\mathbf{j} = (1,0)$.

To attempt to eliminate orbits trapped on invariant tori encircling the stable accelerator mode, we choose an initial condition near the unstable fixed point $\mathbf{y} = \mathbf{0}$, and $\mathbf{x} = (0.5,0.5)$, iterate it a large number of iterations, and break the orbit into n segments of length T to obtain statistics:

$$\mathbf{D} = \frac{1}{n} \sum_{k=1}^n \frac{1}{2T} (\mathbf{y}_{kT} - \mathbf{y}_{(k-1)T}) (\mathbf{y}_{kT} - \mathbf{y}_{(k-1)T}) \quad (5.28)$$

The result is shown in Fig. 5.4(d) with $n=10^3$, $T=10^4$ for the same parameters as Fig. 5.4(c). In contrast to the area preserving case, these results show little dependence of the diffusion coefficients on T . This can be interpreted that the trapping time distribution of a stochastic orbit near the accelerator mode decays rapidly enough that the diffusion series converges. We will study the decay of correlations for this situation in a future paper.

When either of parameters is small, the effect of trapping is conceivably important. Thus, to study the small c regime we use the single, long orbit statistics introduced above. Our numerical results imply that as c approaches zero, D_{aa} limits to the diffusion of the standard map for the given value of a , as would be expected. On the other hand, D_{cc} appears to scale like c^2 as c goes to zero, but does not seem to converge to its quasilinear

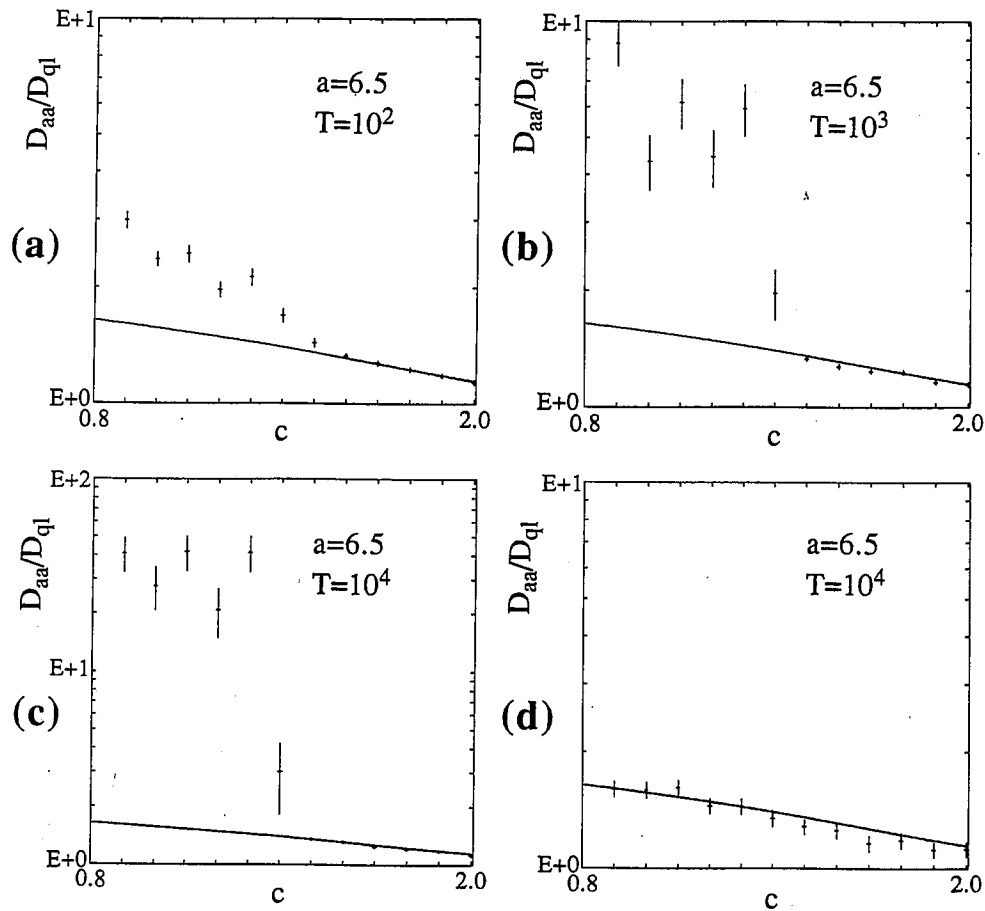


Fig. 5.4 Enhancement of diffusion coefficient D_{aa} due to the presence of the stable accelerator mode with $\mathbf{j}=(1,0)$. Its divergent behavior in time is shown; (a) $T=10^2$, (b) $T=10^3$ and (c) $T=10^4$. The statistical ensemble consists of 10^4 orbits whose initial conditions are uniformly distributed on the $\mathbf{y}=\mathbf{0}$ plane. (d) shows that the effect of accelerator modes can be avoided for moderately long times by choosing the ensemble in a connected stochastic region. The ensemble consists of 10^3 suborbits of length $T=10^4$, obtained by breaking a single stochastic orbit of length $T=10^7$ into pieces. The initial condition is chosen near the unstable fixed point; $x^1=0.5$, $y^1=0.01$, $x^2=0.5$ and $y^2=0.02$.

value consistently for all values of a , as would be expected on the basis of Eq. (5.24). Rather, it is observed that $\lim_{c \rightarrow 0} D_{cc}/D_{ql}$ depends on a in an oscillatory fashion, only approaching one as $a \rightarrow \infty$. This is shown in Fig. 5.5, setting $c = 10^{-5}$ (which is small enough to be considered as the $c \rightarrow 0$ limit), $n=10^3$ and $T=10^3$. Note that the oscillations in D_{cc}/D_{ql} are in phase with those of D_{aa}/D_{ql} which is also shown in Fig. 5.5. The relative behavior of D_{aa} and D_{cc} give an insight to correlations between the motion in the two canonical planes in the parameter regime of small c and large a : the short time correlations in one plane affect even the zero coupling limit of the diffusion in the other plane.

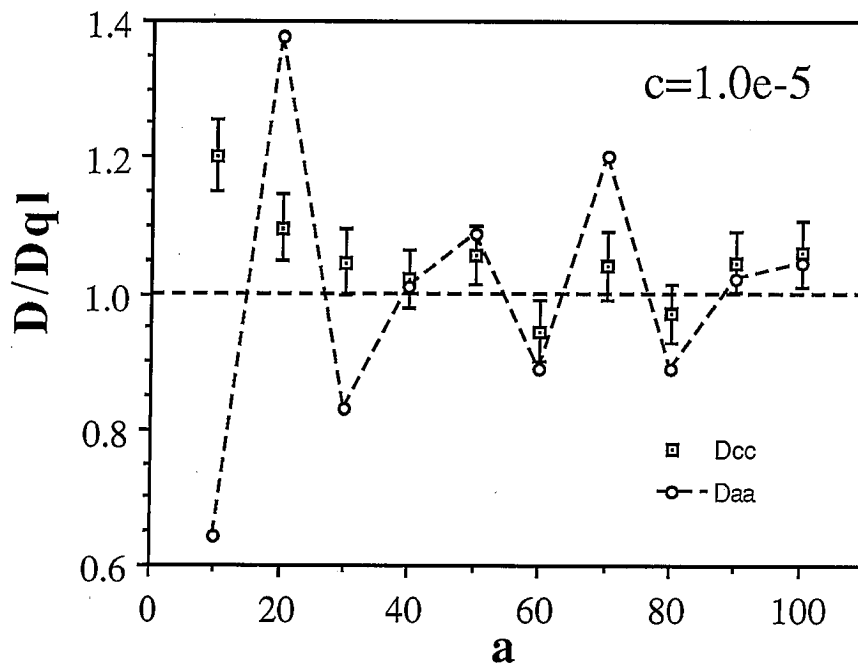


Fig. 5.5 Oscillatory dependence of $\lim_{c \rightarrow 0} D_{cc}/D_{ql}$ on a . Actually $c = 10^{-5}$, which is small enough for D_{cc} to approach its limiting behavior. Both D_{cc} (squares) and D_{aa} (circles) normalized to their quasilinear values are shown. Only D_{cc} is plotted with its RMS deviations. The same statistics as for Figs. 4 is used except a different length for suborbits ($T=10^3$).

5.8. Conclusions

We have derived a formal series for the momentum diffusion tensor of a symplectic map with the second difference (or Lagrangian) form (5.1), in terms of the characteristic functions. The lowest order terms in this series give the quasilinear results for the diffusion tensor. Higher order terms provide oscillations about \mathbf{D}_{ql} as a function of the map parameters, showing that the oscillatory behavior of \mathbf{D} is due to short time correlations. The force correlation tensor, $\mathbf{C}(t)$ can be obtained analytically for $t \leq 3$ using the characteristic function method. For larger t , the expressions for the correlations involve infinite sums. The dominant terms in these sums for large parameters, the principal terms, have an easily calculable form, and an approximate summation of the diffusion series retaining only these terms gives a good approximation to \mathbf{D} as long as the correlation function decays rapidly.

We show that stable accelerator modes enhance the diffusion. For the area preserving case these cause the divergence of D ; however, for higher dimensions this appears to be no longer true when the ensemble of orbits does not include the accelerating invariant tori. The long time behavior of the correlation function for a stochastic orbit is still under investigation.

The thick-layer model of diffusion is appropriate when a is large and c is small. In the limit $c \rightarrow 0$, the motion in the two canonical planes is separable, and when $b=0$ the (x^2, y^2) motion is integrable. The transport in

this limit is described by the quantity $\lim_{c \rightarrow 0} D_{cc}/D_{q1}$. While this function approaches one as a increases, it oscillates with a due to short time correlations in the (x^1, y^1) plane.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The dynamics of a Hamiltonian system with $N+1$ degrees-of-freedom is represented by a flow on the $2N+1$ -dimensional constant energy surface in phase space. The successive returns of a flow to the Poincaré surface of section induce a $2N$ -dimensional symplectic map. Symplectic maps give an efficient way to study Hamiltonian systems since the qualitative behavior of the flows is preserved on the surface of section.

For an integrable mapping, the dynamics is described by uniform rotations on N -dimensional tori. Upon a generic perturbation, all resonant tori are broken up, but a finite number of periodic motions on each resonant torus are still retained. We have shown that reversible, symplectic, twist maps of a $2N$ -dimensional phase space have periodic orbits which can be classified by frequency, symmetry, and Morse index. There are 2^N different types of symmetric orbits for each frequency. For each frequency, we also find 2^N orbits whose action functions have indices ranging from 0 to N . For small perturbations, these two sets of orbits often coincide; however, the symmetric orbits commonly undergo symmetry breaking bifurcations which change their indices. Orbital stability is closely related to the index of the orbit even though the index is not a sufficient test for stability for general N ; an action minimizing orbit (index zero) is typically hyperbolic, though there are possible counter examples. Orbits with index less than N tend to avoid the dominant symmetry plane, which gives rise to the formation of a Sierpinski-carpet-like set in the configuration plane. This is analogous to the formation of cantori in area preserving maps.

The structure of phase space is viewed via resonances and a chaotic layer. Orbits in a resonance can remain trapped for long periods, and while they are trapped have the same frequency as the resonance. Collections of resonances which obey a commensurability relation form channels on the symmetry plane. Chaotic transport from resonance to resonance appears to take place along the edges of these channels. The collection of channels gives the Arnol'd web.

Finding an orbit of a map for a given period and parameters is equivalent to solving a system of nonlinear simultaneous equations. This becomes a difficult task for unstable orbits since in the neighborhood of such orbits, numerical errors grow exponentially upon iteration of the map. Therefore, any mapping iteration scheme becomes impractical for strongly unstable orbits. We have obtained a stable and efficient Newton's algorithm based on the variational method. The computation time and the size of the storage space in this algorithm is linear in the period of the orbit. Another benefit of the algorithm is that the index of the orbit, which is closely related to the orbital stability, is obtained during execution of the algorithm. Finally, the construction of the algorithm relies only on the particular structure of the Hessian matrix of the action function. Since the structure of the Hessian matrix is due to the general form of the Lagrangian generating function, the method is expected to be applicable to a wide range of dynamical systems.

The action formulation gives an analytical way to tackle the geometric structure of phase space. In area preserving maps, partial barriers and turnstiles are constructed from the segments of stable and unstable

manifolds of minimizing orbits. Flux across a partial barrier occurs only through the turnstile. This flux as well as the area under the partial barrier is obtained from the action difference between relevant homoclinic orbits. These results are directly used to study transport in phase space for area preserving maps. Phase space is partitioned with partial barriers into separate regions; each region is associated with a state of a Markov process. The transition probability between states is determined by the flux between the regions. This defines a Markov model for transport. The Markov models are fairly successful in describing transport for area preserving maps; however, their applicability is limited due to the assumption that the dynamics in each region associated with a state is purely random. An extension of the ideas employed in area preserving maps to the higher dimensional cases has not yet been accomplished since the topology of the phase space structures are quite different in each case.

The characteristic function method provides another way to study transport for systems with an arbitrary number of degrees-of-freedom. Though this method does not rely much on the underlying dynamics, it is easily applicable to various maps (satisfying a periodicity condition) and its predictions give excellent agreement with numerical experiments for moderately large parameters. A formal series for the diffusion tensor is obtained in terms of the characteristic functions. The lowest order term in the series gives the quasilinear result for the diffusion tensor. High order terms provide oscillations about the quasilinear value as a function of the parameters, showing that the oscillatory behavior of the diffusion tensor is

mainly due to the short time correlations. The correlation tensor $\mathbf{C}(t)$ is obtained analytically for $t \leq 3$. For larger t , the principal terms give an approximation, which is in good agreement with numerical experiments for moderately large parameters. Accelerator modes exist for doubly periodic maps. For area preserving maps these have been shown to cause the divergence of D ; however, for higher dimensional maps this appears to be no longer true when the ensemble of orbits does not include the accelerating invariant tori.

Understanding the long-time behavior of Hamiltonian systems is extremely important for examining transport processes, especially since phase space is generically a mixture of regular and chaotic regions on all scales. For the area preserving case, it has been shown that the resonant islands (regular regions) trap a chaotic orbit for a long periods and that this trapping dominates the long-time behavior of the correlation function, causing it to decay more slowly than t^{-1} . This implies that a statistical description of transport as diffusive is inappropriate for area preserving maps. It is still an open question whether the transport in the chaotic region of higher degree-of-freedom systems is normally diffusive. To better understand transport in this case, further investigation is needed on the details of dynamics and the undeniably complicated structure of phase space.

In the subject of the present work, there still seem to be many unanswered questions. We list some as follows:

- How does a KAM torus break up? Which is the last (perhaps it has the frequency of spiral mean)? After the break-up, does it form a Sierpinski set, and eventually become a cantor set?
- Do minimizing quasiperiodic orbits exist (generalization of Aubry-Mather theory)?
- What is the role of symplectic flux in transport? Is there some relation of N-dimensional Lagrangian manifolds such as invariant tori to transport?
- How can partial barriers be constructed?
- How can the resonance boundary be constructed? Do the orbits of index less than N play a role?
- Can Arnol'd diffusion be calculated in terms of transport along channels? Is it truly diffusion?
- What dominates the long-time behavior of transport? Does the correlation function fall off faster than t^{-1} ?

Appendix A: Constant Twist Case

Lemma: A mapping with constant twist and zero net flux has a generating function

$$F = \frac{1}{2} (\mathbf{q}' - \mathbf{q}) \cdot \mathbf{b} \cdot (\mathbf{q}' - \mathbf{q}) - V(\mathbf{q}) \quad (\text{A1})$$

where \mathbf{b} is a symmetric matrix, and V is a periodic function.

Proof: If the twist is constant then the generating function can be written as

$$F = \frac{1}{2} [-\mathbf{q}' \cdot \mathbf{B} \cdot \mathbf{q} + f(\mathbf{q}) + g(\mathbf{q}')]$$

where \mathbf{B} is a constant matrix and f and g are arbitrary functions. Periodicity of the mapping implies that the net flux, $\mathcal{F}_m = F(\mathbf{q} + \mathbf{m}, \mathbf{q}' + \mathbf{m}) - F(\mathbf{q}, \mathbf{q}')$ is independent of \mathbf{q} and \mathbf{q}' for all integer vectors \mathbf{m} , or

$$\mathcal{F}_m = f(\mathbf{q} + \mathbf{m}) - f(\mathbf{q}) + g(\mathbf{q}' + \mathbf{m}) - g(\mathbf{q}') - \mathbf{q} \cdot \mathbf{B} \cdot \mathbf{m} - \mathbf{m} \cdot \mathbf{B} \cdot \mathbf{q}' - \mathbf{m} \cdot \mathbf{B} \cdot \mathbf{m} \quad (\text{A2})$$

This implies that f and g can have arbitrary periodic parts, call them \mathcal{V} and \mathcal{U} respectively, and that the remainder must be at most quadratic, thus

$$f(\mathbf{q}) = \mathcal{V}(\mathbf{q}) + \mathbf{q} \cdot \mathbf{C} \cdot \mathbf{q} + \mathbf{q} \cdot \mathbf{D}$$

$$g(\mathbf{q}') = \mathcal{U}(\mathbf{q}') + \mathbf{q}' \cdot \mathbf{E} \cdot \mathbf{q}' + \mathbf{q}' \cdot \mathbf{F}$$

for matrices \mathbf{C} and \mathbf{E} and vectors \mathbf{D} and \mathbf{F} . Substitution into (A2) implies that $\mathbf{B} = \mathbf{C} + \tilde{\mathbf{C}} = \mathbf{E} + \tilde{\mathbf{E}}$, and that $\mathbf{m} \cdot (\mathbf{D} + \mathbf{F}) = \mathcal{F}_m$. Therefore for a solution to exist, the matrix \mathbf{B} must be symmetric. Furthermore, if the net flux is zero, $\mathbf{D} + \mathbf{F} = 0$, and the generating function takes the form

$$F = \frac{1}{2} [(\mathbf{q}' - \mathbf{q}) \cdot \mathbf{B} \cdot (\mathbf{q}' - \mathbf{q}) + (\mathbf{q} - \mathbf{q}') \cdot \mathbf{D} - \mathcal{V}(\mathbf{q}) - u(\mathbf{q})] \quad (\text{A3})$$

Equation (A3) can be simplified by the canonical coordinate change $\bar{\mathbf{q}} = \mathbf{q}$, $\bar{\mathbf{p}} = \mathbf{p} + 1/2 \mathbf{D} + 1/2 \nabla u(\mathbf{q})$ which gives Eq. (A1) in the new coordinates, with $V(\mathbf{q}) \equiv 1/2 [\mathcal{V}(\mathbf{q}) + u(\mathbf{q})]$. \square

Appendix B: Existence of Periodic Orbits

The existence of a minimum for the periodic action W_ω is guaranteed by the following:

Lemma [MacKay 1988]: Let $F(\mathbf{q}, \mathbf{q}')$ have zero net flux and suppose F_{12} is uniformly negative definite, i.e. $\exists B > 0$ such that $\mathbf{x} \cdot F_{12}(\mathbf{q}, \mathbf{q}') \cdot \mathbf{x} \leq -B |\mathbf{x}|^2$. Then $\exists \alpha$, and positive β and γ such that

$$F(\mathbf{q}, \mathbf{q}') \geq \alpha - \beta |\mathbf{q} - \mathbf{q}'| + \gamma |\mathbf{q} - \mathbf{q}'|^2 \quad (\text{B1})$$

Proof:

$$F(\mathbf{q}, \mathbf{q}') = F(\mathbf{q}, \mathbf{q}) + \int_0^1 d\lambda F_2(\mathbf{q}, \xi_\lambda) \cdot (\mathbf{q}' - \mathbf{q})$$

where $\xi_\lambda = (1-\lambda)\mathbf{q} + \lambda\mathbf{q}'$. Similarly applying this to F_2 yields

$$\begin{aligned} F(\mathbf{q}, \mathbf{q}') &= F(\mathbf{q}, \mathbf{q}) + \int_0^1 d\lambda F_2(\xi_\lambda, \xi_\lambda) \cdot (\mathbf{q}' - \mathbf{q}) \\ &\quad - \int_0^1 d\lambda \int_0^\lambda d\mu (\mathbf{q}' - \mathbf{q}) \cdot F_{12}(\xi_\mu, \xi_\lambda) \cdot (\mathbf{q}' - \mathbf{q}) \end{aligned}$$

$$\geq \alpha - \beta | \mathbf{q}' - \mathbf{q} | + | \mathbf{q}' - \mathbf{q} |^2$$

where $\alpha = \min F(\mathbf{q}, \mathbf{q})$ and $\beta = \max | F_2(\mathbf{q}, \mathbf{q}) |$ exist since $F(\mathbf{q}, \mathbf{q})$ is periodic for zero net flux, and $\gamma = 1/2B$. \square

Corollary 1: There exists a minimum of W_ω for each $\omega = (\mathbf{m}, n)$.

Proof: Equation (B1) implies that F has a lower bound, thus W_ω does as well. Furthermore, the set $\{ \mathbf{q}, \mathbf{q}' \mid \mathbf{q} \in \mathcal{T}^N, \mathbf{q}' \in \mathcal{R}^N, F(\mathbf{q}, \mathbf{q}') \leq C \}$ is compact since (B1) implies that $F < C$ corresponds to bounded $| \mathbf{q} - \mathbf{q}' |$. Similarly the set

$$\mathcal{S} = \left\{ \mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1} \mid \mathbf{q}_0 \in \mathcal{T}^N, (\mathbf{q}_1, \dots, \mathbf{q}_{n-1}) \in \mathcal{R}^{(n-1)N}, \right. \\ \left. W_\omega(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1}) < C' \right\}$$

is compact. W_ω is continuous and bounded on \mathcal{S} , thus a minimum exists. \square

Corollary 2: There exist at least $N+1$ (and generically at least 2^N) periodic orbits for each $\omega = (\mathbf{m}, n)$.

Proof: Consider the constrained minimization of W_ω for fixed \mathbf{q}_0 . Such a minimum exists by virtue of the preceding considerations. Define the function

$$F(\mathbf{q}_0) \equiv \min_{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{n-1}} W_\omega(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{n-1})$$

to be this minimum. Of course a critical point of W_ω for fixed \mathbf{q}_0 gives an orbit segment, but not generally a periodic orbit. However, zero net flux implies F is a function on \mathcal{T}^N ; such a function has at least $N+1$ critical points. In fact, Morse theory [Milnor 1963] implies that if the critical points are non-degenerate, there must be at least 2^N ; furthermore, non-degeneracy is generic. In this case there are at least $N! / i! (N-i)!$ orbits with index i . Each critical point of F gives a periodic orbit. \square

Corollary 3: There exist at least 2^N symmetric periodic orbits for each $\omega = (\mathbf{m}, n)$, one for each symmetry class.

Proof: Consider the minimum of the half orbit action:

$$W_k(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{k-1}) \equiv W(\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{k-1}, \mathbf{q}_k) \Big|_{\substack{\mathbf{q}_0 \in \text{Fix}(S) \\ \mathbf{q}_k \in \text{Fix}(S')}}$$

where S is a symmetry and S' is SR_m or TSR_m for n even or odd, respectively, according to Eq. (2.28). Critical points of W_k give a periodic orbit of period n by using Eq. (2.28) with $i=0$ to define the \mathbf{z}_j for $j \in [k+1, n]$. Such critical points exist because W_k is bounded from below and the set $W_k \leq C$ is compact. \square

The symmetric periodic orbits of Corollary 3 do not necessarily coincide with the orbits of Corollary 2.

SUBROUTINE BROYDEN(Y1,Y2,N,ITMAX,ERRY,ERRF,DNRMY,DNRMF,ITER)

CC

C

C RECEIVES Y1,Y2, THE INITIAL GUESS OF MOMENTA ON SYMM PLANE

C SOLVES A NONLINEAR SYSTEM OF 2 UNKNOWNS USING NEWTON'S METHOD

C B IS THE JACOBIAN MATRIX FOR (F1,F2)

C B OR ITS INVERSE IS OBTAINED ITERATIVELY BY BROYDEN'S SCHEME

C INITIAL B=IDENTITY, DY1= -INV(B0)*F0

C

CC

IMPLICIT REAL*8(A-H,O-Z)

REAL*8 B(2,2)

CALL FROESCHLE(Y1,Y2,F1,F2,N)

C INITIAL JACOBIAN B IS GIVEN AS AN IDENTITY

B(1,1)=1.0

B(2,1)=0.0

B(1,2)=0.0

B(2,2)=1.0

SERRY=ERRY**2

SERRF=ERRF**2

ITER=0

20 ITER=ITER+1

C UPDATE Y; DY=-INVB*F

DET B=B(1,1)*B(2,2)-B(2,1)*B(1,2)

DY1=(-B(2,2)*F1+B(1,2)*F2)/DET B

DY2=(B(2,1)*F1-B(1,1)*F2)/DET B

$Y1=Y1+DY1$ $Y2=Y2+DY2$

CALL FROESCHLE(Y1,Y2,F1,F2,N)

 $DNRMY=DY1^{**2}+DY2^{**2}$ $DNRMF=F1^{**2}+F2^{**2}$

C EXIT WHEN CRITERIA FULFILLED OR BAD CONVERGENCE

IF(DNRMY.LE.SERRY) GO TO 40

IF(ITER.GT.ITMAX) GO TO 40

C BROYDEN UPDATE OF JACOBIAN B

 $DY1=DY1/DNRMY$ $DY2=DY2/DNRMY$ $B(1,1)=B(1,1)+F1*DY1$ $B(2,1)=B(2,1)+F2*DY1$ $B(1,2)=B(1,2)+F1*DY2$ $B(2,2)=B(2,2)+F2*DY2$

GO TO 20

40 DNRMY=SQRT(DNRMY)

DNRMF=SQRT(DNRMF)

RETURN

END

SUBROUTINE NEWTON(X,DETA,TRA,N,ITMAX,EPS,IERR,CRIT1,CRIT2,ITER)

CC

C

C RECEIVES THE INITIAL ANGLES X, 2N-DIM VECTOR

C X(N) IS FIXED ON THE SYMMETRY PLANE

C ITERATES THE NEWTON LOOP ITER TIMES

C DETERMINES FINAL ANGLES X WITHIN THE ACCURACY EPS

C DETERMINANTS OF DIAGONAL SUBMATRICES A ARE STORED IN DETA(N)

C TRACES OF DIAGONAL SUBMATRICES A ARE STORED IN TRA(N)

C

CC

PARAMETER (NMAX=13580)

REAL*8 DETA(N),TRA(N)

REAL*8 X(2,N),G(2,NMAX),DELX(2,NMAX)

REAL*8 H(2,2,NMAX+1),A(2,2),INVA(2,2,NMAX)

REAL*8 B(2,2,NMAX),INVB(2,2,NMAX)

REAL*8 C(2,2,NMAX),UP(2,NMAX),DN(2,NMAX),TEMP(2,2)

REAL*8 EPS,DEPS,CRIT1,CRIT2,DNORM

NM=N-1

NONINV=0

IERR=0

ITER=0

DO 5 I=1,NM

DO 5 K=1,2

DELX(K,I)=0.D0

5 CONTINUE


```
10    CONTINUE
      ITER=ITER+1

      DO 15 I=1,NM
      DO 15 K=1,2
          X(K,I)=X(K,I)-DELX(K,I)
15    CONTINUE

      CALL HESS(X,H,N)

      CALL GRAD(X,G,N)

      DO 20 L=1,2
      DO 20 K=1,2
          A(K,L)=H(K,L,1)
20    CONTINUE

      DETA(1)=A(1,1)*A(2,2)-A(1,2)**2
      IF(DETA(1).EQ.0.D0) THEN
          NONINV=1
          RETURN
      ENDIF
      TRA(1)=A(1,1)+A(2,2)

      INVA(1,1,1)=A(2,2)/DETA(1)
      INVA(2,1,1)=-A(2,1)/DETA(1)
      INVA(1,2,1)=INVA(2,1,1)
      INVA(2,2,1)=A(1,1)/DETA(1)

      B(1,1,1)=1.D0
      B(2,1,1)=0.D0
      B(1,2,1)=0.D0
      B(2,2,1)=1.D0
```

```
INVB(1,1,1)=1.D0
INVB(2,1,1)=0.D0
INVB(1,2,1)=0.D0
INVB(2,2,1)=1.D0

DO 30 K=1,2
    UP(K,1)=G(K,1)
30 CONTINUE

DO 70 I=2,NM
    IM=I-1

    DO 40 L=1,2
    DO 40 K=1,2
        B(K,L,I)=B(K,1,IM)*A(1,L)+B(K,2,IM)*A(2,L)
40 CONTINUE

    DO 50 K=1,2
        UP(K,I)=UP(K,IM)+B(K,1,I)*G(1,I)+B(K,2,I)*G(2,I)
50 CONTINUE

    DO 60 L=1,2
    DO 60 K=1,2
        A(K,L)=H(K,L,I)-INVA(K,L,IM)
60 CONTINUE

    DETA(I)=A(1,1)*A(2,2)-A(1,2)**2
    IF(DETA(I).EQ.0.D0) THEN
        NONINV=I
        RETURN
    END IF
    TRA(I)=A(1,1)+A(2,2)
```

```

        INVA(1,1,I)=A(2,2)/DETA(I)
        INVA(2,1,I)=-A(2,1)/DETA(I)
        INVA(1,2,I)=INVA(2,1,I)
        INVA(2,2,I)=A(1,1)/DETA(I)

        DO 70 L=1,2
        DO 70 K=1,2
                INVB(K,L,I)=INVA(K,1,IM)*INVB(1,L,IM)+INVA(K,2,IM)*INVB(2,L,IM)
70    CONTINUE

        DO 80 L=1,2
        DO 80 K=1,2
                C(K,L,NM)=INVB(1,K,NM)*INVA(1,L,NM)+INVB(2,K,NM)*INVA(2,L,NM)
80    CONTINUE

        DN(1,NM)=0.D0
        DN(2,NM)=0.D0

        DO 90 K=1,2
                DELX(K,NM)=C(1,K,NM)*UP(1,NM)+C(2,K,NM)*UP(2,NM)
90    CONTINUE

        DO 120 I=N-2,1,-1
                IP=I+1

                DO 100 L=1,2
                DO 100 K=1,2
                        C(K,L,I)=C(K,1,IP)*INVA(1,L,I)+C(K,2,IP)*INVA(2,L,I)+INVB(L,K,IP)
100   CONTINUE

                DO 110 K=1,2

```

```

          DN(K,I)=DN(K,IP)+C(K,1,IP)*G(1,IP)+C(K,2,IP)*G(2,IP)
110      CONTINUE

          DO 120 K=1,2
          DELX(K,I)=C(1,K,I)*UP(1,I)+C(2,K,I)*UP(2,I)+
1          B(1,K,I)*DN(1,I)+B(2,K,I)*DN(2,I)
120      CONTINUE

          CRIT1=DNORM(G,N)
          CRIT2=DNORM(DELX,N)

          IF(CRIT2.LE.EPS) THEN

              DO 130 L=1,2
              DO 130 K=1,2
                  TEMP(K,L)=INVA(K,1,NM)*INVB(1,L,NM)+INVA(K,2,NM)*INVB(2,L,NM)
130          CONTINUE

              DO 140 L=1,2
              DO 140 K=1,2
                  A(K,L)=H(K,L,N)-INVA(K,L,NM)-C(K,L,1)-TEMP(K,L)-TEMP(L,K)
140          CONTINUE

              DETA(N)=A(1,1)*A(2,2)-A(1,2)**2
              IF (DETA(N).EQ.0.D0) THEN
                  NONINV=N
                  RETURN
              END IF

              TRA(N)=A(1,1)+A(2,2)
              IERR=1
              RETURN
          END IF

```



```

IMPLICIT REAL*8 (A-H,O-Z)
COMMON/PARAM/TPI,AH0,AK0,AK02,AH,AK,AK2
COMMON/PERIOD/AM1,AM2
COMMON/SYMM/SR1,SR2,TSON

N1=(N-1)/2
ONETWO=FLOAT(N-2*N1)
C ONETWO=1 FOR ODD N, =2 FOR EVEN N
FACTOR=(1.E0-TSON)/ONETWO+TSON*ONETWO/2.E0

X1=(SR1+TSON*Y1)/2.D0
X2=(SR2+TSON*Y2)/2.D0

YT1=Y1
YT2=Y2

DO 10 NP=1,N1
    COUPLE=AH*DSIN(TPI*(X1+X2))
    YT1=YT1-AK*DSIN(TPI*X1)-COUPLE
    YT2=YT2-AK2*DSIN(TPI*X2)-COUPLE
    X1=X1+YT1
    X2=X2+YT2
10 CONTINUE

    COUPLE=AH*DSIN(TPI*(X1+X2))
    YT1=YT1-AK*DSIN(TPI*X1)-COUPLE
    YT2=YT2-AK2*DSIN(TPI*X2)-COUPLE

F1=YT1-FACTOR*(AM1+SR1-2.E0*X1)
F2=YT2-FACTOR*(AM2+SR2-2.E0*X2)

RETURN
END

```

```
SUBROUTINE HESS(X,H,N)
```

```
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
C
```

```
C RECEIVES ANGLES X RETURNS THE DIAGONAL BLOCKS IN H
```

```
C OF THE HESSIAN MATRIX OF THE ACTION FUNCTION
```

```
C
```

```
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
```

```
REAL*8 X(2,N),H(2,2,N)
```

```
REAL*8 COUPLE,TPI,AH0,AK0,AK02,AH,AK,AK2
```

```
COMMON/PARAM/TPI,AH0,AK0,AK02,AH,AK,AK2
```

```
DO 10 I=1,N
```

```
    COUPLE=AH0*DCOS(TPI*(X(1,I)+X(2,I)))
```

```
    H(1,1,I)=2.D0-AK0*DCOS(TPI*X(1,I))-COUPLE
```

```
    H(2,1,I)=-COUPLE
```

```
    H(1,2,I)=-COUPLE
```

```
    H(2,2,I)=2.D0-AK02*DCOS(TPI*X(2,I))-COUPLE
```

```
10 CONTINUE
```

```
RETURN
```

```
END
```

SUBROUTINE GRAD(X,G,N)

```

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C   RECEIVES ANGLES AND RETURNS THE GRADIENT OF ACTION FUNCTION
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

```

PARAMETER (NMAX=13580)

REAL*8 X(2,N),G(2,N-1),P(2,NMAX+1),COUPLE

REAL*8 TPI,AH0,AK0,AK02,AH,AK,AK2,AM1,AM2

COMMON/PARAM/TPI,AH0,AK0,AK02,AH,AK,AK2

COMMON/PERIOD/AM1,AM2

P(1,1)=X(1,1)-X(1,N)

P(2,1)=X(2,1)-X(2,N)

DO 10 I=2,N-1

P(1,I)=X(1,I)-X(1,I-1)

P(2,I)=X(2,I)-X(2,I-1)

10 CONTINUE

P(1,N)=AM1-X(1,N-1)+X(1,N)

P(2,N)=AM2-X(2,N-1)+X(2,N)

DO 20 I=1,N-1

COUPLE=AH*DSIN(TPI*(X(1,I)+X(2,I)))

G(1,I)=P(1,I)-P(1,I+1)-AK*DSIN(TPI*X(1,I))-COUPLE

G(2,I)=P(2,I)-P(2,I+1)-AK2*DSIN(TPI*X(2,I))-COUPLE

20 CONTINUE

RETURN

END

List of References

- R. Abraham and J. Marsden, Foundations of Mechanics, (Benjamin/Cummings, Reading, 1978), p. 308.
- T.M. Antonsen and E. Ott, "Diffusion Coefficient for Ions in the presence of a Coherent Lower Hybrid Wave," *Phys. Fluids* **24**, 1635(1981).
- V.I. Arnold, Mathematical Methods of Classical Mechanics, (Springer-Verlag, New York, 1978).
- K.E. Atkinson, An introduction to numerical analysis, (John Wiley & Sons, New York, 1978).
- S. Aubry, "The twist map, the extended Frenkel-Kontorova model and devil's staircase," *Physica* **7D**, 240 (1983).
- S. Aubry and P.Y. Le Daeron, "The discrete Frenkel-Kontorova model and its extensions," *Physica* **8D**, 381 (1983).
- R. Balescu, Equilibrium and Nonequilibrium Statistical Mechanics, (John Wiley & Sons, New York, 1975).
- L. Bernstein, The Jacobi-Perron Algorithm: Its Theory and Application (Springer-Verlag, Heidelberg, 1971).

- D. Bernstein and A. Katok, "Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians," *Invent. math.* **88**, 225 (1987).
- M.V. Berry, "Regular and Irregular Motion", in Topics in Nonlinear Dynamics, AIP Conf. Proc. **46**(1978), ed. S. Jorna.
- J. Binney and S. Tremaine, Galactic Dynamics. (Princeton University Press, Princeton, 1987).
- A.J. Brentjes, Multidimensional Continued Fraction Algorithms (Mathematisch Centrum, Amsterdam, 1981).
- R. Broucke, "Stability of periodic orbits in the elliptic three-body problem," *AIAA Journal* **7**, 1003 (1969).
- C.G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Math. Comp.* **19**, 577 (1965).
- C.G. Broyden, J.E. Dennis Jr. and J.J. More, "On the local and superlinear convergence of Quasi-Newton methods," *J. Inst. Math. Appl.* **12**, 223 (1973).
- J. R. Cary and J.D. Meiss, "Rigorously Diffusive Deterministic Map," *Phys. Rev* **24A**, 2664(1981).
- J. R. Cary, J. D. Meiss, and A. Bhattacharjee, "Statistical Characterization of Periodic Measure-Preserving Mappings," *Physical Review* **A23**, 2744(1981).

- J.R. Cary and R.G. Littlejohn, "Noncanonical Hamiltonian Mechanics and Its Application to Magnetic Field Line Flow," *Annals. Phys.* **151**, 1(1983).
- Q. Chen, J.D. Meiss and I.C. Percival, "Orbit extension method for finding unstable orbits," *Physica* **29D**, 143 (1987).
- Q. Chen, "Area as a devil's staircase in twist maps," *Phys. Lett. A* **123**, 444(1987).
- Q. Chen, R.S. MacKay, and J.D. Meiss, "Cantori for symplectic maps," (1989) submitted for publication.
- B.V. Chirikov, "A universal instability of many-dimensional oscillator systems," *Phys. Rep.* **52**, 265 (1979).
- G. Contopoulos, "Qualitative changes in 3-dimensional dynamical systems," *Astron. Astrophys.* **161**, 244(1986).
- I. Dana, N.W. Murray, and I.C. Percival, "Resonances and Diffusion in Periodic Hamiltonian Maps," *Phys. Rev. Lett.* **62**, 233(1989).
- E. Davoust and R. Broucke, "A manifold of periodic orbits in the planar general three-body problem with equal masses," *Astron. Astrophys.* **112**, 305(1982).
- J.E. Dennis Jr. and J.J. More, "Quasi-Newton methods, motivation and theory," *SIAM Review* **19**, 46 (1977).
- S.F. Dermott and C.D. Murray, "Nature of the Kirkwood gaps in the asteroid belt," *Nature* **301**, 201(1983).

- A.J. Dragt, "Lectures on nonlinear orbit dynamics," AIP Conf. Proc. **87** (1982); in Nonlinear Dynamics and the Beam-Beam Interaction, Brookhaven, 1979, AIP Conf. Proc #57, M.Month and J.Herrera (eds.), p. 236 (AIP, New York, 1979).
- C. Froeschlé, "On the number of isolating integrals in systems with three degrees of freedom," *Astrophys. Space Sci.* **14**, 110 (1971).
- C. Froeschlé, "Numerical study of a four-dimensional mapping," *Astron. Astrophys.* **16**, 172 (1972).
- C. Froeschlé and J.-P. Scheideker, "Numerical study of a four-dimensional mapping. II" *Astron. Astrophys.* **22**, 431 (1973).
- V.L. Ginzburg and M.B. Sevryuk, the result is a consequence of the symplectic Bochner Theorem, in a letter to R.S. MacKay dated Jan. 18, 1988.
- J.M. Greene, "A method for determining a stochastic transition," *J. Math. Phys.* **20**, 1183 (1979).
- J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. (Springer-Verlag, New York, 1983).
- Y. Hagihara, Celestial Mechanics. (MIT, Cambridge, Massachusetts, 1970).
- J.D. Hanson, J.R. Cary, and J.D. Meiss, "Algebraic Decay in Self-Similar Markov Chains," *J. Stat. Phys.* **39**, 327(1985).

G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers (Clarendon Press, Oxford, 1954).

T. Hatori, T. Kamimura, Y.H. Ichikawa, "Turbulent Diffusion for the Radial Twist Map", *Physica* **14D**, 193(1985) .

G.A. Hedlund, "Geodesics on a two-dimensional Riemannian manifold with periodic coefficients," *Ann. Math.* **33**, 719(1932)

R.H.G. Helleman, "Exact Results for Some Linear and Nonlinear Beam-Beam Effects," in Nonlinear Dynamics and the Beam-Beam Interaction, AIP Conf. Proc. **57**(1979).

J.E.Howard, A.J. Lichtenberg, M.A. Lieberman, and R.H. Cohen, "Four-dimensional mapping model for two-frequency electron cyclotron resonance heating," *Physica* **20D**, 259(1986).

J.E. Howard and R.S. MacKay, "Linear stability of symplectic maps," *J. Math. Phys.* **28**, 1036(1987).

Y. H. Ichikawa, T. Kamimura, and T. Hatori, "Stochastic Diffusion in the Standard Map", *Phys. Rev. A* (1987) .

K. Kaneko and R.J. Bagley, "Arnold diffusion, ergodicity and intermittency in a coupled standard mapping," *Phys. Lett.* **110A**, 435 (1985).

C.F.F. Karney, A. B. Rechester, and R.B. White, "Effect of Noise on the Standard Mapping", *Physica* **4D**, 425(1982).

C.F.F. Karney, "Long Time Correlations in the Stochastic Regime", *Physica* **8D**, 360(1983) .

A.Y. Khinchin, Continued Fractions (Univ. of Chicago Press, Chicago, 1964).

S. Kim and S. Ostlund, "Simultaneous rational approximations in the study of dynamical systems," *Phys. Rev.* **A35** (1987).

C. Kueny and P.J. Morrison, in preparation (1989).

A.J. Lichtenberg and M.A. Lieberman, Regular and Stochastic Motion (Springer-Verlag, New York,1983).

A.J. Lichtenberg, M. A. Lieberman and N.W. Murray, "The Effect of Quasi-Accelerator Modes on Diffusion", *Physica* **28D**, 371(1987).

R.G. Littlejohn, "Variational Principles of Guiding Center Motion", *J. Plasma Phys.* **29**, 111(1983).

R.S. MacKay, "Renormalization in area preserving maps", Ph.D. Thesis, (Univ. Microfilms, Ann Arbor, 1982).

R.S. MacKay, "A renormalization approach to invariant circles in area-preserving maps," *Physica* **7D**, 283 (1983).

R.S. MacKay and J.D. Meiss, "Linear stability of periodic orbits in Lagrangian systems," *Phys. Lett.* **98A**, 92 (1983).

R.S. MacKay, J.D. Meiss, and I.C. Percival, "Stochasticity and transport in Hamiltonian systems," *Phys. Rev. Lett.* **52**, 697(1984).

R.S. MacKay, J.D. Meiss, and I.C. Percival, "Transport in Hamiltonian systems," *Physica* **13D**, 55(1984).

R.S. MacKay, J.D. Meiss and I.C. Percival, "Resonances in area-preserving maps," *Physica* **27D**, 1(1987).

R.S. MacKay, J.D. Meiss, and J. Stark, "Converse KAM theory for symplectic twist maps," University of Warwick preprint(1988).

J.M. Mao, I.I. Satija, and Bambi Hu, "Period doubling in four-dimensional symplectic maps," *Phys. Rev.* **A34**, 4325 (1986).

J.M. Mao and R.H.G. Helleman, "New Feigenbaum constants for four-dimensional volume-preserving symmetric maps," *Phys. Rev.* **A35**, 1847(1987).

J.M. Mao and R.H.G. Helleman, "Non-symmetric four-dimensional volume-preserving maps: Universality classes of period doubling," *Phys. Rev.* **A37**, 3475 (1987).

J.N. Mather, "Existence of quasi-periodic orbits for twist homeomorphisms of the annulus," *Topology* **21**, 457(1982).

J.N. Mather, "A Criterion for the Non-existence of Invariant Circles," *Publ. Math. IHES* **63**, 153(1986).

J. D. Meiss, J. R. Cary, C. Grebogi, J. D. Crawford, A. N. Kaufman and H.D.I. Abarbanel, "Correlations of Periodic Area-Preserving Maps," *Physica* **6D**, 375(1983).

J.D. Meiss and E. Ott, "Markov-Tree Model of Intrinsic Transport in Hamiltonian Systems," Phys. Rev. Lett. **55**, 2741(1985).

J.D. Meiss, "Class Renormalization: Islands around Islands," Phys. Rev. **A34**, 2375(1986).

B. Mestel and I.C. Percival, "Newton method for highly unstable orbits," Physica **24D**, 172 (1987).

J. Milnor, Morse Theory, Annals of Mathematical Studies **51**, (Princeton Univ. Press, 1963).

M. Month and J.C. Herrera, Nonlinear Dynamics and the Beam-Beam Interaction, AIP Conf. Proc. **57**(1979).

J. Moser, "New aspects in the theory of stability of Hamiltonian systems," Comm. pure and applied Math., vol **XI**, 81(1958).

J. Moser, Stable and random motions in dynamical systems, (Princeton University Press, Princeton, 1973).

N.W. Murray, M.A. Lieberman, and A.J. Lichtenberg, "Corrections to Quasilinear Diffusion in Area Preserving Maps", Phys. Rev. **32A**, 2413(1985).

N.N. Nekhoroshev, "An Exponential Estimate for the Time of Stability of Nearly-Integrable Hamiltonian Systems," Russ. Math. Surveys **32:6**, 1 (1965).

J.M. Ortega and W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables, (Academic Press, New York, 1970).

I.C. Percival, "Variational Principles for Invariant Tori and Cantori," in Nonlinear Dynamics and the Beam-Beam Interaction, AIP Conf. Proc. **57**(1979).

E. Polak, Computational method in optimization, (Academic Press, New York, 1971).

W.H. Press et. al., Numerical Recipes, (Cambridge Univ. Press, Cambridge, 1986).

A.B. Rechester, and R.B. White, "Calculation of Turbulent Diffusion for the Chirikov-Taylor Model", Phys. Rev. Lett. **44**, 1586(1980).

A.B. Rechester, M.N. Rosenbluth and R.B. White, "Fourier-space Paths Applied to the Calculation of Diffusion for the Chirikov-Taylor Model", Phys. Rev. **23A**, 2664(1981) .

R. Rimmer, "Symmetry and bifurcation of fixed points of area preserving maps," Journal of Diff. Eq. **29**, 329 (1978).

V. Rom-Kedar and S. Wiggins, "Transport in Two dimensional Maps", preprint (1989).

W. Sierpinski, "Sur une courbe cantorienne qui contient une image biunivoque et continue de toute courbe donnee," Compte Rendus Acad. Sci. Paris **162**, 629 (1916).

J. L. Tennyson, M.A. Lieberman and A.J. Lichtenberg, "Diffusion in Near-Integrable Hamiltonian Systems with Three Degrees of Freedom," in Nonlinear Dynamics and the Beam-Beam Interaction, M. Month and J.C. Herrera (eds.) AIP Conf. Proc. **57**, 272(1979).

D. K. Umberger and J. D. Farmer, "Fat Fractals on the Energy Surface", Phys. Rev. Lett. **55**, 661(1985).

R. DeVogelaere, in Contributions to the Theory of Nonlinear Oscillations, (Princeton Univ. Press, Princeton, 1958), S. Lefschetz (ed), vol **IV**, p. 53.

E.T. Whittaker, A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, (Cambridge Univ. Press, Cambridge, 1964).

S. Wiggins, Global Bifurcations and Chaos, (Springer-Verlag, New York, 1988).

J. Wisdom, "The origin of the Kirkwood gaps: a mapping for asteroidal motion near the 3/1 commensurability," Astron J. **87**, 577 (1982).

VITA

Hyungtae Kook was born in Seoul, Korea on April 20, 1958, the son of Kiho Kook and Mikyung Lee. After completing his work at Yang Chung High School in Seoul, Korea in 1977, he entered Seoul National University in Seoul, Korea, where he received the degree of Bachelor of Science in Physics in February, 1981. In March, 1981, he entered The Graduate School of Seoul National University. During the following years he was employed as a teaching assistant in department of Physics. He received the degree of Master of Science in Physics from Seoul National University in February, 1983. During the following year, he served in the army. In September, 1984, he entered The Graduate School of The University of Minnesota and was employed as a teaching assistant in department of Physics. In January, 1985, he transferred to The University of Texas at Austin, Texas. Since June, 1985, he has been employed as a research assistant by The Institute for Fusion Studies at The University of Texas at Austin.

Permanent address: 326-4 Pungnab-dong
 Songpa-gu
 Seoul, Korea

This dissertation was set with Microsoft® Word by the author.